Some tests for mean vectors with monotone missing data

（単調欠測データをもつ平均ベクトルに対するいくつかの検定）

Ayaka Yagi

Department of Applied Mathematics, Faculty of Science

Tokyo University of Science

1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

Acknowledgments

The author would like to express her sincere gratitude to Professor Takashi Seo for his constant encouragement, advice, and support of my Ph.D. study and research. The author appreciates all his contributions of time, ideas and discussions. Without his guidance and persistent help, this dissertation would not have been possible. The author is grateful to Professor Zofia Hanusz, University of Life Sciences, Emeritus Professor Muni S. Srivastava, University of Toronto, and Emeritus Professor Yasunori Fujikoshi, Hiroshima University, who have supported her during her study. The author's research was partly supported by a Grant-in-Aid for JSPS Fellows (15J00414). Finally, the author would like to thank her parents and sister. They were always supporting and encouraging her with their best wishes.

Contents

Chapter 1

Introduction

In this study, we consider some tests for mean vectors when each data set has a monotone missing data pattern. We deal with the Hotelling's T^2-type test statistics and the likelihood ratio test (LRT) statistics. First we consider the Hotelling's T^2-type test statistics for one-sample, two-sample, and multi-sample problems. We give simplified T^2-type test statistics and approximate upper percentiles of the statistics. Using the approximations, approximate simultaneous confidence intervals for the mean components are obtained. We also consider approximate simultaneous confidence intervals for pairwise comparisons among mean vectors and comparisons with a control are obtained using Bonferroni's approximation procedure. The accuracy and asymptotic behavior of the approximations are investigated by Monte Carlo simulation.

Second, as for LRT, we give the likelihood ratio (LR) for one-sample problem. And then, we derive modified test (MT) and modified likelihood ratio test (MLRT) statistics by using decomposition of the LR. In addition, we deal with the LRT, MT, and MLRT statistics to test the equality of mean vectors in a one-way MANOVA when each dataset has a monotone pattern of missing observations. The accuracy of the approximation for the chi-square distribution is investigated using a Monte Carlo simulation. Throughout this paper, we assume that the data are missing completely at random (MCAR).

The remainder of this paper is as follows. In Chapter 2, we discuss the simplified T^2-type test statistics in one-sample problem with three-step and general step monotone missing data. For the one-sample problem with k-step monotone missing data, Jinadasa and Tracy (1992) obtained closed form expressions for the maximum likelihood estimators (MLEs) of the mean vector and the covariance matrix of the multivariate normal distribution. In particular, Anderson (1957) and Anderson and Olkin (1985) considered a two-step monotone missing data pattern. Kanda and Fujikoshi (1998) discussed the properties of the MLEs in the case of k-step monotone missing data using the conditional approach.

Tests for a mean vector with monotone missing data have been discussed by many authors. For discussions related to the statistics based on the Hotelling's T^2 statistic in one-sample problem, see Krishnamoorthy and Pannala (1999), Chang and Richards (2009), Seko, Yamazaki and Seo (2012), Yagi and Seo (2014), and Kawasaki and Seo (2016), among others. For example, Chang and Richards (2009) gave the Hotelling's T^2-type statistic and its some properties, and Seko et al. (2012) proposed approximate upper percentiles of the Hotelling's T^2-type statistic with two-step monotone missing data. Krishnamoorthy and Pannala (1999) derived a pivotal quantity, similar to the Hotelling's T^2 statistic and gave an approximation to the null distribution of the statistic with monotone missing data. In Section 2.1, we give the simplified T^2-type statistic and approximate upper percentiles of its null distribution with three-step monotone missing data. This is a summary of Yagi and Seo (2014). A generalization of Section 2.1 is given in Section 2.2, which is a summary of Yagi and Seo (2017).

In Chapter 3, as an extension of Chapter 2 to the two-sample or multi-sample problem, the simplified T^2-type test statistic in two-sample or multi-sample problem with three-step and general step monotone missing data are given in Section 3.1 and Section 3.2, respectively. In the two-sample problem, for two-step monotone missing data, Seko, Kawasaki and Seo (2011) gave the Hotelling's T^2-type statistic and their approximate upper percentiles for the null distributions, using the linear interpolation approximation. Yu, Krishnamoorthy and Pannala (2006) proposed a pivotal quantity, similar to the Hotelling's T^2 statistic and discussed the approximate distributions for the statistic. Indeed, we note that the simplified Hotelling's T^2-type statistics coincide with the pivotal quantities similar to the Hotelling's T^2 statistic in Krishnamoorthy and Pannala (1999) and Yu et al. (2006). We also present simultaneous confidence intervals for multiple comparisons among mean vectors under the two-sample and multi-sample problems. These are summaries of Yagi and Seo (2015b) and Yagi and Seo (2017).

On the other hand, since the simplified T^2-type test is not the LR test in the case of the data with monotone missing data pattern unlike the complete data case, the LR test for one-sample problem is discussed in Chapter 4. The LR test for a mean vector is also discussed by many authors. Krishnamoorthy and Pannala (1998) gave the decomposition of LR and provided comparisons with several approximation procedures. Then, Seko et al. (2012) discussed the LRT statistic and the linear interpolation approximation to the null distribution in the two-step monotone missing case. For a discussion on developing estimation and testing procedures for the mean vector and the scale matrix of the elliptical distributions with monotone missing data, see Batsidis and Zografos (2006). In this chapter, the MT and MLRT statistics of the one-sample test for a normal mean vector with monotone missing data are obtained, which is a summary of Yagi, Seo and Srivastava (2017a). We present that the LR for the one-sample test of the mean vector with monotone missing data can be expressed as the products of the LR of the test for a mean vector and those of subvector, and we derive the asymptotic expansion by the perturbation method. A related discussion of a test for a subvector and a decomposition with complete data was given by Siotani et al. (1985). Under nonnormality with complete data, Gupta, Xu and Fujikoshi (2006) discussed the asymptotic expansion of the distribution of Rao's U-statistic, which is proposed as test for a subvector or additional information. For the simultaneous testing of the mean vector and the covariance matrix with monotone missing data, see Hao and Krishnamoorthy (2001), Tsukada (2014), Hosoya and Seo (2015, 2016), among others.

Finally, in Chapter 5, as an extension of Chapter 4 to the multi-sample problem, the LRT, MT and MLRT statistics in a one-way MANOVA is considered. This is a summary of Yagi, Seo and Srivastava (2017b).

Chapter 2

A test for a mean vector

In this chapter, we consider the problem of testing for a mean vector and simultaneous confidence intervals when the data have three-step (Section 2.1) or general k-step (Section 2.2) monotone pattern of missing observations. The MLEs of the mean vector and the covariance matrix with a three-step or general step monotone missing data pattern are presented. We propose an approximate upper percentile of simplified T^2-type statistic to test a mean vector. Further, we obtain the approximate simultaneous confidence intervals for any and all linear compounds of the mean, and testing the equality of mean components is discussed. Finally, the accuracy of the approximation is investigated by Monte Carlo simulation and a numerical example is given to illustrate the method.

The case in which the missing observations are of the monotone type has been considered by several authors, including Rao (1956), Anderson (1957), and Bhargava (1962). Jinadasa and Tracy (1992) obtained closed form expressions for the MLEs of the mean vector and the covariance matrix of the multivariate normal distribution in the case of the k-step monotone missing data. Kanda and Fujikoshi (1998) discussed the distribution of the MLEs in the case of the k-step monotone missing data.

In this chapter, we consider the problem of testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs. $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ when the data have monotone pattern of missing observations and $\boldsymbol{\mu}_0$ is known. In the case of a two-step monotone missing data, Chang and Richards (2009) and Seko et al. (2012) derived a Hotelling's T^2-type statistic and some properties. For the case of a k-step monotone missing data pattern ($k \geq 3$), Krishnamoorthy and Pannala (1999) derived a simplified T^2-type statistic since the usual T^2-type statistic becomes very complicated to derive the exact covariance matrix for MLE of the mean vector (see, Kanda and Fujikoshi, 1998, p.185). They proposed an approximation to the upper percentile.

In Section 2.1, using other definitions, we give the simplified T^2-type statistic and propose its approximate upper percentile in the case of a three-step monotone missing data. Our approximation procedure is essentially based on that given in Seko et al. (2012). The related discussion is given by Hao and Krishnamoorthy (2001), Little and Rubin (2002), Chang and Richards (2009), among others.

Section 2.1 is organized in the following way. In Section 2.1.1, we present the MLEs of the mean vector and the covariance matrix with a three-step monotone missing data using the notations and derivation used by Jinadasa and Tracy (1992). These results are simple and useful in order to derive a simplified T^2-type statistic and the covariance for the MLE of the mean vector. In Section 2.1.2, we present the simplified T^2-type statistic for testing the mean vector and its approximate upper percentile. In Section 2.1.3, the approximate simultaneous confidence intervals for any and all linear compounds of the mean are obtained. Further, we discuss testing the equality of mean components. In Sections 2.1.4 and 2.1.5, we give some simulation results and a numerical example, respectively. Indeed, Section 2.1 is organized based on Yagi and Seo (2014).

On the other hand, in Section 2.2, which is summarized using the parts of Yagi and Seo (2017), we consider the distribution of the simplified T^2-type test statistic in the case of general k-step monotone missing data. That is, we give an extension of three-step case in Section 2.1. In Section 2.2.1, some preliminary notations are presented, and the MLEs of the mean vector, and the common covariance matrix are obtained in the case of k-step monotone missing data. In Section 2.2.2, we give the simplified T^2-type statistic to test for a mean vector and their approximate upper percentiles. Further, we give approximate simultaneous confidence intervals for any and all linear compounds of the mean and the testing equality of mean components. In Section 2.2.3, we give the simulation results. We present the numerical comparisons of our linear interpolation approximation with the approximation by Krishnamoorthy and Pannala (1999). Finally, we state our conclusions in Section 2.3.

2.1 Three-step monotone missing data

In this section, we propose a simplified T^2-type statistic and its approximate upper percentile in the case of a three-step monotone missing data, similar to that in the case of a two-step monotone missing data. We deal with the problem of testing for mean vector with a three-step monotone missing data:

$$\begin{pmatrix}
x_{11} & \cdots & x_{1p_3} & x_{1,p_3+1} & \cdots & x_{1p_2} & x_{1,p_2+1} & \cdots & x_{1p_1} \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{n_1 1} & \cdots & x_{n_1 p_3} & x_{n_1,p_3+1} & \cdots & x_{n_1 p_2} & x_{n_1,p_2+1} & \cdots & x_{n_1 p_1} \\
x_{n_1+1,1} & \cdots & x_{n_1+1,p_3} & x_{n_1+1,p_3+1} & \cdots & x_{n_1+1,p_2} & * & \cdots & * \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{n_1+n_2,1} & \cdots & x_{n_1+n_2,p_3} & x_{n_1+n_2,p_3+1} & \cdots & x_{n_1+n_2,p_2} & * & \cdots & * \\
x_{n_1+n_2+1,1} & \cdots & x_{n_1+n_2+1,p_3} & * & \cdots & * & * & \cdots & * \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{n_1+n_2+n_3,1} & \cdots & x_{n_1+n_2+n_3,p_3} & * & \cdots & * & * & \cdots & *
\end{pmatrix},$$

where $p = p_1 > p_2 > p_3 > 0$, $n_1 > p$, and "$*$" indicates a missing observation. That is, we have a complete data set for n_1 observations with p_1 dimensions and two incomplete data sets which have n_2 observations with p_2 dimensions and n_3 observations with p_3 dimensions. Further, let $\boldsymbol{x}$ be distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let $\boldsymbol{x}_i = (\boldsymbol{x})_i$ be the vector of the first p_i elements of $\boldsymbol{x}$. Then, $\boldsymbol{x}_i (= (x_1, x_2, \ldots, x_{p_i})')$ is distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2, 3$, where $\boldsymbol{\mu}_i = (\boldsymbol{\mu})_i = (\mu_1, \mu_2, \ldots, \mu_{p_i})'$ and $\boldsymbol{\Sigma}_i$ is the principal submatrix of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}_1)$ of order $p_i \times p_i$.

Let $(\boldsymbol{\Sigma}_i)_j$ be the principal submatrix of $\boldsymbol{\Sigma}_i$ of order $p_j \times p_j$, $1 \le i < j \le 3$. We define

$$\boldsymbol{\Sigma}_{i+1} = (\boldsymbol{\Sigma}_1)_{i+1}, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{i+1} & \boldsymbol{\Sigma}_{i+1,2} \\ \boldsymbol{\Sigma}'_{i+1,2} & \boldsymbol{\Sigma}_{i+1,3} \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_i = \begin{pmatrix} \boldsymbol{\Sigma}_{i+1} & \boldsymbol{\Sigma}_{(i,2)} \\ \boldsymbol{\Sigma}'_{(i,2)} & \boldsymbol{\Sigma}_{(i,3)} \end{pmatrix}, \quad i = 1, 2.$$

For example, we can express $\boldsymbol{\Sigma}_1$ as

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \overbrace{\boldsymbol{\Sigma}_2}^{p_2} & \overbrace{\boldsymbol{\Sigma}_{22}}^{p_1-p_2} \\ \boldsymbol{\Sigma}'_{22} & \boldsymbol{\Sigma}_{23} \end{pmatrix} \begin{matrix} \}p_2 \\ \}p_1 - p_2 \end{matrix} \quad \text{or} \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} \overbrace{\boldsymbol{\Sigma}_3}^{p_3} & \overbrace{\boldsymbol{\Sigma}_{32}}^{p_1-p_3} \\ \boldsymbol{\Sigma}'_{32} & \boldsymbol{\Sigma}_{33} \end{pmatrix} \begin{matrix} \}p_3 \\ \}p_1 - p_3 \end{matrix} .$$

Also, we have

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \overbrace{\boldsymbol{\Sigma}_3}^{p_3} & \overbrace{\boldsymbol{\Sigma}_{(2,2)}}^{p_2-p_3} \\ \boldsymbol{\Sigma}'_{(2,2)} & \boldsymbol{\Sigma}_{(2,3)} \end{pmatrix} \begin{matrix} \}p_3 \\ \}p_2 - p_3 \end{matrix} ,$$

where $\boldsymbol{\Sigma}_2$ is the upper left submatrix of $\boldsymbol{\Sigma}_1$.

If $\boldsymbol{x}_{ij}$ denotes the jth observation on $\boldsymbol{x}_i$, then the three-step monotone missing data set is of the form

$$\left(\begin{array}{ccccc} \multicolumn{5}{c}{\boldsymbol{x}'_{11}} \\ \multicolumn{5}{c}{\vdots} \\ \multicolumn{5}{c}{\boldsymbol{x}'_{1n_1}} \\ \hdashline \multicolumn{3}{c}{\boldsymbol{x}'_{21}} & * & \cdots * \\ \multicolumn{3}{c}{\vdots} & \vdots & \vdots \\ \multicolumn{3}{c}{\boldsymbol{x}'_{2n_2}} & * & \cdots * \\ \hdashline \boldsymbol{x}'_{31} & * \cdots * & & * & \cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{x}'_{3n_3} & * \cdots * & & * & \cdots * \end{array} \right),$$

Such a data set is called a three-step monotone missing data pattern. For a k-step monotone sample or a k-step monotone missing data pattern, see Bhargava (1962), Srivastava and Carter (1983), Little and Rubin (2002), Srivastava (2002), among others.

2.1.1 MLEs of the mean vector and the covariance matrix

Let the MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ be denoted by $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$, respectively. If the data have three-step monotone pattern missing observations, then we have the following theorem based on the derivation of Jinadasa and Tracy (1992).

Theorem 2.1 (Yagi and Seo, 2014) *If the data have three-step monotone pattern missing observations, then the MLE of the mean vector is given by*

$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \widehat{\boldsymbol{T}}_2 \boldsymbol{d}_2 + \widehat{\boldsymbol{T}}_2 \widehat{\boldsymbol{T}}_3 \boldsymbol{d}_3,$$

where

$$\overline{\boldsymbol{x}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\boldsymbol{x}_{ij},\ i = 1,2,3,$$

$$\boldsymbol{d}_2 = \frac{n_2}{N_3}\left\{\overline{\boldsymbol{x}}_2 - (\overline{\boldsymbol{x}}_1)_2\right\},\ \boldsymbol{d}_3 = \frac{n_3}{N_4}\left[\overline{\boldsymbol{x}}_3 - \frac{1}{N_3}\left\{n_1(\overline{\boldsymbol{x}}_1)_3 + n_2(\overline{\boldsymbol{x}}_2)_3\right\}\right],$$

$$\widehat{\boldsymbol{T}}_2 = \begin{pmatrix}\boldsymbol{I}_{p_2}\\ \widehat{\boldsymbol{\Sigma}}'_{(1,2)}\widehat{\boldsymbol{\Sigma}}_2^{-1}\end{pmatrix},\ \widehat{\boldsymbol{T}}_3 = \begin{pmatrix}\boldsymbol{I}_{p_3}\\ \widehat{\boldsymbol{\Sigma}}'_{(2,2)}\widehat{\boldsymbol{\Sigma}}_3^{-1}\end{pmatrix},\quad N_{i+1} = \sum_{j=1}^{i} n_j,\ i = 1,2,3,$$

and then, the MLE of the covariance matrix is given by

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}} = {} & \frac{1}{N_2}\boldsymbol{E}_1 + \frac{1}{N_3}\boldsymbol{G}_2\left[\boldsymbol{E}_2 + \frac{N_2N_3}{n_2}\boldsymbol{d}_2\boldsymbol{d}_2' - \frac{n_2}{N_2}\boldsymbol{L}_{11}\right]\boldsymbol{G}_2' \\ & + \frac{1}{N_4}\boldsymbol{G}_2\boldsymbol{G}_3\left[\boldsymbol{E}_3 + \frac{N_3N_4}{n_3}\boldsymbol{d}_3\boldsymbol{d}_3' - \frac{n_3}{N_3}\boldsymbol{L}_{21}\right]\boldsymbol{G}_3'\boldsymbol{G}_2',\end{aligned}$$

where

$$\boldsymbol{E}_i = \sum_{j=1}^{n_i}(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)',\ i = 1,2,3,\quad \boldsymbol{G}_2 = \begin{pmatrix}\boldsymbol{I}_{p_2}\\ \boldsymbol{L}_{12}'\boldsymbol{L}_{11}^{-1}\end{pmatrix},\ \boldsymbol{G}_3 = \begin{pmatrix}\boldsymbol{I}_{p_3}\\ \boldsymbol{L}_{22}'\boldsymbol{L}_{21}^{-1}\end{pmatrix},$$

$$\boldsymbol{L}_1 = \boldsymbol{E}_1,\ \boldsymbol{L}_2 = \boldsymbol{L}_{11} + \boldsymbol{E}_2 + \frac{N_2N_3}{n_2}\boldsymbol{d}_2\boldsymbol{d}_2',\quad \boldsymbol{L}_i = \begin{pmatrix}\boldsymbol{L}_{i1} & \boldsymbol{L}_{i2}\\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3}\end{pmatrix},\ i = 1,2.$$

Figure 2.1 shows the data set with a three-step monotone missing data pattern that is used to calculate $\boldsymbol{d}_2$ and $\boldsymbol{d}_3$, respectively.

The values of both MLEs coincide with those of Kanda and Fujikoshi (1998) derived by the conditional approach. In this paper, we present the MLEs for the case of a three-step monotone missing data in order to obtain a simplified T^2-type statistic for testing the mean vector.

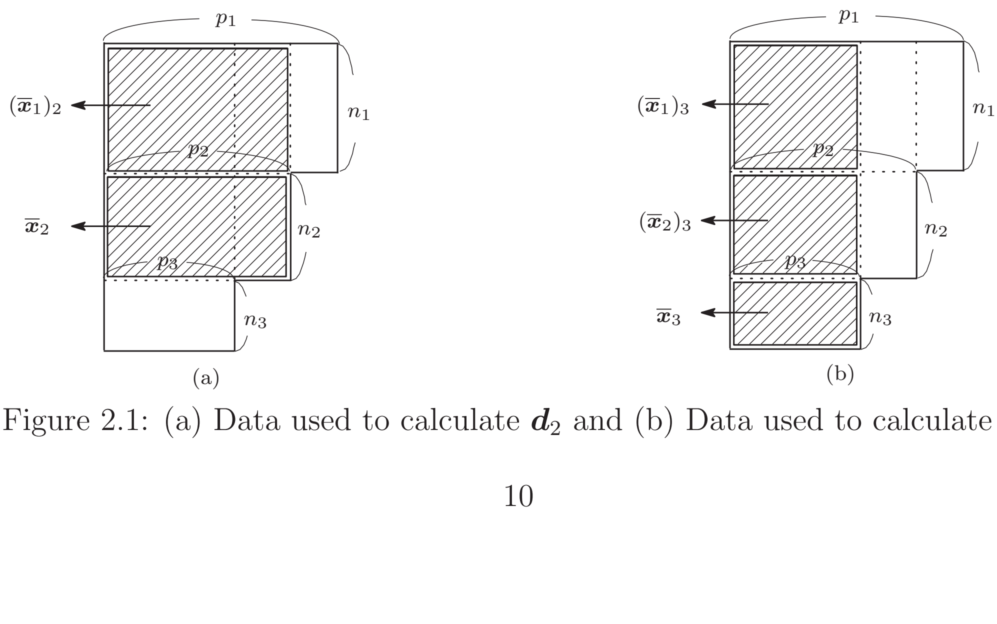

Figure 2.1: (a) Data used to calculate $\boldsymbol{d}_2$ and (b) Data used to calculate $\boldsymbol{d}_3$

2.1.2 Simplified T^2-type statistic

In this section, we consider the following hypothesis test with a three-step monotone missing data pattern:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \text{ vs. } H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0,$$

where $\boldsymbol{\mu}_0$ is known. Without loss of generality, we can assume that $\boldsymbol{\mu}_0 = \mathbf{0}$. To test the hypothesis H_0, we consider a usual Hotelling's T^2-type statistic given by

$$T^2 = \widehat{\boldsymbol{\mu}}' \, \widehat{\boldsymbol{\Gamma}}^{-1} \widehat{\boldsymbol{\mu}},$$

where $\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \widehat{\boldsymbol{T}}_2 \boldsymbol{d}_2 + \widehat{\boldsymbol{T}}_2 \widehat{\boldsymbol{T}}_3 \boldsymbol{d}_3$ and $\widehat{\boldsymbol{\Gamma}}$ is an estimator of $\boldsymbol{\Gamma} = \mathrm{Cov}(\widehat{\boldsymbol{\mu}})$. In order to discuss the distribution of T^2, the covariance matrix of $\widehat{\boldsymbol{\mu}}$, $\mathrm{Cov}(\widehat{\boldsymbol{\mu}})$, is a key quantity. However, we use the simplified T^2-type statistic with $\widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\mu}})$ instead of $\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})$ since $\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})$ is complicated and $\widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\mu}})$ and $\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})$ are asymptotically equivalent.

Therefore, we adopt that

$$\widetilde{T}^2 = \widehat{\boldsymbol{\mu}}' \, \widetilde{\boldsymbol{\Gamma}}^{-1} \widehat{\boldsymbol{\mu}},$$

where $\widetilde{\boldsymbol{\Gamma}} = \widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\mu}})$ and $\widetilde{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \boldsymbol{T}_2 \boldsymbol{d}_2 + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3$. Then we have $\mathrm{Cov}(\widetilde{\boldsymbol{\mu}}) = \mathrm{E}[\widetilde{\boldsymbol{\mu}} \widetilde{\boldsymbol{\mu}}'] - \boldsymbol{\mu}\boldsymbol{\mu}'$, where

$$\begin{aligned}
\mathrm{E}[\widetilde{\boldsymbol{\mu}} \widetilde{\boldsymbol{\mu}}'] =& \mathrm{E}[(\overline{\boldsymbol{x}}_1 + \boldsymbol{T}_2 \boldsymbol{d}_2 + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3)(\overline{\boldsymbol{x}}_1 + \boldsymbol{T}_2 \boldsymbol{d}_2 + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3)'] \\
=& \mathrm{E}[\overline{\boldsymbol{x}}_1 \overline{\boldsymbol{x}}_1' + \overline{\boldsymbol{x}}_1 \boldsymbol{d}_2' \boldsymbol{T}_2' + \overline{\boldsymbol{x}}_1 \boldsymbol{d}_3' \boldsymbol{T}_3' \boldsymbol{T}_2' + \boldsymbol{T}_2 \boldsymbol{d}_2 \overline{\boldsymbol{x}}_1' + \boldsymbol{T}_2 \boldsymbol{d}_2 \boldsymbol{d}_2' \boldsymbol{T}_2' \\
&+ \boldsymbol{T}_2 \boldsymbol{d}_2 \boldsymbol{d}_3' \boldsymbol{T}_3' \boldsymbol{T}_2' + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3 \overline{\boldsymbol{x}}_1' + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3 \boldsymbol{d}_2' \boldsymbol{T}_2' + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3 \boldsymbol{d}_3' \boldsymbol{T}_3' \boldsymbol{T}_2'].
\end{aligned}$$

Further, using the results,

$$\begin{aligned}
&\mathrm{E}[\overline{\boldsymbol{x}}_1 \overline{\boldsymbol{x}}_1'] = \frac{1}{n_1}\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}', \quad \mathrm{E}[\overline{\boldsymbol{x}}_1 \boldsymbol{d}_2'] = -\frac{n_2}{N_2 N_3}\begin{pmatrix} \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_{22}' \end{pmatrix}, \quad \mathrm{E}[\overline{\boldsymbol{x}}_1 \boldsymbol{d}_3'] = -\frac{n_3}{N_3 N_4}\begin{pmatrix} \boldsymbol{\Sigma}_3 \\ \boldsymbol{\Sigma}_{32}' \end{pmatrix}, \\
&\mathrm{E}[\boldsymbol{d}_2 \boldsymbol{d}_2'] = \frac{n_2}{N_2 N_3}\boldsymbol{\Sigma}_2, \qquad \mathrm{E}[\boldsymbol{d}_2 \boldsymbol{d}_3'] = \boldsymbol{O}, \quad \text{and} \quad \mathrm{E}[\boldsymbol{d}_3 \boldsymbol{d}_3'] = \frac{n_3}{N_3 N_4}\boldsymbol{\Sigma}_3,
\end{aligned}$$

we can obtain

$$\widetilde{\boldsymbol{\Gamma}} = \widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\mu}}) = \frac{1}{N_2}\widehat{\boldsymbol{\Sigma}}_1 - \frac{n_2}{N_2 N_3}\widehat{\boldsymbol{U}} - \frac{n_3}{N_3 N_4}\widehat{\boldsymbol{V}},$$

where

$$\widehat{\boldsymbol{U}} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_2 \\ \widehat{\boldsymbol{\Sigma}}_{22}' \end{pmatrix} \widehat{\boldsymbol{T}}_2', \;\; \widehat{\boldsymbol{V}} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_3 \\ \widehat{\boldsymbol{\Sigma}}_{32}' \end{pmatrix} \widehat{\boldsymbol{T}}_3' \widehat{\boldsymbol{T}}_2'.$$

Therefore, we can determine a simplified T^2-type statistic. We note that, under H_0, the simplified T^2-type statistic is asymptotically distributed as a χ^2 distribution with p degrees of freedom when $n_1, N_4 \to \infty$ with $n_1/N_4 \to \delta \in (0, 1]$. However, it is noted that χ^2 approximation is not a good approximation to the upper percentile of the simplified T^2-type statistic when the sample is not large. Using the same concept adopted for two-step monotone missing data by Seko et al. (2012), we propose the approximate upper percentile of $\widetilde{T}^2$ statistic since it is difficult to find the exact upper percentiles of the $\widetilde{T}^2$ statistic. That is

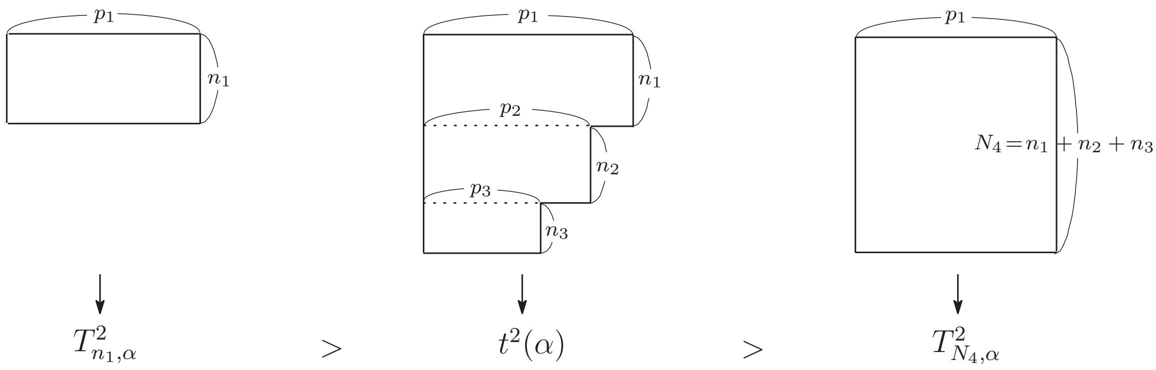

Figure 2.2: Approximation for the upper percentiles of $\widetilde{T}^2$

$$t^2_{\text{YS}\cdot\text{L1}}(\alpha) = \left\{1 - \frac{n_2p_2 + n_3p_3}{(n_2+n_3)p_1}\right\} T^2_{n_1,\alpha} + \frac{n_2p_2+n_3p_3}{(n_2+n_3)p_1} T^2_{N_4,\alpha},$$

where

$$T^2_{n_1,\alpha} = \frac{n_1p_1}{n_1-p_1}F_{p_1,n_1-p_1,\alpha}, \quad T^2_{N_4,\alpha} = \frac{N_4p_1}{N_4-p_1}F_{p_1,N_4-p_1,\alpha},$$

and $F_{p,q,\alpha}$ is the upper 100α percentile of the F distribution with p and q degrees of freedom. We note that $T^2_{n_1} = n_1\overline{\boldsymbol{x}}_1'\widehat{\boldsymbol{\Sigma}}^{-1}_{\text{ML}}\overline{\boldsymbol{x}}_1$ is distributed as $n_1p_1/(n_1-p_1)F_{p_1,n_1-p_1}$ since $\widehat{\boldsymbol{\Sigma}}_{\text{ML}} = (1/n_1)\sum_{j=1}^{n_1}(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_1)(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_1)'$ is not unbiased estimator but maximum likelihood estimator based on $n_1 \times p_1$ complete data set. As in Figure 2.2, $T^2_{n_1,\alpha}$ and $T^2_{N_4,\alpha}$ are calculated from an $n_1 \times p_1$ complete data set (left-hand side) and an $N_4 \times p_1$ complete data set (right-hand side), respectively. We also noted from Figure 2.2 that the upper percentiles of $\widetilde{T}^2$, $t^2(\alpha)$ may be between $T^2_{N_4,\alpha}$ and $T^2_{n_1,\alpha}$, and the value of $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ is an approximation for $t^2(\alpha)$ using the linear interpolation for the coordinates $(n_1p_1, T^2_{n_1,\alpha})$ and $(N_4p_1, T^2_{N_4,\alpha})$. Indeed, $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ converges to $T^2_{n_1,\alpha}$ as $n_2 \to 0, n_3 \to 0$, and $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ converges to $T^2_{N_4,\alpha}$ as $p_2 \to p_1, p_3 \to p_1$. As a remark, Seko et al. (2012) gave the approximation using unbiased estimator of $\boldsymbol{\Sigma}$ but we adopt the maximum likelihood estimator as our approximation. In this sense, it may be noted that our approximation gives a slight improvement.

2.1.3 Simultaneous confidence intervals and testing equality of mean components

We consider the simultaneous confidence intervals for any and all linear compounds of the mean when the data have three-step monotone missing observations. Using the approximate upper percentiles of $\widetilde{T}^2$ from Section 2.1.2, for any nonnull vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_p)'$, the approximate simultaneous confidence intervals for $\boldsymbol{a}'\boldsymbol{\mu}$ are given by

$$\boldsymbol{a}'\widehat{\boldsymbol{\mu}} - \sqrt{t^2_{\text{YS}\cdot\text{L1}}(\alpha)\boldsymbol{a}'\widetilde{\boldsymbol{\Gamma}}\boldsymbol{a}} \le \boldsymbol{a}'\boldsymbol{\mu} \le \boldsymbol{a}'\widehat{\boldsymbol{\mu}} + \sqrt{t^2_{\text{YS}\cdot\text{L1}}(\alpha)\boldsymbol{a}'\widetilde{\boldsymbol{\Gamma}}\boldsymbol{a}}, \quad \forall \boldsymbol{a} \in \boldsymbol{R}^p - \{\boldsymbol{0}\}.$$

For a two-step monotone missing data, see Seko et al. (2012). Further, we consider the testing equality of mean components for the case of a three-step monotone missing data, that is,

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_p \text{ vs. } H_1 : \text{At least two means are different.}$$

In this case, let $\boldsymbol{y}_{ij} = \boldsymbol{C}_i \boldsymbol{x}_{ij}, i = 1, 2, 3, j = 1, 2, \ldots, n_i$, where $\boldsymbol{C}_i$ is a $(p_i - 1) \times p_i$ matrix such that $\boldsymbol{C}_i \mathbf{1} = \mathbf{0}$ and $\boldsymbol{C}_i \boldsymbol{C}_i' = \boldsymbol{I}_{p_i-1}$, then $\boldsymbol{y}_{ij}$'s are distributed as $N_{p_i-1}(\boldsymbol{C}_i \boldsymbol{\mu}_i,\ \boldsymbol{I}_{p_i-1})$, because without loss of generality, we can assume that $\boldsymbol{\Sigma} = \boldsymbol{I}$ when we consider the T^2-type statistic with a monotone missing data. Hence, the simplified T^2-type statistic is given by

$$\widetilde{T}_c^2 = \widehat{\boldsymbol{\mu}}^{*\prime} \widetilde{\boldsymbol{\Gamma}}^{*^{-1}} \widehat{\boldsymbol{\mu}}^*,$$

where $\widehat{\boldsymbol{\mu}}^*$ is the MLE of $\boldsymbol{\mu}^* = \boldsymbol{C}_1 \boldsymbol{\mu}$ and $\widetilde{\boldsymbol{\Gamma}}^* = \widehat{\mathrm{Cov}}(\widetilde{\boldsymbol{\mu}}^*)$, $\widetilde{\boldsymbol{\mu}}^* = \boldsymbol{C}_1 \widetilde{\boldsymbol{\mu}}$. Therefore, essentially, using the same values of $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ obtained in Section 2.1.2 for the approximate upper percentile of the $\widetilde{T}_c^2$ statistic, we can test the equality of mean components with a three-step monotone missing data.

2.1.4 Simulation studies

We compute the upper percentiles of the simplified T^2-type statistic with a three-step monotone missing data using the Monte Carlo simulation. One million simulations were conducted for each combination of selected values of $p_i, n_i, i = 1, 2, 3$ and α. It is interesting to see how the approximations are close to the exact upper percentiles. Simulation results related to this problem are summarized in Tables 2.1–2.3. Computations are made for the following two cases:

$$\begin{aligned} \text{Case I } : & (p_1, p_2, p_3) = (6, 4, 2),\ (12, 8, 4), \\ & n_1 = 30, 50, 100, 200, 300,\ n_2, n_3 = 10, 20,\ \alpha = 0.05, 0.01, \end{aligned}$$

where the sets of (n_1, n_2, n_3) are combinations of n_1, n_2 and n_3.

$$\begin{aligned} \text{Case II } : & (p_1, p_2, p_3) = (12, 4, 2), \\ & (n_1, n_2, n_3) = (30w, 10w, 10w),\ w = 1(1)5, 8, 12,\ \alpha = 0.05, 0.01. \end{aligned}$$

Tables 2.1 and 2.2 list the simulated upper percentiles of $\widetilde{T}^2$, $\widetilde{t}^2_{\text{simu}}(\alpha)$, the approximate upper percentiles of $\widetilde{T}^2$, $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$, and the upper percentiles of χ^2 distribution with p degrees of freedom, $\chi^2_p(\alpha)$ for Case I. It may be noted from Tables 2.1 and 2.2 that the simulated values are closer to the upper percentiles of χ^2 distribution when the sample size n_1 becomes large. Therefore, we note that the χ^2 approximation $\chi^2_p(\alpha)$ is not good for cases where n_1 is small. However, it is seen that the proposed approximation $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ is considerably good even for cases where n_1 is not large. In addition, Tables 2.1 and 2.2 list the simulated coverage probabilities for $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ and $\chi^2_p(\alpha)$ for Case I. The simulated coverage probabilities for $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ and $\chi^2_p(\alpha)$ are defined as

$$\mathrm{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YS}\cdot\text{L1}}(\alpha)\},\quad \mathrm{CP}(\chi^2_p(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > \chi^2_p(\alpha)\},$$

Table 2.1: The simulated and the approximate values for $\widetilde{T}^2$, χ^2 approximation, and the simulated coverage probabilities when $(p_1, p_2, p_3) = (6, 4, 2)$

Sample size			Upper percentile			Coverage probability	
n_1	n_2	n_3	$\widetilde{t}^2_{\text{simu}}(\alpha)$	$t^2_{\text{YS}\cdot\text{L1}}(\alpha)$	$\chi^2_p(\alpha)$	$\text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$	$\text{CP}(\chi^2_p(\alpha))$
				$\alpha = 0.05$			
30	10	10	17.68	17.29	12.59	0.946	0.851
50	10	10	15.36	15.25	12.59	0.948	0.896
100	10	10	13.92	13.89	12.59	0.950	0.925
200	10	10	13.22	13.24	12.59	0.950	0.938
300	10	10	13.02	13.02	12.59	0.950	0.942
30	20	10	17.25	16.78	12.59	0.944	0.859
50	20	10	15.17	15.02	12.59	0.948	0.900
100	20	10	13.84	13.82	12.59	0.950	0.926
200	20	10	13.23	13.22	12.59	0.950	0.938
300	20	10	13.03	13.01	12.59	0.950	0.942
30	10	20	17.56	17.18	12.59	0.946	0.853
50	10	20	15.28	15.17	12.59	0.948	0.897
100	10	20	13.91	13.86	12.59	0.949	0.925
200	10	20	13.25	13.23	12.59	0.950	0.938
300	10	20	13.03	13.02	12.59	0.950	0.942
30	20	20	17.16	16.77	12.59	0.945	0.860
50	20	20	15.11	14.99	12.59	0.948	0.901
100	20	20	13.85	13.80	12.59	0.949	0.926
200	20	20	13.23	13.21	12.59	0.950	0.938
300	20	20	13.01	13.01	12.59	0.950	0.942
				$\alpha = 0.01$			
30	10	10	25.53	24.81	16.81	0.988	0.940
50	10	10	21.40	21.22	16.81	0.990	0.966
100	10	10	18.98	18.94	16.81	0.990	0.980
200	10	10	17.86	17.86	16.81	0.990	0.986
300	10	10	17.55	17.51	16.81	0.990	0.987
30	20	10	24.76	23.91	16.81	0.988	0.945
50	20	10	21.02	20.84	16.81	0.989	0.968
100	20	10	18.86	18.82	16.81	0.990	0.981
200	20	10	17.83	17.83	16.81	0.990	0.986
300	20	10	17.49	17.50	16.81	0.990	0.987
30	10	20	25.27	24.63	16.81	0.989	0.941
50	10	20	21.25	21.09	16.81	0.990	0.967
100	10	20	18.90	18.89	16.81	0.990	0.980
200	10	20	17.87	17.85	16.81	0.990	0.985
300	10	20	17.50	17.50	16.81	0.990	0.987
30	20	20	24.54	23.91	16.81	0.988	0.946
50	20	20	20.92	20.78	16.81	0.990	0.969
100	20	20	18.84	18.79	16.81	0.990	0.981
200	20	20	17.84	17.82	16.81	0.990	0.986
300	20	20	17.47	17.49	16.81	0.990	0.987

Table 2.2: The simulated and the approximate values for $\widetilde{T}^2$, χ^2 approximation, and the simulated coverage probabilities when $(p_1, p_2, p_3) = (12, 8, 4)$

Sample size			Upper percentile			Coverage probability	
n_1	n_2	n_3	$\widetilde{t}^2_{\text{simu}}(\alpha)$	$t^2_{\text{YS}\cdot\text{L1}}(\alpha)$	$\chi^2_p(\alpha)$	CP$(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$	CP$(\chi^2_p(\alpha))$
				$\alpha = 0.05$			
30	10	10	42.70	39.35	21.03	0.931	0.598
50	10	10	30.55	29.85	21.03	0.944	0.779
100	10	10	25.10	24.99	21.03	0.949	0.881
200	10	10	22.95	22.93	21.03	0.950	0.919
300	10	10	22.27	22.28	21.03	0.950	0.931
30	20	10	41.16	37.15	21.03	0.925	0.620
50	20	10	29.89	29.04	21.03	0.942	0.791
100	20	10	24.92	24.77	21.03	0.948	0.884
200	20	10	22.90	22.87	21.03	0.950	0.920
300	20	10	22.26	22.25	21.03	0.950	0.931
30	10	20	42.40	39.09	21.03	0.931	0.601
50	10	20	30.37	29.60	21.03	0.943	0.782
100	10	20	25.06	24.90	21.03	0.948	0.881
200	10	20	22.94	22.90	21.03	0.950	0.919
300	10	20	22.30	22.27	21.03	0.950	0.930
30	20	20	41.01	37.35	21.03	0.928	0.622
50	20	20	29.74	28.93	21.03	0.943	0.792
100	20	20	24.89	24.71	21.03	0.948	0.884
200	20	20	22.89	22.85	21.03	0.950	0.921
300	20	20	22.28	22.24	21.03	0.950	0.931
				$\alpha = 0.01$			
30	10	10	60.93	54.96	26.22	0.984	0.752
50	10	10	40.48	39.41	26.22	0.988	0.900
100	10	10	32.16	32.01	26.22	0.990	0.961
200	10	10	28.96	28.97	26.22	0.990	0.979
300	10	10	28.04	28.02	26.22	0.990	0.983
30	20	10	58.48	51.46	26.22	0.981	0.773
50	20	10	39.41	38.17	26.22	0.988	0.908
100	20	10	31.88	31.68	26.22	0.990	0.963
200	20	10	28.91	28.89	26.22	0.990	0.979
300	20	10	27.98	27.99	26.22	0.990	0.983
30	10	20	60.48	54.65	26.22	0.984	0.756
50	10	20	40.22	39.04	26.22	0.988	0.902
100	10	20	32.10	31.87	26.22	0.989	0.961
200	10	20	28.98	28.93	26.22	0.990	0.979
300	10	20	28.02	28.01	26.22	0.990	0.983
30	20	20	58.52	51.86	26.22	0.982	0.774
50	20	20	39.31	38.02	26.22	0.988	0.909
100	20	20	31.80	31.59	26.22	0.989	0.963
200	20	20	28.90	28.85	26.22	0.990	0.979
300	20	20	28.06	27.97	26.22	0.990	0.983

Table 2.3: The simulated and the approximate values for $\widetilde{T}^2$, χ^2 approximation, and the simulated coverage probabilities when $(p_1, p_2, p_3) = (12, 4, 2)$

Sample size			Upper percentile			Coverage probability	
n_1	n_2	n_3	$\widetilde{t}^2_{\text{simu}}(\alpha)$	$t^2_{\text{YS}\cdot\text{L1}}(\alpha)$	$\chi^2_p(\alpha)$	$\text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$	$\text{CP}(\chi^2_p(\alpha))$
				$\alpha = 0.05$			
30	10	10	45.82	43.09	21.03	0.937	0.561
60	20	20	28.92	28.41	21.03	0.945	0.809
90	30	30	25.71	25.45	21.03	0.947	0.869
120	40	40	24.33	24.19	21.03	0.948	0.895
150	50	50	23.59	23.48	21.03	0.949	0.908
240	80	80	22.56	22.50	21.03	0.949	0.926
360	120	120	22.06	21.99	21.03	0.949	0.934
				$\alpha = 0.01$			
30	10	10	65.85	61.19	26.22	0.986	0.718
60	20	20	37.87	37.18	26.22	0.989	0.920
90	30	30	33.05	32.70	26.22	0.989	0.955
120	40	40	30.94	30.81	26.22	0.990	0.968
150	50	50	29.89	29.78	26.22	0.990	0.974
240	80	80	28.39	28.34	26.22	0.990	0.981
360	120	120	27.69	27.60	26.22	0.990	0.985

respectively. It may be noted from Tables 2.1 and 2.2 that the simulated coverage probabilities for $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$, $\text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$ are considerably close to the nominal level $1-\alpha$ even for cases where n_1 is small. Therefore, it can be concluded that our approximation procedure is very accurate even for small samples. Table 2.3 lists the values of $t^2_{\text{simu}}(\alpha)$, $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ and $\chi^2_p(\alpha)$ for Case II. In addition, the values of $\text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$ and $\text{CP}(\chi^2_p(\alpha))$ for Case II are listed in Table 2.3. We note that the missing rate of the data sets in Case II is constant $(= 0.3)$, where

$$\text{the missing rate} = 1 - \frac{1}{N_4 p_1} \sum_{i=1}^{3} n_i p_i.$$

It appears from Table 2.3 that our approximation is considerably good even when (n_1, n_2, n_3) is small.

2.1.5 Numerical example

In this section, we shall discuss an example to illustrate the approximation developed in this paper. In this example, we treat a three-step monotone missing data taken from Wei and Lachin (1984). A data material is presented where serum cholesterol on each patient has been measured five different time points: before start (baseline) and at 6, 12, 20 and 24 months after study start. In our analysis, we consider only the case of three-step monotone missing data. We are interested in the change from the baseline at each post-baseline time point. Thus, we have the three-step monotone missing data of $(p_1, p_2, p_3) = (4, 3, 2)$ and $(n_1, n_2, n_3) = (36, 7, 12)$. We consider the hypothesis $H_0 : \boldsymbol{\mu} = (\mu_2 - \mu_1,\ \mu_3 - \mu_1,\ \mu_4 - \mu_1,\ \mu_5 - \mu_1)' = \mathbf{0}$. From this three-step monotone missing data,

Table 2.4: 99% simultaneous confidence intervals

$\alpha = 0.01$	$\widetilde{t}^2_{\text{simu}}(\alpha)$	$t^2_{\text{YS}\cdot\text{L1}}(\alpha)$	$\chi^2_p(\alpha)$
$\mu_2 - \mu_1$	(2.77, 36.98)	(2.85, 36.90)	(4.72, 35.02)
$\mu_3 - \mu_1$	(6.53, 44.31)	(6.62, 44.22)	(8.69, 42.15)
$\mu_4 - \mu_1$	(6.02, 48.11)	(6.12, 48.01)	(8.42, 45.70)
$\mu_5 - \mu_1$	(−1.91, 60.29)	(−1.76, 60.14)	(1.65, 56.73)

we can compute the maximum likelihood estimate for $\boldsymbol{\mu}$, which is given by

$$\widehat{\boldsymbol{\mu}} = (19.87,\ 25.42,\ 27.06,\ 29.19)'.$$

Also, an estimate for the covariance matrix $\boldsymbol{\Gamma}$ for $\widetilde{\boldsymbol{\mu}}$ is given by

$$\widetilde{\boldsymbol{\Gamma}} = \widehat{\text{Cov}}(\widetilde{\boldsymbol{\mu}}) = \begin{pmatrix} 17.29 & 9.98 & 12.74 & 13.98 \\ 9.98 & 21.08 & 11.20 & 18.80 \\ 12.74 & 11.20 & 26.17 & 17.55 \\ 13.98 & 18.80 & 17.55 & 57.13 \end{pmatrix}.$$

Therefore, the value of the test statistic $\widetilde{T}^2$ is $40.58 > \widetilde{t}^2_{\text{simu}}(0.01) = 16.93$. Thus, the hypothesis H_0 is rejected at the significance level of 0.01. When we use $t^2_{\text{YS}\cdot\text{L1}}(0.01) = 16.77$ or $\chi^2_4(0.01) = 13.28$, the hypothesis H_0 is also rejected and the simultaneous confidence intervals for the change from the baseline at each time point can be obtained. The approximate simultaneous confidence intervals for $\mu_j - \mu_1, j = 2, \ldots, 5$ with level $1 - \alpha = 0.99$ as in Section 2.1.3 can be computed and are summarized in Table 2.4. It can be seen from Table 2.4 that $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ gives very similar confidence intervals to the simulation value $\widetilde{t}^2_{\text{simu}}(\alpha)$, while $\chi^2_p(\alpha)$ gives an incorrect result for $\mu_5 - \mu_1$. This implies that the approximation $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ is good approximate upper percentile of $\widetilde{T}^2$ statistic.

2.2 k-step monotone missing data

In this section, we consider the distribution of the Hotelling's T^2-type test statistic to test a mean vector with monotone missing data. We give a simplified T^2-type statistic and propose the approximate upper percentiles of the simplified T^2-type statistic in the case of data with general k-step monotone missing data pattern. We also consider the approximate simultaneous confidence intervals for any and all linear compounds of the mean and the testing equality of mean components. Finally, the accuracy and asymptotic behavior of the approximations are investigated by Monte Carlo simulation.

2.2.1 Monotone missing data and MLE

We first present some notations, definitions, and the setting in this paper, and we derive the MLEs. Let $\boldsymbol{x}$ be distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and let $\boldsymbol{x}_i = (\boldsymbol{x})_i$ be the subvector of $\boldsymbol{x}$ containing the first p_i components of $\boldsymbol{x}$. Then, $\boldsymbol{x}_i (= (x_1, x_2, \ldots, x_{p_i})')$ is distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2, \ldots, k$, with $p = p_1 > p_2 > \cdots > p_k$, where $\boldsymbol{\mu}_i = (\boldsymbol{\mu})_i = (\mu_1, \mu_2, \ldots, \mu_{p_i})'$ and $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ principal submatrix of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}_1)$. Further, for the covariance matrix, for $1 \leq i < j \leq k$, let $(\boldsymbol{\Sigma}_i)_j$ be the principal submatrix of $\boldsymbol{\Sigma}_i$ of order $p_j \times p_j$; we define

$$\boldsymbol{\Sigma}_i = (\boldsymbol{\Sigma}_1)_i, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{i2} \\ \boldsymbol{\Sigma}'_{i2} & \boldsymbol{\Sigma}_{i3} \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_{i-1} = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{(i-1,2)} \\ \boldsymbol{\Sigma}'_{(i-1,2)} & \boldsymbol{\Sigma}_{(i-1,3)} \end{pmatrix}, \quad i = 2, 3, \ldots, k.$$

We use these notations, which are based on Jinadasa and Tracy (1992), throughout this paper.

Suppose we have n_1 observations on $\boldsymbol{x}_1$, n_2 observations on $\boldsymbol{x}_2$, $\cdots$, n_k observations on $\boldsymbol{x}_k$. If $\boldsymbol{x}_{ij}$ denotes the jth observation on $\boldsymbol{x}_i$, then the k-step monotone missing data set is of the form

$$\begin{pmatrix}
\boldsymbol{x}'_{11} & & \\
\vdots & & \\
\boldsymbol{x}'_{1n_1} & & \\
\boldsymbol{x}'_{21} & & * \cdots * \\
\vdots & & \vdots \quad \vdots \\
\boldsymbol{x}'_{2n_2} & & * \cdots * \\
\ddots & & \\
\boldsymbol{x}'_{k1} & * \cdots * & * \cdots * \\
\vdots & \vdots \quad \vdots & \vdots \quad \vdots \\
\boldsymbol{x}'_{kn_k} & * \cdots * & * \cdots *
\end{pmatrix},$$

where "$*$" indicates a missing observation. Let $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{in_i}$ be distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ for $i = 1, 2, \ldots, k$, with $n_1 - 1 \geq p_1$, and let

$$\overline{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}, \quad \boldsymbol{E}_i = \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)', \; i = 1, 2, \ldots, k.$$

Further, we define

$$N_1 = 0, \quad N_{i+1} = N_i + n_i \left(= \sum_{j=1}^{i} n_j\right), \quad i = 1, 2, \ldots, k,$$

$$\boldsymbol{d}_1 = \overline{\boldsymbol{x}}_1, \quad \boldsymbol{d}_i = \frac{n_i}{N_{i+1}} \left[\overline{\boldsymbol{x}}_i - \frac{1}{N_i} \sum_{j=1}^{i-1} n_j (\overline{\boldsymbol{x}}_j)_i\right], \quad i = 2, 3, \ldots, k,$$

$$\boldsymbol{f}_1 = \boldsymbol{d}_1, \quad \boldsymbol{f}_i = \boldsymbol{U}_i \boldsymbol{d}_i, \quad i = 2, 3, \ldots, k, \quad \boldsymbol{U}_1 = \boldsymbol{T}_1, \quad \boldsymbol{U}_i = \boldsymbol{U}_{i-1} \boldsymbol{T}_i, \quad i = 2, 3, \ldots, k,$$

$$\boldsymbol{T}_1 = \boldsymbol{I}_{p_1}, \quad \boldsymbol{T}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{\Sigma}'_{(i,2)} \boldsymbol{\Sigma}_{i+1}^{-1} \end{pmatrix}, \quad \widehat{\boldsymbol{T}}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \widehat{\boldsymbol{\Sigma}}'_{(i,2)} \widehat{\boldsymbol{\Sigma}}_{i+1}^{-1}, \end{pmatrix}, \quad i = 1, 2, \ldots, k-1,$$

$$\boldsymbol{H}_1 = \boldsymbol{E}_1, \quad \boldsymbol{H}_i = \boldsymbol{E}_i + \frac{N_i N_{i+1}}{n_i}\boldsymbol{d}_i\boldsymbol{d}_i', \quad i = 2, 3, \ldots, k$$

$$\boldsymbol{L}_1 = \boldsymbol{H}_1, \quad \boldsymbol{L}_i = (\boldsymbol{L}_{i-1})_i + \boldsymbol{H}_i, \quad i = 2, 3, \ldots, k$$

$$\boldsymbol{L}_{i1} = (\boldsymbol{L}_i)_{i+1}, \quad \boldsymbol{L}_i = \begin{pmatrix} \boldsymbol{L}_{i1} & \boldsymbol{L}_{i2} \\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1,$$

$$\boldsymbol{F}_1 = \boldsymbol{G}_1, \quad \boldsymbol{F}_i = \boldsymbol{F}_{i-1}\boldsymbol{G}_i, \quad i = 2, 3, \ldots, k$$

$$\boldsymbol{G}_1 = \boldsymbol{I}_{p_1}, \quad \boldsymbol{G}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{L}_{i2}'\boldsymbol{L}_{i1}^{-1} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1.$$

Then, $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ are given in the following theorem.

Theorem 2.2 (Jinadasa and Tracy, 1992) *The MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for the monotone sample are*

$$\widehat{\boldsymbol{\mu}} = \sum_{i=1}^{k} \widehat{\boldsymbol{f}}_i$$

with

$$\widehat{\boldsymbol{f}}_1 = \boldsymbol{d}_1, \quad \widehat{\boldsymbol{f}}_i = \boldsymbol{T}_1\widehat{\boldsymbol{T}}_2 \cdots \widehat{\boldsymbol{T}}_i\boldsymbol{d}_i, \quad i = 2, 3, \ldots, k \quad \textit{and}$$

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n_1}\boldsymbol{H}_1 + \sum_{i=2}^{k} \frac{1}{N_{i+1}}\boldsymbol{F}_i\left[\boldsymbol{H}_i - \frac{n_i}{N_i}\boldsymbol{L}_{i-1,1}\right]\boldsymbol{F}_i'.$$

2.2.2 A simplified T^2-type statistic and simultaneous confidence intervals

In this section, we consider the following hypothesis test with a k-step monotone missing data pattern:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \text{ vs. } H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0,$$

where $\boldsymbol{\mu}_0$ is known. Without loss of generality, we can assume that $\boldsymbol{\mu}_0 = \boldsymbol{0}$. To test the hypothesis H_0, we consider the simplified T^2-type statistic given by

$$\widetilde{T}^2 = \widehat{\boldsymbol{\mu}}'\,\widetilde{\boldsymbol{\Gamma}}^{-1}\widehat{\boldsymbol{\mu}},$$

where $\widehat{\boldsymbol{\mu}}(= \sum_{i=1}^{k}\widehat{\boldsymbol{f}}_i)$ is given in Theorem 2.2, $\widetilde{\boldsymbol{\Gamma}} = \widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}]$, and $\widetilde{\boldsymbol{\mu}} = \sum_{i=1}^{k}\boldsymbol{f}_i$. Then, we have the following theorem.

Theorem 2.3 (Yagi and Seo, 2017) *If the data have a k-step monotone pattern of missing observations, then an estimator of the covariance matrix of $\widetilde{\boldsymbol{\mu}}$ is given by*

$$\widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}] = \frac{1}{n_1}\widehat{\boldsymbol{\Sigma}}_1 - \sum_{i=2}^{k} \frac{n_i}{N_i N_{i+1}}\widehat{\boldsymbol{U}}_i\widehat{\boldsymbol{\Sigma}}_i\widehat{\boldsymbol{U}}_i',$$

where

$$\widehat{\boldsymbol{U}}_1 = \boldsymbol{T}_1, \quad \widehat{\boldsymbol{U}}_i = \widehat{\boldsymbol{U}}_{i-1}\widehat{\boldsymbol{T}}_i, \quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{T}_1 = \boldsymbol{I}_{p_1}, \quad \widehat{\boldsymbol{T}}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \widehat{\boldsymbol{\Sigma}}'_{(i,2)}\widehat{\boldsymbol{\Sigma}}_{i+1}^{-1} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1.$$

proof First, since $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \mathrm{E}[\widetilde{\boldsymbol{\mu}}\widetilde{\boldsymbol{\mu}}'] - \boldsymbol{\mu}\boldsymbol{\mu}'$ and $\widetilde{\boldsymbol{\mu}} = \sum_{i=1}^{k} \boldsymbol{f}_i$, we have

$$\mathrm{E}[\widetilde{\boldsymbol{\mu}}\widetilde{\boldsymbol{\mu}}'] = \mathrm{E}[\boldsymbol{f}_1\boldsymbol{f}_1'] + \sum_{r=2}^{k} \mathrm{E}[\boldsymbol{f}_r\boldsymbol{f}_r'] + 2\left[\sum_{s=2}^{k} \mathrm{E}[\boldsymbol{f}_1\boldsymbol{f}_s'] + \sum_{\substack{r=2 \\ r<s}}^{k}\sum_{s=3}^{k} \mathrm{E}[\boldsymbol{f}_r\boldsymbol{f}_s']\right].$$

Further, using the following results,

$$\mathrm{E}[\boldsymbol{f}_1\boldsymbol{f}_1'] = \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \boldsymbol{\mu}_1\boldsymbol{\mu}_1', \quad \mathrm{E}[\boldsymbol{f}_r\boldsymbol{f}_r'] = \frac{n_r}{N_rN_{r+1}}\boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{U}_r', \quad r = 2, 3, \ldots, k,$$
$$\mathrm{E}[\boldsymbol{f}_1\boldsymbol{f}_s'] = -\frac{n_s}{N_sN_{s+1}}\begin{pmatrix} \boldsymbol{\Sigma}_s \\ \boldsymbol{\Sigma}_{s2}' \end{pmatrix}\boldsymbol{U}_s', \quad s = 2, 3, \ldots, k,$$

and

$$\mathrm{E}[\boldsymbol{f}_r\boldsymbol{f}_s'] = \boldsymbol{O}, \quad 2 \leq r < s \leq k,$$

we obtain

$$\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \sum_{r=2}^{k} \frac{n_r}{N_rN_{r+1}}\left[\boldsymbol{U}_r\boldsymbol{\Sigma}_r - 2\begin{pmatrix} \boldsymbol{\Sigma}_r \\ \boldsymbol{\Sigma}_{r2}' \end{pmatrix}\right]\boldsymbol{U}_r',$$

where $\boldsymbol{U}_1 = \boldsymbol{T}_1,\ \boldsymbol{U}_i = \boldsymbol{U}_{i-1}\boldsymbol{T}_i,\ i = 2, 3, \ldots, k.$ Therefore,

$$\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \frac{1}{n_1}\boldsymbol{\Sigma}_1 - \sum_{r=2}^{k} \frac{n_r}{N_rN_{r+1}}\boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{U}_r'$$

since $\boldsymbol{U}_r\boldsymbol{\Sigma}_r = \begin{pmatrix} \boldsymbol{\Sigma}_r \\ \boldsymbol{\Sigma}_{r2}' \end{pmatrix}$. After replacing the unknown parameters in this equation by their MLEs, we get the result. $\square$

For a two-step monotone missing data pattern, Yagi and Seo (2015a) gave $\mathrm{Cov}(\widehat{\boldsymbol{\mu}})$ as well as $\mathrm{Cov}(\widetilde{\boldsymbol{\mu}})$, and Seko et al. (2012) discussed the usual Hotelling's T^2-type statistic, $T^2 = \widehat{\boldsymbol{\mu}}'\,\widehat{\boldsymbol{\Gamma}}^{-1}\widehat{\boldsymbol{\mu}}$, and its null distribution using other definitions. Comparing our notation with that of Krishnamoorthy and Pannala (1999), we can confirm that the simplified T^2-type statistic, $\widetilde{T}^2$, coincides with the pivotal quantity of the Hotelling's T^2 statistic in Krishnamoorthy and Pannala (1999, p.397).

We note that under H_0, the simplified T^2-type statistic is asymptotically distributed as a χ^2 distribution with p degrees of freedom when $n_1, N_{k+1} \to \infty$ with $n_1/N_{k+1} \to \delta \in$

$(0,1]$. The derivation idea of the condition for sample sizes is as follows. For $k=2$, the condition is given by $n_1, N_3 \to \infty$ with $n_1/N_3 \to \delta \in (0,1]$ (see, Chang and Richards, 2009, p.1891). In the general k-step case, the condition can be written as

$$\begin{aligned} & n_1, N_3, N_4 \ldots, N_{k+1} \to \infty \text{ with } n_1/N_3 \to \delta_3 \in (0,1], \\ & n_1/N_4 \to \delta_4 \in (0,1], \ldots, \text{and } n_1/N_{k+1} \to \delta_{k+1} \in (0,1], \end{aligned}$$

where $N_{i+1} = \sum_{j=1}^{i} n_j$, $i = 1, 2, \ldots, k$. That is, this condition is equivalent to

$$n_1, N_{k+1} \to \infty \text{ with } n_1/N_{k+1} \to \delta \in (0,1]$$

since "$n_1/N_3 \to \delta_3 \in (0,1], n_1/N_4 \to \delta_4 \in (0,1], \ldots,$ and $n_1/N_k \to \delta_k \in (0,1]$" is included in "$n_1/N_{k+1} \to \delta \in (0,1]$". However, it has been noted that the χ^2 approximation is not a good approximation to the upper percentile of the simplified T^2-type statistic when the sample is not large. Using the same concept for three-step monotone missing data used in Section 2.1.2, we propose the approximate upper percentile of the $\widetilde{T}^2$ statistic since it is difficult to find the exact upper percentiles of the $\widetilde{T}^2$ statistic. The two kinds of approximate upper 100α percentiles of the $\widetilde{T}^2$ statistic are given by

$$\begin{aligned} t^2_{\text{YS}\cdot\text{L1}}(\alpha) &= (1-\omega_1)T^2_{n_1,\alpha} + \omega_1 T^2_{N_{k+1},\alpha}, \\ t^2_{\text{YS}\cdot F1}(\alpha) &= \frac{n_1^* p_1}{n_1^* - p_1} F_{p_1, n_1^* - p_1, \alpha}, \end{aligned}$$

where

$$\begin{aligned} T^2_{n_1,\alpha} &= \frac{n_1 p_1}{n_1 - p_1} F_{p_1, n_1 - p_1, \alpha}, \ T^2_{N_{k+1},\alpha} = \frac{N_{k+1} p_1}{N_{k+1} - p_1} F_{p_1, N_{k+1} - p_1, \alpha}, \\ \omega_1 &= \frac{\sum_{i=2}^{k} n_i p_i}{p_1 \sum_{i=2}^{k} n_i}, \ n_1^* = \frac{1}{p_1} \sum_{i=1}^{k} n_i p_i, \end{aligned}$$

and $F_{p,q,\alpha}$ is the upper 100α percentile of the F distribution with p and q degrees of freedom.

Further, we consider the simultaneous confidence intervals for any and all linear compounds of the mean when the data have k-step monotone missing observations. Using the approximate upper percentiles of $\widetilde{T}^2$, for any nonnull vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)'$, the approximate simultaneous confidence intervals for $\boldsymbol{c}'\boldsymbol{\mu}$ are given by

$$\boldsymbol{c}'\widehat{\boldsymbol{\mu}} - \sqrt{t^2_{\text{app}\cdot 1}(\alpha)\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}\boldsymbol{c}} \leq \boldsymbol{c}'\boldsymbol{\mu} \leq \boldsymbol{c}'\widehat{\boldsymbol{\mu}} + \sqrt{t^2_{\text{app}\cdot 1}(\alpha)\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}\boldsymbol{c}}, \quad \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\}.$$

where $t^2_{\text{app}\cdot 1}(\alpha)$ is the value of $t^2_{\text{YS}\cdot\text{L1}}(\alpha)$ or $t^2_{\text{YS}\cdot F1}(\alpha)$.

2.2.3 Simulation studies

In this section, we investigate the accuracy and asymptotic behavior of the approximations for the upper percentiles of the simplified T^2-type statistic by Monte Carlo simulation. We provide the simulated upper percentiles and their approximations for selected parameters.

Table 2.5: Simulated and approximate values and coverage probabilities when $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability			
n_1	$n_2 = \cdots = n_5$	$\widetilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L1}$	$t^2_{{\rm YS}\cdot F1}$	$t^2_{\rm KP}$	${\rm CP}_{\rm YS\cdot L1}$	${\rm CP}_{{\rm YS}\cdot F1}$	${\rm CP}_{\rm KP}$	${\rm CP}_{\chi^2}$
				$\alpha = 0.05$					
19	5	367.43	234.37	76.53	–	0.898	0.567	–	0.094
20	5	237.96	163.63	72.10	237.66	0.900	0.643	0.950	0.128
30	5	63.20	57.08	50.13	60.33	0.927	0.886	0.940	0.429
40	5	46.14	44.02	42.06	43.89	0.938	0.924	0.937	0.598
50	5	39.79	38.71	37.90	38.06	0.943	0.936	0.937	0.693
100	5	31.07	30.86	30.79	30.26	0.948	0.947	0.942	0.849
200	5	27.81	27.76	27.75	27.43	0.949	0.949	0.945	0.907
400	5	26.36	26.34	26.34	26.17	0.950	0.950	0.948	0.930
800	5	25.68	25.66	25.66	25.57	0.950	0.950	0.949	0.940
19	10	355.91	227.80	51.37	–	0.901	0.419	–	0.117
20	10	229.07	157.51	50.13	225.80	0.903	0.501	0.949	0.155
30	10	59.06	53.73	42.06	57.45	0.928	0.838	0.944	0.471
40	10	43.97	41.90	37.90	42.25	0.937	0.903	0.940	0.630
50	10	38.32	37.25	35.36	36.99	0.942	0.925	0.940	0.716
100	10	30.68	30.43	30.22	29.99	0.948	0.945	0.943	0.856
200	10	27.72	27.64	27.61	27.35	0.949	0.949	0.946	0.909
400	10	26.32	26.31	26.31	26.15	0.950	0.950	0.948	0.931
800	10	25.65	25.65	25.65	25.57	0.950	0.950	0.949	0.941
				$\alpha = 0.01$					
19	5	942.52	541.02	113.59	–	0.974	0.724	–	0.154
20	5	519.17	325.87	105.67	510.04	0.973	0.793	0.990	0.204
30	5	91.50	80.66	68.40	82.83	0.982	0.964	0.984	0.589
40	5	62.47	58.84	55.65	56.58	0.986	0.981	0.982	0.759
50	5	52.02	50.57	49.28	48.09	0.988	0.986	0.983	0.840
100	5	39.10	38.89	38.78	37.43	0.990	0.989	0.986	0.945
200	5	34.52	34.45	34.44	33.71	0.990	0.990	0.988	0.974
400	5	32.44	32.45	32.45	32.08	0.990	0.990	0.989	0.983
800	5	31.51	31.50	31.50	31.32	0.990	0.990	0.989	0.987
19	10	914.54	530.70	70.41	–	0.974	0.565	–	0.185
20	10	515.45	316.31	68.40	489.88	0.973	0.654	0.989	0.241
30	10	86.11	75.57	55.65	79.40	0.982	0.937	0.986	0.633
40	10	59.02	55.68	49.28	54.47	0.986	0.972	0.984	0.788
50	10	50.04	48.41	45.48	46.71	0.988	0.981	0.984	0.858
100	10	38.56	38.27	37.95	37.07	0.989	0.989	0.986	0.949
200	10	34.34	34.28	34.24	33.61	0.990	0.990	0.988	0.975
400	10	32.46	32.41	32.40	32.06	0.990	0.990	0.989	0.983
800	10	31.45	31.49	31.49	31.31	0.990	0.990	0.990	0.987

Note. ${\rm CP}_{\rm YS\cdot L1} = {\rm CP}(t^2_{\rm YS\cdot L1}(\alpha))$, ${\rm CP}_{{\rm YS}\cdot F1} = {\rm CP}(t^2_{{\rm YS}\cdot F1}(\alpha))$, ${\rm CP}_{\rm KP} = {\rm CP}(t^2_{\rm KP}(\alpha))$, ${\rm CP}_{\chi^2} = {\rm CP}(\chi^2_p(\alpha))$, $\chi^2_{15}(0.05) = 25.00$, $\chi^2_{15}(0.01) = 30.58$.

We also present the numerical comparisons of our approximation proposed in Section 2.2.2 and the approximation using F distribution in Krishnamoorthy and Pannala (1999). We compute the upper percentiles of the simplified T^2-type statistic with k-step monotone missing data using Monte Carlo simulation (10^6 runs). That is, the $\widetilde{T}^2$ statistic is computed 10^6 times based on the normal random vectors generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$, $i = 1, 2, \ldots, k$. Note that the simplified T^2-type statistic with two-step monotone missing data is lower triangular invariant (e.g., see Krishnamoorthy and Pannala, 1999 and Romer

Table 2.6: Simulated and approximate values and coverage probabilities when $(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}) = (20, 18, 16, 14, 12, 10, 8, 6, 4, 2)$

Sample Size		Upper Percentile				Coverage Probability			
n_1	$n_2 = \cdots = n_{10}$	$\widetilde{t}^2_{\text{simu}}$	$t^2_{\text{YS}\cdot\text{L1}}$	$t^2_{\text{YS}\cdot F1}$	t^2_{KP}	$\text{CP}_{\text{YS}\cdot\text{L1}}$	$\text{CP}_{\text{YS}\cdot F1}$	CP_{KP}	CP_{χ^2}
				$\alpha = 0.05$					
24	5	531.75	373.34	69.54	–	0.915	0.368	–	0.060
25	5	333.66	252.88	67.91	332.87	0.918	0.447	0.950	0.083
30	5	120.46	107.28	61.48	118.41	0.929	0.696	0.947	0.208
40	5	68.61	65.16	53.61	64.92	0.936	0.858	0.935	0.418
50	5	55.72	53.88	48.99	52.86	0.940	0.903	0.934	0.554
100	5	40.89	40.49	40.01	39.63	0.947	0.942	0.939	0.795
200	5	35.82	35.71	35.66	35.20	0.949	0.948	0.944	0.886
400	5	33.52	33.53	33.52	33.25	0.950	0.950	0.947	0.921
800	5	32.47	32.46	32.46	32.32	0.950	0.950	0.948	0.936
24	10	510.29	368.57	50.37	–	0.920	0.284	–	0.095
25	10	316.45	248.27	49.96	306.21	0.924	0.354	0.947	0.125
30	10	108.24	103.34	48.12	109.07	0.943	0.609	0.951	0.277
40	10	62.23	62.18	45.35	60.09	0.950	0.812	0.941	0.491
50	10	51.75	51.55	43.37	49.78	0.949	0.878	0.938	0.613
100	10	39.84	39.54	38.40	38.80	0.947	0.937	0.941	0.814
200	10	35.51	35.39	35.22	34.97	0.949	0.947	0.944	0.891
400	10	33.47	33.43	33.41	33.19	0.950	0.949	0.947	0.922
800	10	32.43	32.44	32.43	32.30	0.950	0.950	0.948	0.937
				$\alpha = 0.01$					
24	5	1415.31	873.19	92.91	–	0.977	0.512	–	0.105
25	5	756.91	509.35	90.38	726.25	0.978	0.601	0.989	0.142
30	5	194.70	162.55	80.52	178.83	0.981	0.840	0.987	0.328
40	5	93.16	87.16	68.77	83.17	0.986	0.951	0.981	0.589
50	5	72.39	69.48	62.03	65.50	0.987	0.973	0.981	0.727
100	5	50.48	49.97	49.29	47.82	0.989	0.988	0.984	0.916
200	5	43.44	43.37	43.29	42.25	0.990	0.990	0.987	0.965
400	5	40.47	40.41	40.40	39.83	0.990	0.990	0.988	0.980
800	5	39.00	38.97	38.97	38.68	0.990	0.990	0.989	0.985
24	10	1379.64	866.39	64.03	–	0.978	0.401	–	0.155
25	10	737.05	502.77	63.43	677.89	0.979	0.488	0.988	0.202
30	10	179.42	156.96	60.77	169.46	0.984	0.762	0.988	0.414
40	10	84.31	82.97	56.81	77.26	0.989	0.924	0.984	0.664
50	10	66.62	66.21	54.00	61.67	0.990	0.961	0.983	0.780
100	10	48.87	48.66	47.05	46.79	0.990	0.986	0.985	0.927
200	10	43.07	42.93	42.70	41.96	0.990	0.989	0.987	0.967
400	10	40.32	40.28	40.24	39.75	0.990	0.990	0.988	0.980
800	10	38.93	38.94	38.93	38.66	0.990	0.990	0.989	0.986

Note. $\text{CP}_{\text{YS}\cdot\text{L1}} = \text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha))$, $\text{CP}_{\text{YS}\cdot F1} = \text{CP}(t^2_{\text{YS}\cdot F1}(\alpha))$, $\text{CP}_{\text{KP}} = \text{CP}(t^2_{\text{KP}}(\alpha))$, $\text{CP}_{\chi^2} = \text{CP}(\chi^2_p(\alpha))$, $\chi^2_{20}(0.05) = 31.41$, $\chi^2_{20}(0.01) = 37.57$.

and Richards, 2013). Tables 2.5 and 2.6 give the simulated upper 100α percentiles of the $\widetilde{T}^2$ statistic with five-step and ten-step monotone missing data patterns. That is, we provide $\tilde{t}^2_{\text{simu}} (= \tilde{t}^2_{\text{simu}}(\alpha))$ for the following cases:

Five-step Case: $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$,

$n_1 = 19, 20(10)50, 100, 200, 400, 800,\ n_2 = n_3 = \cdots = n_5 = 5, 10,$

$\alpha = 0.05, 0.01.$

Ten-step Case: $(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}) = (20, 18, 16, 14, 12, 10, 8, 6, 4, 2)$,
$n_1 = 24, 25, 30(10)50, 100, 200, 400, 800,\ n_2 = n_3 = \cdots = n_{10} = 5, 10$,
$\alpha = 0.05, 0.01$.

These tables also give the approximations to the upper 100α percentiles of the $\widetilde{T}^2$ statistic, that is, $t^2_{\text{YS}\cdot\text{L1}}(= t^2_{\text{YS}\cdot\text{L1}}(\alpha))$ and $t^2_{\text{YS}\cdot F1}(= t^2_{\text{YS}\cdot F1}(\alpha))$ in Section 2.2.2, and $t^2_{\text{KP}}(= t^2_{\text{KP}}(\alpha))$. We denote $t^2_{\text{KP}}(\alpha)$ as the approximation in Krishnamoorthy and Pannala (1999). In addition, we provide the actual coverage probabilities for the approximate upper 100α percentiles in Tables 2.5 and 2.6, which are given by

$$\text{CP}(t^2_{\text{YS}\cdot\text{L1}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YS}\cdot\text{L1}}(\alpha)\}, \quad \text{CP}(t^2_{\text{YS}\cdot F1}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YS}\cdot F1}(\alpha)\},$$
$$\text{CP}(t^2_{\text{KP}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{KP}}(\alpha)\}, \quad \text{CP}(\chi^2_p(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > \chi^2_p(\alpha)\}.$$

It may be noted from Tables 2.5 and 2.6 that the simulated values, $\widetilde{t}^2_{\text{simu}}(\alpha)$, are closer to the upper percentiles of the χ^2 distribution when the sample size n_1 becomes large. However, the upper 100α percentiles of the χ^2 distribution, $\chi^2_p(\alpha)$, are not good approximations to those of the $\widetilde{T}^2$ statistic even for moderately large sample sizes. At the same time, the proposed approximate upper percentiles $t^2_{\text{YS}\cdot\text{L1}}$ and $t^2_{\text{YS}\cdot F1}$ as well as t^2_{KP} are good for moderately large sample sizes; in particular, $t^2_{\text{YS}\cdot\text{L1}}$ is considerably good when n_1 is greater than 40. We note that the condition that $t^2_{\text{YS}\cdot\text{L1}}$ and $t^2_{\text{YS}\cdot F1}$ can be defined is $n_1 - 1 \geq p_1$ and the condition for t^2_{KP} is $n_1 - 5 \geq p_1$. For example, when $p_1 = 15$ and $n_1 = 19$ in Table 2.5, the value of t^2_{KP} cannot be computed but that of $t^2_{\text{YS}\cdot\text{L1}}$ or $t^2_{\text{YS}\cdot F1}$ can be computed.

2.3 Conclusions

In conclusion, we have developed the approximate upper percentiles of the simplified T^2-type statistic for testing a mean vector with monotone missing data. We presented the numerical comparisons of our approximation proposed in this paper with the approximation using F distribution in Krishnamoorthy and Pannala (1999). The proposed approximate values as well as the approximation by Krishnamoorthy and Pannala (1999) can be calculated easily and the approximations are considerably better than the χ^2 approximation even when the sample size is small.

Chapter 3

Testing the equality of mean vectors and simultaneous confidence intervals

In this chapter, we consider the problem of testing the equality of two mean vectors when the data have a three-step (Section 3.1) or general k-step (Section 3.2) monotone pattern of missing observations. We give a simplified T^2-type statistic and propose an approximate upper percentile of the statistic where each data set has a three-step or k-step monotone missing data pattern and the population covariance matrices are equal. Further, we obtain the Hotelling's T^2-type and simplified T^2-type statistics and their approximate upper percentiles in the case of data with unequal two-step monotone missing data patterns. We also consider multivariate multiple comparisons for mean vectors. Approximate simultaneous confidence intervals for pairwise comparisons among mean vectors and comparisons with a control are obtained using Bonferroni's approximate upper percentiles of the $T^2_{\mathrm{max\cdot p}}$ and $T^2_{\mathrm{max\cdot c}}$ statistics, respectively. Finally, the accuracy of the approximations is investigated via Monte Carlo simulation.

Section 3.1 extends the one-sample problem investigated by Yagi and Seo (2014) (see, Section 2.1) to the two-sample or m-sample problem in the case of three-step monotone missing data. In other words, we use the concepts of Yagi and Seo (2014) to develop an approximate upper percentile of the simplified T^2-type statistic for the two-sample problem. In Section 3.1, we consider the problem of testing the equality of two mean vectors when the data have a three-step monotone pattern of missing observations. In the case of the two-step monotone missing data pattern, Seko et al. (2011) derived a Hotelling's T^2-type statistic, the likelihood ratio test statistic and their approximate upper percentiles. In addition, Yu, Krishnamoorthy and Pannala (2006) gave the results of two-sample case similar to that of one-sample case in Krishnamoorthy and Pannala (1999). That is, they derived the simplified T^2-type statistic and its approximate distribution using another approach. Recently, Seko (2012) discussed tests for mean vectors with two-step monotone missing data for the m-sample problem.

Section 3.1 is a summary of Yagi and Seo (2015b) and organized as follows. In Section 3.1.1, we present some preliminary notations and the MLEs of the mean vector and the covariance matrix for the m-sample problem that includes the two-sample problem. In

Section 3.1.2, we present a simplified T^2-type statistic to test the equality of two mean vectors and its approximate upper percentiles in the case of three-step monotone missing data. In Section 3.1.3, we discuss the Hotelling's T^2-type statistic and the simplified T^2-type statistic when two data sets have different two-step monotone missing data patterns. In Section 3.1.4, we present approximate simultaneous confidence intervals for multiple comparisons among mean vectors under the two-sample or m-sample problem. In order to obtain the simultaneous confidence intervals, we derive approximate upper percentiles of the $T^2_{\max}$-type statistics via Bonferroni's approximation. Finally, in Section 3.1.5, we present some simulation results.

In Section 3.2, we consider testing the equality of mean vectors in the case of general k-step monotone missing data, which is summarized using the parts of Yagi and Seo (2017). In Section 3.2.1, some preliminary notations are presented, and the MLEs of the mean vectors, and the common covariance matrix for the m-sample problem are obtained in the case of k-step monotone missing data. In Section 3.2.2, we give the simplified T^2-type statistic to test the equality of two mean vectors and their approximate upper percentiles in the case of k-step monotone missing data. Under the m-sample problem ($m \geq 3$), we give approximate simultaneous confidence intervals for multiple comparisons among mean vectors in Section 3.2.3. Finally, we give the simulation results in Section 3.2.4 and state our conclusions in Section 3.3.

3.1 Two-step or three-step monotone missing data

Consider the problem of testing the equality of mean vectors with two-step or three-step monotone missing data.

3.1.1 Three-step monotone missing data and MLE

As preliminaries, we present some notations for the vector and matrix needed to express the three-step monotone missing data for the general m-sample problem. Using the assumptions and notations in Section 2.1, we consider the MLEs of the mean vectors and the common covariance matrix for the m-sample problem.

Let $\boldsymbol{x}_{i1}^{(\ell)}, \boldsymbol{x}_{i2}^{(\ell)}, \ldots, \boldsymbol{x}_{in_i^{(\ell)}}^{(\ell)}$ be distributed as $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$ for $i = 1, 2, 3$ and $\ell = 1, 2, \ldots, m$, where $\boldsymbol{\mu}_i^{(\ell)} = (\mu_1^{(\ell)}, \mu_2^{(\ell)}, \ldots, \mu_{p_i}^{(\ell)})'$ and $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ covariance matrix, where $p = p_1 > p_2 > p_3 > 0$, $\nu_1 - m \geq p$ and $\nu_1 = \sum_{\ell=1}^{m} n_1^{(\ell)}$. Let

$$\overline{\boldsymbol{x}}_i^{(\ell)} = \frac{1}{n_i^{(\ell)}} \sum_{j=1}^{n_i^{(\ell)}} \boldsymbol{x}_{ij}^{(\ell)}, \quad \boldsymbol{E}_i^{(\ell)} = \sum_{j=1}^{n_i^{(\ell)}} (\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i^{(\ell)})(\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i^{(\ell)})', \ i = 1, 2, 3.$$

Then the MLEs of $\boldsymbol{\mu}^{(\ell)}$ and $\boldsymbol{\Sigma}$ are given in the following theorem.

Theorem 3.1 (Yagi and Seo, 2015b) *Let $\boldsymbol{x}_{ij}^{(\ell)}$, $i = 1, 2, 3$, $j = 1, 2, \ldots, n_i^{(\ell)}$, $\ell = 1, 2, \ldots, m$ be the j-th random vector of the i-th step from the ℓ-th population distributed* *as* $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$. *Then, the MLEs of* $\boldsymbol{\mu}^{(\ell)}$, $\ell = 1, 2, \ldots, m$ *are given by*

$$\widehat{\boldsymbol{\mu}}^{(\ell)} = \overline{\boldsymbol{x}}_1^{(\ell)} + \widehat{\boldsymbol{T}}_2^{[pl]}\boldsymbol{d}_2^{(\ell)} + \widehat{\boldsymbol{T}}_2^{[pl]}\widehat{\boldsymbol{T}}_3^{[pl]}\boldsymbol{d}_3^{(\ell)},$$

where

$$\boldsymbol{d}_2^{(\ell)} = \frac{n_2^{(\ell)}}{N_3^{(\ell)}}\left[\overline{\boldsymbol{x}}_2^{(\ell)} - (\overline{\boldsymbol{x}}_1^{(\ell)})_2\right], \ \boldsymbol{d}_3^{(\ell)} = \frac{n_3^{(\ell)}}{N_4^{(\ell)}}\left[\overline{\boldsymbol{x}}_3^{(\ell)} - \frac{1}{N_3^{(\ell)}}\left\{n_1^{(\ell)}(\overline{\boldsymbol{x}}_1^{(\ell)})_3 + n_2^{(\ell)}(\overline{\boldsymbol{x}}_2^{(\ell)})_3\right\}\right],$$

$$N_1^{(\ell)} = 0, \ N_{i+1}^{(\ell)} = \sum_{j=1}^{i} n_j^{(\ell)}, \ \widehat{\boldsymbol{T}}_2^{[pl]} = \begin{pmatrix} \boldsymbol{I}_{p_2} \\ \widehat{\boldsymbol{\Sigma}}_{(1,2)}^{[pl]\prime}\widehat{\boldsymbol{\Sigma}}_2^{[pl]-1} \end{pmatrix}, \ \widehat{\boldsymbol{T}}_3^{[pl]} = \begin{pmatrix} \boldsymbol{I}_{p_3} \\ \widehat{\boldsymbol{\Sigma}}_{(2,2)}^{[pl]\prime}\widehat{\boldsymbol{\Sigma}}_3^{[pl]-1} \end{pmatrix},$$

and then, the MLE of $\boldsymbol{\Sigma}$ *is given by*

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}}^{[pl]} &= \frac{1}{M_2}\sum_{\ell=1}^{m}\boldsymbol{E}_1^{(\ell)} + \frac{1}{M_3}\left[\sum_{\ell=1}^{m}\boldsymbol{F}_2^{[pl]}\left\{\boldsymbol{E}_2^{(\ell)} + \frac{N_2^{(\ell)}N_3^{(\ell)}}{n_2^{(\ell)}}\boldsymbol{d}_2^{(\ell)}\boldsymbol{d}_2^{(\ell)\prime} - \frac{\nu_2}{M_2}\boldsymbol{L}_{11}^{(\ell)}\right\}\boldsymbol{F}_2^{[pl]\prime}\right] \\ &\quad + \frac{1}{M_4}\left[\sum_{\ell=1}^{m}\boldsymbol{F}_3^{[pl]}\left\{\boldsymbol{E}_3^{(\ell)} + \frac{N_3^{(\ell)}N_4^{(\ell)}}{n_3^{(\ell)}}\boldsymbol{d}_3^{(\ell)}\boldsymbol{d}_3^{(\ell)\prime} - \frac{\nu_3}{M_3}\boldsymbol{L}_{21}^{(\ell)}\right\}\boldsymbol{F}_3^{[pl]\prime}\right],\end{aligned}$$

where

$$\begin{aligned}
&M_{i+1} = \sum_{\ell=1}^{m} N_{i+1}^{(\ell)}, \ i = 1, 2, 3, \ \ \nu_i = \sum_{\ell=1}^{m} n_i^{(\ell)}, \ i = 2, 3, \\
&\boldsymbol{F}_2^{[pl]} = \boldsymbol{G}_2^{[pl]}, \ \ \boldsymbol{F}_3^{[pl]} = \boldsymbol{G}_2^{[pl]}\boldsymbol{G}_3^{[pl]}, \\
&\boldsymbol{G}_2^{[pl]} = \begin{pmatrix} \boldsymbol{I}_{p_2} \\ \left(\sum_{\ell=1}^{m}\boldsymbol{L}_{12}^{(\ell)}\right)'\left(\sum_{\ell=1}^{m}\boldsymbol{L}_{11}^{(\ell)}\right)^{-1} \end{pmatrix}, \ \ \boldsymbol{G}_3^{[pl]} = \begin{pmatrix} \boldsymbol{I}_{p_3} \\ \left(\sum_{\ell=1}^{m}\boldsymbol{L}_{22}^{(\ell)}\right)'\left(\sum_{\ell=1}^{m}\boldsymbol{L}_{21}^{(\ell)}\right)^{-1} \end{pmatrix}, \\
&\boldsymbol{L}_1^{(\ell)} = \boldsymbol{E}_1^{(\ell)}, \ \ \boldsymbol{L}_2^{(\ell)} = \boldsymbol{L}_{11}^{(\ell)} + \boldsymbol{E}_2^{(\ell)} + \frac{N_2^{(\ell)}N_3^{(\ell)}}{n_2^{(\ell)}}\boldsymbol{d}_2^{(\ell)}\boldsymbol{d}_2^{(\ell)\prime}, \\
&\boldsymbol{L}_{i1}^{(\ell)} = (\boldsymbol{L}_i^{(\ell)})_{i+1}, \ \ \boldsymbol{L}_i^{(\ell)} = \begin{pmatrix} \boldsymbol{L}_{i1}^{(\ell)} & \boldsymbol{L}_{i2}^{(\ell)} \\ \boldsymbol{L}_{i2}^{(\ell)\prime} & \boldsymbol{L}_{i3}^{(\ell)} \end{pmatrix}, \ i = 1, 2.
\end{aligned}$$

We note that the result of Theorem 3.1 can be obtained in a straightforward manner based on the case of the one-sample problem investigated in Section 2.1. In the next section, in order to obtain a simplified T^2-type statistic to test the equality of two mean vectors, we present the covariance matrix of $\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}$, where $\widetilde{\boldsymbol{\mu}}^{(\ell)} = \overline{\boldsymbol{x}}_1^{(\ell)} + \boldsymbol{T}_2\boldsymbol{d}_2^{(\ell)} + \boldsymbol{T}_2\boldsymbol{T}_3\boldsymbol{d}_3^{(\ell)}$. The simplified T^2-type statistic with $\widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}]$ instead of $\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$ is adopted because $\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$ is complicated for more than three-step case whereas $\widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}]$ is asymptotically equivalent to $\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$.

3.1.2 Simplified T^2-type statistic

In this section, we consider the hypothesis test

$$H_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} \quad \text{vs.} \quad H_1 : \boldsymbol{\mu}^{(1)} \neq \boldsymbol{\mu}^{(2)}$$

when two data sets have the same three-step monotone missing data pattern. In order to test the hypothesis H_0, under the assumption of a common population covariance matrix, we adopt the simplified T^2-type statistic given by

$$\widetilde{T}^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \, \widetilde{\boldsymbol{\Gamma}}^{[pl]^{-1}} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}),$$

where $\widehat{\boldsymbol{\mu}}^{(\ell)} = \overline{\boldsymbol{x}}_1^{(\ell)} + \widehat{\boldsymbol{T}}_2^{[pl]} \boldsymbol{d}_2^{(\ell)} + \widehat{\boldsymbol{T}}_2^{[pl]} \widehat{\boldsymbol{T}}_3^{[pl]} \boldsymbol{d}_3^{(\ell)}, \ \ell = 1, 2$, and $\widetilde{\boldsymbol{\Gamma}}^{[pl]}$ is an estimator of Cov$[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}]$, $\widetilde{\boldsymbol{\mu}}^{(\ell)} = \overline{\boldsymbol{x}}_1^{(\ell)} + \boldsymbol{T}_2 \boldsymbol{d}_2^{(\ell)} + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3^{(\ell)}, \ \ell = 1, 2$. Then, using the result for the one-sample problem in Section 2.1, we have

$$\widetilde{\boldsymbol{\Gamma}}^{[pl]} = \left(\sum_{\ell=1}^{2} \frac{1}{N_2^{(\ell)}} \right) \widehat{\boldsymbol{\Sigma}}_1^{[pl]} - \left(\sum_{\ell=1}^{2} \frac{n_2^{(\ell)}}{N_2^{(\ell)} N_3^{(\ell)}} \right) \widehat{\boldsymbol{U}}_2^{[pl]} - \left(\sum_{\ell=1}^{2} \frac{n_3^{(\ell)}}{N_3^{(\ell)} N_4^{(\ell)}} \right) \widehat{\boldsymbol{U}}_3^{[pl]},$$

where $\widehat{\boldsymbol{\Sigma}}^{[pl]}$ is the MLE for $m = 2$ in Theorem 3.1,

$$\widehat{\boldsymbol{U}}_2^{[pl]} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_2^{[pl]} \\ \widehat{\boldsymbol{\Sigma}}_{22}^{[pl]'} \end{pmatrix} \widehat{\boldsymbol{T}}_2^{[pl]'}, \ \widehat{\boldsymbol{U}}_3^{[pl]} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_3^{[pl]} \\ \widehat{\boldsymbol{\Sigma}}_{32}^{[pl]'} \end{pmatrix} \widehat{\boldsymbol{T}}_3^{[pl]'} \widehat{\boldsymbol{T}}_2^{[pl]'}.$$

We note that under H_0, $\widetilde{T}^2$ is asymptotically distributed as a χ^2 distribution with p degrees of freedom when $n_1^{(\ell)}, N_4^{(\ell)} \to \infty$ with $n_1^{(\ell)}/N_4^{(\ell)} \to \delta^{(\ell)} \in (0, 1], \ \ell = 1, 2$. However, note that the χ^2 approximation is not an accurate approximate upper percentile of the $\widetilde{T}^2$ statistic when the sample size is small. Then, the two approximate upper 100α percentiles of the $\widetilde{T}^2$ statistic when two data sets have the same three-step monotone missing data pattern are given by

$$t^2_{\text{YS}\cdot\text{L}2}(\alpha) = (1 - d) T^2_{n,\alpha} + d T^2_{N,\alpha},$$
$$t^2_{\text{YS}\cdot F2}(\alpha) = \frac{n^* p_1}{n^* - p_1 - 1} F_{p_1, n^* - p_1 - 1, \alpha},$$

where

$$d = \frac{\sum_{\ell=1}^2 \sum_{i=2}^3 n_i^{(\ell)} p_i}{p_1 \sum_{\ell=1}^2 \sum_{i=2}^3 n_i^{(\ell)}}, \ T^2_{n,\alpha} = \frac{n p_1}{n - p_1 - 1} F_{p_1, n - p_1 - 1, \alpha}, \ T^2_{N,\alpha} = \frac{N p_1}{N - p_1 - 1} F_{p_1, N - p_1 - 1, \alpha},$$
$$n = n_1^{(1)} + n_1^{(2)}, \ N = N_4^{(1)} + N_4^{(2)}, \ n^* = \frac{1}{p_1} \sum_{\ell=1}^{2} \sum_{i=1}^{3} n_i^{(\ell)} p_i$$

and $F_{p,q,\alpha}$ is the upper 100α percentile of the F distribution with p and q degrees of freedom.

For the $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ values, $T^2_{N,\alpha}$ and $T^2_{n,\alpha}$ are calculated from the complete data sets $((N_4^{(1)} \times p_1)+(N_4^{(2)} \times p_1))$ and $((n_1^{(1)} \times p_1)+(n_1^{(2)} \times p_1))$, respectively. We note that the upper percentiles of $\widetilde{T}^2$ may lie between $T^2_{N,\alpha}$ and $T^2_{n,\alpha}$. Then, we propose $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ as an approximate upper percentile of $\widetilde{T}^2$ by linear interpolation of the coordinates $((n_1^{(1)}+n_1^{(2)})p_1, T^2_{n,\alpha})$ and $((N_4^{(1)}+N_4^{(2)})p_1, T^2_{N,\alpha})$. This approach is essentially based on that adopted in Section 2.1. As another approximation procedure, we propose $t^2_{\mathrm{YS\cdot}F2}(\alpha)$ by adjusting the degrees of freedom of the F distribution. If two data sets have the same type of two-step monotone missing data pattern, then the approximation adjusting the degrees of freedom of the F distribution are given in Theorem 10 of Yagi and Seo (2015a).

3.1.3 Simplified T^2-type and Hotelling's T^2-type statistics with different two-step monotone missing data patterns

In this section, we test the equality of two mean vectors when two data sets have different two-step monotone missing data patterns. That is, two data sets Π_1 and Π_2 are of the forms given in Figure 3.1. Then, we can reduce Theorem 3.1 to the following result.

Corollary 3.2 (Yagi and Seo, 2015b) *Let* $\boldsymbol{x}_{ij}^{(\ell)}$, $i=1,2,3$, $j=1,2,\ldots,n_i^{(\ell)}$, $\ell=1,2$ *be the j-th random vector of the i-th step from the ℓ-th population distributed as* $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$, *where* $n_3^{(1)}=0$, $n_2^{(2)}=0$. *Then, the MLEs of* $\boldsymbol{\mu}^{(\ell)}$, $\ell=1,2$ *and* $\boldsymbol{\Sigma}$ *are given by*

$$\begin{aligned}
\widehat{\boldsymbol{\mu}}^{(1)} &= \overline{\boldsymbol{x}}_1^{(1)} + \widehat{\boldsymbol{T}}_2^{[pl]} \boldsymbol{d}_2^{(1)}, \\
\widehat{\boldsymbol{\mu}}^{(2)} &= \overline{\boldsymbol{x}}_1^{(2)} + \widehat{\boldsymbol{T}}_2^{[pl]} \widehat{\boldsymbol{T}}_3^{[pl]} \boldsymbol{d}_3^{(2)}, \\
\widehat{\boldsymbol{\Sigma}}^{[pl]} &= \frac{1}{N_2^{(1)}+N_2^{(2)}} \sum_{\ell=1}^{2} \boldsymbol{E}_1^{(\ell)} \\
&\quad + \frac{1}{N_3^{(1)}+N_3^{(2)}} \boldsymbol{F}_2^{[pl]} \left\{ \boldsymbol{E}_2^{(1)} + \frac{N_2^{(1)} N_3^{(1)}}{n_2^{(1)}} \boldsymbol{d}_2^{(1)} \boldsymbol{d}_2^{(1)\prime} - \frac{n_2^{(1)}}{N_2^{(1)}+N_2^{(2)}} (\boldsymbol{L}_{11}^{(1)} + \boldsymbol{L}_{11}^{(2)}) \right\} \boldsymbol{F}_2^{[pl]\prime} \\
&\quad + \frac{1}{N_4^{(1)}+N_4^{(2)}} \boldsymbol{F}_3^{[pl]} \left\{ \boldsymbol{E}_3^{(2)} + \frac{N_3^{(2)} N_4^{(2)}}{n_3^{(2)}} \boldsymbol{d}_3^{(2)} \boldsymbol{d}_3^{(2)\prime} - \frac{n_3^{(2)}}{N_3^{(1)}+N_3^{(2)}} (\boldsymbol{L}_{21}^{(1)} + \boldsymbol{L}_{21}^{(2)}) \right\} \boldsymbol{F}_3^{[pl]\prime}.
\end{aligned}$$

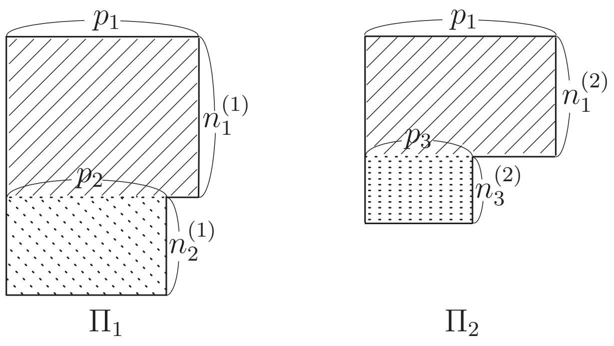

Figure 3.1: Different two-step monotone missing data patterns

Further, the simplified T^2-type statistic is reduced to

$$\widetilde{T}^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \, \widetilde{\boldsymbol{\Gamma}}^{[pl]^{-1}} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}),$$

where

$$\begin{aligned}
\widetilde{\boldsymbol{\mu}}^{(1)} &= \overline{\boldsymbol{x}}_1^{(1)} + \boldsymbol{T}_2 \boldsymbol{d}_2^{(1)}, \quad \widetilde{\boldsymbol{\mu}}^{(2)} = \overline{\boldsymbol{x}}_1^{(2)} + \boldsymbol{T}_2 \boldsymbol{T}_3 \boldsymbol{d}_3^{(2)}, \\
\widetilde{\boldsymbol{\Gamma}}^{[pl]} &= \left(\sum_{\ell=1}^{2} \frac{1}{N_2^{(\ell)}} \right) \widehat{\boldsymbol{\Sigma}}_1^{[pl]} - \frac{n_2^{(1)}}{N_2^{(1)} N_3^{(1)}} \widehat{\boldsymbol{U}}_2^{[pl]} - \frac{n_3^{(2)}}{N_3^{(2)} N_4^{(2)}} \widehat{\boldsymbol{U}}_3^{[pl]}, \\
\widehat{\boldsymbol{U}}_2^{[pl]} &= \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_2^{[pl]} \\ \widehat{\boldsymbol{\Sigma}}_{22}^{[pl]'} \end{pmatrix} \widehat{\boldsymbol{T}}_2^{[pl]'}, \quad \widehat{\boldsymbol{U}}_3^{[pl]} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_3^{[pl]} \\ \widehat{\boldsymbol{\Sigma}}_{32}^{[pl]'} \end{pmatrix} \widehat{\boldsymbol{T}}_3^{[pl]'} \widehat{\boldsymbol{T}}_2^{[pl]'},
\end{aligned}$$

and $\widehat{\boldsymbol{\Sigma}}^{[pl]}$ is the MLE in Corollary 3.2. Indeed, Kanda and Fujikoshi (1998) obtained $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}]$ with two-step monotone missing data under the one-sample problem. Under the two-sample problem, we can easily obtain $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$ with different two-step monotone missing data: therefore, we can obtain the Hotelling's T^2-type statistic as

$$T^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \, \widehat{\boldsymbol{\Gamma}}^{[pl]^{-1}} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}),$$

where

$$\begin{aligned}
&\widehat{\boldsymbol{\mu}}^{(1)} = \overline{\boldsymbol{x}}_1^{(1)} + \widehat{\boldsymbol{T}}_2^{[pl]} \boldsymbol{d}_2^{(1)}, \ \widehat{\boldsymbol{\mu}}^{(2)} = \overline{\boldsymbol{x}}_1^{(2)} + \widehat{\boldsymbol{T}}_2^{[pl]} \widehat{\boldsymbol{T}}_3^{[pl]} \boldsymbol{d}_3^{(2)}, \quad \widehat{\boldsymbol{\Gamma}}^{[pl]} = \sum_{\ell=1}^{2} \left(\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}^{(\ell)}] + \widehat{\boldsymbol{R}}_{(\ell)}^{[pl]} \right), \\
&\widehat{\boldsymbol{R}}_{(1)}^{[pl]} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \dfrac{n_2^{(1)} p_2}{N_2^{(1)} N_3^{(1)} (N_2^{(1)} - p_2 - 2)} \widehat{\boldsymbol{\Sigma}}_{23\cdot 2}^{[pl]} \end{pmatrix}, \ \widehat{\boldsymbol{\Sigma}}_{23\cdot 2}^{[pl]} = \widehat{\boldsymbol{\Sigma}}_{23}^{[pl]} - \widehat{\boldsymbol{\Sigma}}_{22}^{[pl]'} (\widehat{\boldsymbol{\Sigma}}_2^{[pl]})^{-1} \widehat{\boldsymbol{\Sigma}}_{22}^{[pl]}, \\
&\widehat{\boldsymbol{R}}_{(2)}^{[pl]} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \dfrac{n_3^{(2)} p_3}{N_2^{(2)} N_4^{(2)} (N_2^{(2)} - p_3 - 2)} \widehat{\boldsymbol{\Sigma}}_{33\cdot 3}^{[pl]} \end{pmatrix}, \ \widehat{\boldsymbol{\Sigma}}_{33\cdot 3}^{[pl]} = \widehat{\boldsymbol{\Sigma}}_{33}^{[pl]} - \widehat{\boldsymbol{\Sigma}}_{32}^{[pl]'} (\widehat{\boldsymbol{\Sigma}}_3^{[pl]})^{-1} \widehat{\boldsymbol{\Sigma}}_{32}^{[pl]}, \\
&N_2^{(1)} > p_2 + 2, \ N_2^{(2)} > p_3 + 2.
\end{aligned}$$

For details on the case in which two data sets have the same two-step monotone missing data pattern, see Yagi and Seo (2015a). Further, in the case of data with different two-step monotone missing data patterns, the two approximate upper 100α percentiles of the T^2 (or $\widetilde{T}^2$) statistic are given by

$$\begin{aligned}
t^2_{\mathrm{YS\cdot L2}}(\alpha) &= (1-d) T^2_{n,\alpha} + d T^2_{N,\alpha}, \\
t^2_{\mathrm{YS\cdot}F2}(\alpha) &= \frac{n^* p_1}{n^* - p_1 - 1} F_{p_1, n^* - p_1 - 1, \alpha},
\end{aligned}$$

where

$$d = \frac{n_2^{(1)} p_2 + n_3^{(2)} p_3}{\left(n_2^{(1)} + n_3^{(2)}\right) p_1},\ T_{n,\alpha}^2 = \frac{np_1}{n - p_1 - 1} F_{p_1, n-p_1-1,\alpha},$$

$$T_{N,\alpha}^2 = \frac{Np_1}{N - p_1 - 1} F_{p_1, N-p_1-1,\alpha},\ n = n_1^{(1)} + n_1^{(2)},$$

$$N = n_1^{(1)} + n_2^{(1)} + n_1^{(2)} + n_3^{(2)},\ n^* = \frac{1}{p_1}\left(p_1 \sum_{\ell=1}^{2} n_1^{(\ell)} + n_2^{(1)} p_2 + n_3^{(2)} p_3\right).$$

Next, under the two-sample problem, we consider the simultaneous confidence intervals when each data set has three-step monotone missing observations.

For any nonnull vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)'$, the simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$ with the confidence level $(1 - \alpha)$ are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) - \sqrt{L} \leq \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}) \leq \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) + \sqrt{L},\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L = t^2(\alpha)\boldsymbol{c}'\widehat{\boldsymbol{\Gamma}}^{[pl]}\boldsymbol{c}$, $t^2(\alpha)$ is the upper 100α percentile of the $T^2(= (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\boldsymbol{\Gamma}}^{[pl]^{-1}}$ $(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}))$ statistic, and $\widehat{\boldsymbol{\Gamma}}^{[pl]}$ is an estimator of $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$. However, it is not easy to obtain $t^2(\alpha)$. Therefore, using the approximate upper percentiles of the $\widetilde{T}^2$ statistic, $t^2_{\mathrm{YS}\cdot\mathrm{L}2}(\alpha)$ or $t^2_{\mathrm{YS}\cdot F2}(\alpha)$, for any nonnull vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)'$, the approximate simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$ are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) - \sqrt{L_{\mathrm{app}}} \leq \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}) \leq \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) + \sqrt{L_{\mathrm{app}}},\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L_{\mathrm{app}} = t^2_{\mathrm{app}}(\alpha)\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}^{[pl]}\boldsymbol{c}$ and the value of $t^2_{\mathrm{app}}(\alpha)$ is $t^2_{\mathrm{YS}\cdot\mathrm{L}2}(\alpha)$ or $t^2_{\mathrm{YS}\cdot F2}(\alpha)$.

3.1.4 Simultaneous confidence intervals for multiple comparisons among mean vectors

Under the m-sample problem, we consider the simultaneous confidence intervals for pairwise multiple comparisons among mean vectors when each data set has three-step monotone missing observations. Similarly, we also construct the simultaneous confidence intervals for comparisons with a control. Let $\boldsymbol{x}_{i1}^{(\ell)}, \boldsymbol{x}_{i2}^{(\ell)}, \ldots, \boldsymbol{x}_{in_i^{(\ell)}}^{(\ell)}$ be distributed as $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$ for $i = 1, 2, 3$ and $\ell = 1, 2, \ldots, m$. Further, we define the $T^2_{\max\cdot\mathrm{p}}$ statistic as

$$T^2_{\max\cdot\mathrm{p}} = \max_{1 \leq a < b \leq m} T^2_{ab},$$

where $T^2_{ab} = (\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})'\widehat{\boldsymbol{\Gamma}}_{ab}^{[pl]^{-1}}(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})$, and $\widehat{\boldsymbol{\Gamma}}_{ab}^{[pl]}$ is an estimator of $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}]$. Then, for the case of pairwise multiple comparisons, the simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(a)} - \boldsymbol{\mu}^{(b)})$, $1 \leq a < b \leq m$ are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L_{\mathrm{p}}} \leq \boldsymbol{c}'(\boldsymbol{\mu}^{(a)} - \boldsymbol{\mu}^{(b)}) \leq \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L_{\mathrm{p}}},$$
$$1 \leq a < b \leq m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L_{\rm p} = t^2_{\max\cdot{\rm p}}(\alpha)\boldsymbol{c}'\widehat{\boldsymbol{\Gamma}}^{[pl]}_{ab}\boldsymbol{c}$ and $t^2_{\max\cdot{\rm p}}(\alpha)$ is the upper percentile of the $T^2_{\max\cdot{\rm p}}$ statistic.

Similarly, for the case of comparisons with a control, let $\boldsymbol{\mu}^{(1)}$ be a control and define the $T^2_{\max\cdot{\rm c}}$ statistic as

$$T^2_{\max\cdot{\rm c}} = \max_{2\le b\le m} T^2_{1b},$$

where $T^2_{1b} = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})'\widehat{\boldsymbol{\Gamma}}^{[pl]^{-1}}_{1b}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})$ and $\widehat{\boldsymbol{\Gamma}}^{[pl]}_{1b}$ is an estimator of ${\rm Cov}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}]$. Then, the simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(b)})$, $2 \le b \le m$ are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L_{\rm c}} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(b)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L_{\rm c}},$$
$$2 \le b \le m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L_{\rm c} = t^2_{\max\cdot{\rm c}}(\alpha)\boldsymbol{c}'\widehat{\boldsymbol{\Gamma}}^{[pl]}_{1b}\boldsymbol{c}$ and $t^2_{\max\cdot{\rm c}}(\alpha)$ is the upper percentile of the $T^2_{\max\cdot{\rm c}}$ statistic. However, it is not easy to obtain $t^2_{\max\cdot{\rm p}}(\alpha)$ and $t^2_{\max\cdot{\rm c}}(\alpha)$ even under non-missing multivariate normality (see Seo and Siotani, 1992; Seo, Mano and Fujikoshi, 1994). Therefore, in this paper, we adopt Bonferroni's approximation which is one of the solutions to this problem. Let $n^{(1)}_i = n^{(2)}_i = \cdots = n^{(m)}_i$, $i = 1, 2, 3$: then, the null distributions of T^2_{ab} or T^2_{1b} are identical. Therefore, the approximate simultaneous confidence intervals for pairwise comparisons and comparisons with a control are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L^*_{\rm p}} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(a)} - \boldsymbol{\mu}^{(b)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L^*_{\rm p}},$$
$$1 \le a < b \le m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

and

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L^*_{\rm c}} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(b)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L^*_{\rm c}},$$
$$2 \le b \le m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

respectively, where

$$L^*_{\rm p} = t^2_{\rm Bon}(\alpha_1)\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}^{[pl]}_{ab}\boldsymbol{c},\ L^*_{\rm c} = t^2_{\rm Bon}(\alpha_2)\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}^{[pl]}_{1b}\boldsymbol{c},$$

the value of $t^2_{\rm Bon}(\alpha_i)$ is $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_i)$ or $t^2_{{\rm YS}\cdot Fm}(\alpha_i)$, $i = 1, 2$, and

$$\alpha_1 = \frac{2\alpha}{m(m-1)},\ \alpha_2 = \frac{\alpha}{m-1}.$$

We note that $\widetilde{\boldsymbol{\Gamma}}^{[pl]}_{ab}$ and $\widetilde{\boldsymbol{\Gamma}}^{[pl]}_{1b}$ are estimated by the use of $\widehat{\boldsymbol{\Sigma}}^{[pl]}$ in Theorem 3.1, and $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_i)$ or $t^2_{{\rm YS}\cdot Fm}(\alpha_i)$ is the approximation that extended $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha)$ or $t^2_{{\rm YS}\cdot Fm}(\alpha)$ in Section 3.1.2 to the case of m-sample problem. That is, $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_1)$, $t^2_{{\rm YS}\cdot Fm}(\alpha_1)$, $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_2)$, and $t^2_{{\rm YS}\cdot Fm}(\alpha_2)$ are $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_{\rm p})$, $t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm p})$, $t^2_{{\rm YS}\cdot{\rm L}m}(\alpha_{\rm c})$, and $t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm c})$ in Section 3.2.3 when $k = 3$, respectively.

3.1.5 Simulation studies

In order to investigate the accuracy of some of the approximations, we compute the upper percentiles of the T^2, $\widetilde{T}^2$, $\widetilde{T}^2_{\max\cdot\mathrm{p}}$ and $\widetilde{T}^2_{\max\cdot\mathrm{c}}$ statistics via Monte Carlo simulation, where $\widetilde{T}^2_{\max\cdot\mathrm{p}} = \max_{1\le a<b\le m}(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})'\widetilde{\boldsymbol{\Gamma}}^{[pl]^{-1}}_{ab}(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})$, and $\widetilde{T}^2_{\max\cdot\mathrm{c}} = \max_{2\le b\le m}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})'\widetilde{\boldsymbol{\Gamma}}^{[pl]^{-1}}_{1b}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})$. For each parameter, the simulation involved 1,000,000 trials based on three-step and two-step monotone missing data sets. That is, 1,000,000 of the values of the test statistics are computed for each set $(\alpha, p_i, n_i^{(\ell)})$ of parameters based on the normal random vectors $\boldsymbol{x}_{ij}^{(\ell)}$'s generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$.

Computations are carried out for the following cases:

$$
\begin{aligned}
\text{Case I}: &(p_1,p_2,p_3) = (6,4,2),(12,8,4),\ n_1^{(1)} = n_1^{(2)} = 20,50,100,\\
&(n_2^{(1)}, n_3^{(1)}) = (n_2^{(2)}, n_3^{(2)}) = (10,10),(20,20),\ \alpha = 0.05, 0.01,\\
\text{Case II}: &(p_1,p_2,p_1,p_3) = (6,4,6,2),\\
&n_1^{(1)} = n_1^{(2)} = 20,50,100,\ n_2^{(1)} = n_3^{(2)} = 10,20,\ \alpha = 0.05,0.01,\\
\text{Case III}: &m = 6,10,\ (p_1,p_2,p_3) = (12,8,4),\ \alpha = 0.05,0.01,\\
&n_1^{(\ell)} = 20,50,100,\ (n_2^{(\ell)}, n_3^{(\ell)}) = (10,10),(20,20),\ \ell = 1,2,\dots,m.
\end{aligned}
$$

The simulation results related to the upper percentiles of $\widetilde{T}^2$ statistic and their approximations in the case of three-step monotone missing data are summarized in Table 3.1. Table 3.1 lists the simulated upper 100α percentiles of the $\widetilde{T}^2$ statistic ($\widetilde{t}^2_{\mathrm{simu}}(\alpha)$), the approximate upper 100α percentiles of $\widetilde{T}^2$ ($t^2_{\mathrm{YS\cdot L2}}(\alpha)$, $t^2_{\mathrm{YS}\cdot F2}(\alpha)$), and the upper 100α percentiles of the χ^2 distribution with p degrees of freedom ($\chi^2_p(\alpha)$). In Table 3.1, we denote $\widetilde{t}^2_{\mathrm{simu}}(\alpha)$, $t^2_{\mathrm{YS\cdot L2}}(\alpha)$, and $t^2_{\mathrm{YS}\cdot F2}(\alpha)$ as $\widetilde{t}^2_{\mathrm{simu}}$, $t^2_{\mathrm{YS\cdot L2}}$, and $t^2_{\mathrm{YS}\cdot F2}$ respectively. In addition, we provide the simulated coverage probabilities for the approximate upper 100α percentiles given by

$$
\begin{aligned}
&\widetilde{\mathrm{CP}}(t^2_{\mathrm{YS\cdot L2}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\mathrm{YS\cdot L2}}(\alpha)\},\quad \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS}\cdot F2}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\mathrm{YS}\cdot F2}(\alpha)\},\\
&\quad\widetilde{\mathrm{CP}}(\chi^2_p(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > \chi^2_p(\alpha)\}.
\end{aligned}
$$

It may be noted from Table 3.1 that the simulated values are not close to the upper percentiles of the χ^2 distribution even when the sample size $n_1^{(\ell)}$ is moderately large. However, it is seen that the proposed approximations are accurate even for cases where $n_1^{(\ell)}$ is not large. In particular, it is noted that the values of $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ are highly accurate for all cases. In other words, the simulated coverage probabilities for $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ are considerably close to the nominal level $1-\alpha$. Further, we compute the simulated values and their approximations for the unbalanced cases when $n_1^{(1)} \neq n_1^{(2)}$. It may be noted that the proposed approximations are accurate and that $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ is also considerably close to the simulated value $\widetilde{t}^2_{\mathrm{simu}}(\alpha)$. Thus, it can be concluded that the approximation $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ is highly accurate even for small samples and unbalanced cases when the data have a three-step monotone pattern of missing observations.

Table 3.2 lists the simulated values $t^2_{\mathrm{simu}}(\alpha)$ and $\widetilde{t}^2_{\mathrm{simu}}(\alpha)$, the approximate values $t^2_{\mathrm{YS\cdot L2}}(\alpha)$, $t^2_{\mathrm{YS}\cdot F2}(\alpha)$ and $\chi^2_p(\alpha)$ for the case of two-step monotone missing data. Further,

Table 3.1: Simulated and approximate values and coverage probabilities for Case I

Sample Size			Upper Percentile			Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$\widetilde{t}^2_{\text{simu}}$	$t^2_{\text{YS}\cdot\text{L2}}$	$t^2_{\text{YS}\cdot F2}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L2}}$	$\widetilde{\text{CP}}_{\text{YS}\cdot F2}$	$\widetilde{\text{CP}}_{\chi^2}$
				$(p_1, p_2, p_3) = (6, 4, 2)$				
				$\alpha = 0.05$				
20	10	10	16.22	16.01	15.46	0.947	0.940	0.877
50	10	10	13.97	13.93	13.89	0.949	0.949	0.923
100	10	10	13.29	13.28	13.27	0.950	0.950	0.937
20	20	20	15.79	15.63	14.63	0.948	0.933	0.886
50	20	20	13.86	13.80	13.69	0.949	0.947	0.926
100	20	20	13.22	13.23	13.21	0.950	0.950	0.938
				$\alpha = 0.01$				
20	10	10	22.79	22.44	21.49	0.989	0.986	0.957
50	10	10	19.04	18.97	18.90	0.990	0.990	0.980
100	10	10	17.94	17.91	17.90	0.990	0.990	0.985
20	20	20	22.05	21.83	20.12	0.989	0.984	0.962
50	20	20	18.77	18.76	18.57	0.990	0.989	0.981
100	20	20	17.84	17.83	17.80	0.990	0.990	0.986
				$(p_1, p_2, p_3) = (12, 8, 4)$				
				$\alpha = 0.05$				
20	10	10	34.25	32.56	30.10	0.937	0.912	0.713
50	10	10	25.15	24.96	24.81	0.948	0.946	0.880
100	10	10	23.01	22.96	22.95	0.949	0.949	0.918
20	20	20	33.07	31.36	27.22	0.936	0.886	0.734
50	20	20	24.80	24.57	24.20	0.947	0.943	0.886
100	20	20	22.89	22.83	22.77	0.949	0.948	0.920
				$\alpha = 0.01$				
20	10	10	46.34	43.64	39.65	0.986	0.976	0.853
50	10	10	32.23	31.92	31.70	0.989	0.989	0.960
100	10	10	29.01	29.00	28.97	0.990	0.990	0.979
20	20	20	44.73	41.85	35.28	0.985	0.963	0.869
50	20	20	31.60	31.35	30.79	0.989	0.988	0.964
100	20	20	28.82	28.81	28.72	0.990	0.990	0.979

Note. $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L2}} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L2}}(\alpha))$, $\widetilde{\text{CP}}_{\text{YS}\cdot F2} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot F2}(\alpha))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_p(\alpha))$, $\chi^2_6(0.05) = 12.59$, $\chi^2_6(0.01) = 16.81$, $\chi^2_{12}(0.05) = 21.03$, $\chi^2_{12}(0.01) = 26.22$.

the simulated coverage probabilities for their approximate values are provided in Table 3.2, i.e.,

$$\text{CP}(t^2_{\text{YS}\cdot\text{L2}}(\alpha)) = 1 - \Pr\{T^2 > t^2_{\text{YS}\cdot\text{L2}}(\alpha)\}, \quad \text{CP}(t^2_{\text{YS}\cdot F2}(\alpha)) = 1 - \Pr\{T^2 > t^2_{\text{YS}\cdot F2}(\alpha)\},$$
$$\text{CP}(\chi^2_p(\alpha)) = 1 - \Pr\{T^2 > \chi^2_p(\alpha)\},$$

and $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L2}}(\alpha))$, $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot F2}(\alpha))$, and $\widetilde{\text{CP}}(\chi^2_p(\alpha))$. In Table 3.2, we denote $\text{CP}(t^2_{\text{YS}\cdot\text{L2}}(\alpha))$, $\text{CP}(t^2_{\text{YS}\cdot F2}(\alpha))$, $\text{CP}(\chi^2_p(\alpha))$, $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L2}}(\alpha))$, $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot F2}(\alpha))$, and $\widetilde{\text{CP}}(\chi^2_p(\alpha))$ as $\text{CP}_{\text{YS}\cdot\text{L2}}$, $\text{CP}_{\text{YS}\cdot F2}$, CP_{χ^2}, $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L2}}$, $\widetilde{\text{CP}}_{\text{YS}\cdot F2}$, and $\widetilde{\text{CP}}_{\chi^2}$, respectively. In this simulation study, it may be seen that the upper percentile of the $\widetilde{T}^2$ statistic is larger than that of the T^2 statistic, and that the values of $t^2_{\text{YS}\cdot\text{L2}}(\alpha)$ is larger than that of $t^2_{\text{YS}\cdot F2}(\alpha)$ for all cases. Further, it may be

Table 3.2: Simulated and approximate values and coverage probabilities for Case II $((p_1, p_2, p_1, p_3) = (6, 4, 6, 2))$

Sample Size		Upper Percentile			Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(1)} = n_3^{(2)}$	$\widetilde{t}^2_{\mathrm{simu}}$	$t^2_{\mathrm{YS\cdot L2}}$	$t^2_{\mathrm{YS\cdot}F2}$	$\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L2}}$	$\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot}F2}$	$\widetilde{\mathrm{CP}}_{\chi^2}$
				$\alpha = 0.05$			
20	10	16.64	16.42	16.18	0.947	0.944	0.869
50	10	14.07	14.03	14.02	0.949	0.949	0.922
100	10	13.31	13.31	13.30	0.950	0.950	0.936
20	20	16.25	16.01	15.46	0.947	0.939	0.877
50	20	13.97	13.93	13.89	0.949	0.949	0.924
100	20	13.29	13.28	13.27	0.950	0.950	0.937
				$\alpha = 0.01$			
20	10	23.41	23.13	22.70	0.989	0.988	0.952
50	10	19.22	19.13	19.11	0.990	0.990	0.979
100	10	17.94	17.95	17.95	0.990	0.990	0.985
20	20	22.82	22.44	21.49	0.989	0.986	0.956
50	20	18.97	18.97	18.90	0.990	0.990	0.980
100	20	17.95	17.91	17.90	0.990	0.990	0.985
$n_1^{(\ell)}$	$n_2^{(1)} = n_3^{(2)}$	t^2_{simu}	$t^2_{\mathrm{YS\cdot L2}}$	$t^2_{\mathrm{YS\cdot}F2}$	$\mathrm{CP}_{\mathrm{YS\cdot L2}}$	$\mathrm{CP}_{\mathrm{YS\cdot}F2}$	CP_{χ^2}
				$\alpha = 0.05$			
20	10	16.07	16.42	16.18	0.954	0.951	0.882
50	10	14.00	14.03	14.02	0.951	0.950	0.923
100	10	13.29	13.31	13.30	0.950	0.950	0.937
20	20	15.40	16.01	15.46	0.957	0.951	0.896
50	20	13.84	13.93	13.89	0.951	0.951	0.926
100	20	13.25	13.28	13.27	0.950	0.950	0.938
				$\alpha = 0.01$			
20	10	22.55	23.13	22.70	0.991	0.990	0.958
50	10	19.12	19.13	19.11	0.990	0.990	0.979
100	10	17.92	17.95	17.95	0.990	0.990	0.985
20	20	21.56	22.44	21.49	0.992	0.990	0.965
50	20	18.80	18.97	18.90	0.991	0.990	0.981
100	20	17.91	17.91	17.90	0.990	0.990	0.985

Note. $\chi^2_6(0.05) = 12.59$, $\chi^2_6(0.01) = 16.81$.

noted from Table 3.2 that the approximation $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ is highly accurate approximate upper percentile of $\widetilde{T}^2$ for unbalanced case of the number of dimensions.

In order to compare the approximate values with the simulated values in the two cases of pairwise comparisons and comparisons with a control, computations are carried out for the Case III. Tables 3.3–3.4 list the simulated upper 100α percentiles of the $\widetilde{T}^2_{\mathrm{max\cdot p}}$ statistic $(\widetilde{t}^2_{\mathrm{simu\cdot p}}(\alpha))$, the simulated upper $100\alpha_1$ percentiles of the $\widetilde{T}^2_{ab}$ statistic $(\widetilde{t}^2_{\mathrm{simu\cdot Bon}}(\alpha_1))$, the approximate upper $100\alpha_1$ percentiles of the $\widetilde{T}^2_{ab}$ statistic $(t^2_{\mathrm{YS\cdot L}m}(\alpha_1),\ t^2_{\mathrm{YS\cdot}Fm}(\alpha_1))$ and the upper $100\alpha_1$ percentiles of the χ^2 distribution with p degrees of freedom $(\chi^2_p(\alpha_1))$, where $\alpha_1 = 2\alpha/[m(m-1)]$. We note that the values of $\widetilde{t}^2_{\mathrm{simu\cdot Bon}}(\alpha_1)$ are simulated values obtained via Monte Carlo simulation. In the tables, we denote $\widetilde{t}^2_{\mathrm{simu\cdot p}}(\alpha)$, $\widetilde{t}^2_{\mathrm{simu\cdot Bon}}(\alpha_1)$, $t^2_{\mathrm{YS\cdot L}m}(\alpha_1)$ and $t^2_{\mathrm{YS\cdot}Fm}(\alpha_1)$ as $\widetilde{t}^2_{\mathrm{simu\cdot p}}$, $\widetilde{t}^2_{\mathrm{simu\cdot Bon}}$, $t^2_{\mathrm{YS\cdot L}m}$, and $t^2_{\mathrm{YS\cdot}Fm}$, respectively. In addition, we provide the simulated coverage probabilities given by

Table 3.3: Simulated and approximate values and coverage probabilities for pairwise comparisons for Case Ⅲ ($m = 6$ and $(p_1, p_2, p_3) = (12, 8, 4)$)

Sample Size			Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{p}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m}$	$\widetilde{\text{CP}}_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\chi^2}$
					$\alpha = 0.05$				
20	10	10	34.95	35.71	35.44	34.56	0.956	0.945	0.831
50	10	10	31.40	31.94	31.92	31.84	0.957	0.956	0.916
100	10	10	30.25	30.88	30.74	30.73	0.957	0.956	0.937
20	20	20	34.21	34.86	34.79	33.15	0.957	0.935	0.851
50	20	20	31.16	31.61	31.68	31.48	0.957	0.955	0.921
100	20	20	30.18	30.73	30.66	30.62	0.957	0.956	0.939
					$\alpha = 0.01$				
20	10	10	41.46	41.89	41.60	40.45	0.990	0.987	0.938
50	10	10	36.84	37.08	37.09	36.99	0.991	0.991	0.977
100	10	10	35.39	35.72	35.60	35.59	0.991	0.991	0.984
20	20	20	40.50	40.85	40.77	38.65	0.991	0.984	0.948
50	20	20	36.56	36.77	36.80	36.54	0.991	0.990	0.978
100	20	20	35.27	35.82	35.50	35.46	0.991	0.991	0.985

Note. $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_1))$, $\widetilde{\text{CP}}_{\text{YS}\cdot Fm} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_1))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_p(\alpha_1))$, $\alpha_1 = \alpha/15$, $\chi^2_{12}(0.05/15) = 29.49$, $\chi^2_{12}(0.01/15) = 34.03$.

Table 3.4: Simulated and approximate values and coverage probabilities for pairwise comparisons for Case Ⅲ ($m = 10$ and $(p_1, p_2, p_3) = (12, 8, 4)$)

Sample Size			Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{p}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m}$	$\widetilde{\text{CP}}_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\chi^2}$
					$\alpha = 0.05$				
20	10	10	36.48	37.11	37.14	36.52	0.958	0.950	0.865
50	10	10	33.98	34.58	34.52	34.46	0.958	0.957	0.925
100	10	10	33.05	33.55	33.61	33.60	0.958	0.958	0.943
20	20	20	35.87	36.34	36.64	35.46	0.959	0.944	0.882
50	20	20	33.80	34.38	34.34	34.19	0.958	0.955	0.929
100	20	20	33.02	33.34	33.54	33.52	0.958	0.957	0.943
					$\alpha = 0.01$				
20	10	10	42.20	42.38	42.54	41.77	0.991	0.989	0.957
50	10	10	39.05	39.54	39.33	39.26	0.991	0.991	0.980
100	10	10	37.95	38.12	38.22	38.21	0.991	0.991	0.986
20	20	20	41.46	41.69	41.93	40.48	0.991	0.987	0.964
50	20	20	38.89	39.23	39.11	38.93	0.991	0.990	0.982
100	20	20	37.94	37.80	38.14	38.11	0.991	0.991	0.986

Note. $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_1))$, $\widetilde{\text{CP}}_{\text{YS}\cdot Fm} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_1))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_p(\alpha_1))$, $\alpha_1 = \alpha/45$, $\chi^2_{12}(0.05/45) = 32.62$, $\chi^2_{12}(0.01/45) = 37.01$.

$$\widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_1)) = 1 - \Pr\{\widetilde{T}^2_{\max\cdot\text{p}} > t^2_{\text{YS}\cdot\text{L}m}(\alpha_1)\},$$
$$\widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_1)) = 1 - \Pr\{\widetilde{T}^2_{\max\cdot\text{p}} > t^2_{\text{YS}\cdot Fm}(\alpha_1)\},$$
$$\widetilde{\text{CP}}(\chi^2_p(\alpha_1)) = 1 - \Pr\{\widetilde{T}^2_{\max\cdot\text{p}} > \chi^2_p(\alpha_1)\}.$$

It may be noted from Tables 3.3–3.4 that the simulated values for $\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}(\alpha_1)$ are larger

Table 3.5: Simulated and approximate values and coverage probabilities for comparisons with a control for Case III ($m = 6$ and $(p_1, p_2, p_3) = (12, 8, 4)$)

Sample Size			Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{c}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m}$	$\widetilde{\text{CP}}_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\chi^2}$
					$\alpha = 0.05$				
20	10	10	30.80	31.32	31.14	30.42	0.954	0.945	0.861
50	10	10	27.84	28.28	28.24	28.17	0.955	0.954	0.923
100	10	10	26.84	27.26	27.26	27.25	0.956	0.956	0.940
20	20	20	30.21	30.74	30.60	29.26	0.955	0.937	0.874
50	20	20	27.65	28.07	28.04	27.88	0.955	0.953	0.926
100	20	20	26.82	27.22	27.19	27.16	0.955	0.955	0.941
					$\alpha = 0.01$				
20	10	10	37.36	37.64	37.41	36.45	0.990	0.987	0.952
50	10	10	33.41	33.71	33.58	33.50	0.991	0.990	0.979
100	10	10	32.11	32.26	32.31	32.30	0.991	0.991	0.986
20	20	20	36.55	36.78	36.70	34.92	0.990	0.985	0.958
50	20	20	33.16	33.35	33.33	33.11	0.991	0.990	0.981
100	20	20	32.03	32.14	32.22	32.18	0.991	0.990	0.986

Note. $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_2))$, $\widetilde{\text{CP}}_{\text{YS}\cdot Fm} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_2))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_p(\alpha_2))$, $\alpha_2 = \alpha/5$, $\chi^2_{12}(0.05/5) = 26.22$, $\chi^2_{12}(0.01/5) = 30.96$.

Table 3.6: Simulated and approximate values and coverage probabilities for comparisons with a control for Case III ($m = 10$ and $(p_1, p_2, p_3) = (12, 8, 4)$)

Sample Size			Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{c}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m}$	$\widetilde{\text{CP}}_{\text{YS}\cdot Fm}$	$\widetilde{\text{CP}}_{\chi^2}$
					$\alpha = 0.05$				
20	10	10	30.97	31.61	31.56	31.08	0.957	0.951	0.895
50	10	10	28.94	29.55	29.50	29.46	0.958	0.957	0.934
100	10	10	28.25	28.73	28.78	28.77	0.957	0.957	0.946
20	20	20	30.51	31.17	31.17	30.25	0.958	0.946	0.904
50	20	20	28.82	29.38	29.36	29.24	0.957	0.956	0.936
100	20	20	28.20	28.76	28.73	28.70	0.957	0.957	0.947
					$\alpha = 0.01$				
20	10	10	36.86	37.06	37.14	36.52	0.991	0.989	0.968
50	10	10	34.20	34.53	34.52	34.46	0.991	0.991	0.983
100	10	10	33.35	33.52	33.61	33.60	0.991	0.991	0.987
20	20	20	36.29	36.60	36.64	35.46	0.991	0.987	0.972
50	20	20	34.09	34.26	34.34	34.19	0.991	0.990	0.984
100	20	20	33.29	33.63	33.54	33.52	0.991	0.991	0.988

Note. $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}m} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_2))$, $\widetilde{\text{CP}}_{\text{YS}\cdot Fm} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_2))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_p(\alpha_2))$, $\alpha_2 = \alpha/9$, $\chi^2_{12}(0.05/9) = 27.99$, $\chi^2_{12}(0.01/9) = 32.62$.

than the simulated values for $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$ because Bonferroni's approximation is always an overestimate for the $T^2_{\max}$-type statistic: this can be shown theoretically. It may be seen from the tables that the approximate values of $t^2_{\text{YS}\cdot\text{L}m}(\alpha_1)$ and $t^2_{\text{YS}\cdot Fm}(\alpha_1)$ are closer to the simulated values of $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$ when the sample size becomes large. The simulation studies show that $t^2_{\text{YS}\cdot\text{L}}(\alpha_1)$ is close to $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$ and is a conservative approximation.

Further, Tables 3.5–3.6 list the results for the case of comparisons with a control. We provide $\widetilde{t}^2_{\text{simu}\cdot\text{c}}(\alpha)$, $\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}(\alpha_2)$, $t^2_{\text{YS}\cdot Lm}(\alpha_2)$, $t^2_{\text{YS}\cdot Fm}(\alpha_2)$ and $\chi^2_p(\alpha_2)$, as well as $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot Lm}(\alpha_2))$, $\widetilde{\text{CP}}(t^2_{\text{YS}\cdot Fm}(\alpha_2))$ and $\widetilde{\text{CP}}(\chi^2_p(\alpha_2))$, where $\alpha_2 = \alpha/(m-1)$. The accuracy of the approximations is similar to that in the case of pairwise comparisons.

3.2 k-step monotone missing data

In this section, we consider testing the equality of mean vectors and simultaneous confidence intervals when each data set has a monotone missing data pattern. We give a simplified T^2-type statistic and propose the approximate upper percentiles of the statistic in the case of data with general k-step monotone missing data patterns. We also consider multivariate multiple comparisons for mean vectors with general k-step monotone missing data. Approximate simultaneous confidence intervals for pairwise comparisons among mean vectors and comparisons with a control are obtained using Bonferroni's approximation procedure. Finally, the accuracy and asymptotic behavior of the approximations are investigated by Monte Carlo simulation.

3.2.1 MLEs of the mean vectors and the covariance matrix

Using the notations in Section 2.2, we consider the MLEs of the mean vectors and the common covariance matrix for the m-sample problem.

Let $\boldsymbol{x}^{(\ell)}_{i1}, \boldsymbol{x}^{(\ell)}_{i2}, \ldots, \boldsymbol{x}^{(\ell)}_{in^{(\ell)}_i}$ be distributed as $N_{p_i}(\boldsymbol{\mu}^{(\ell)}_i, \boldsymbol{\Sigma}_i)$ for $i = 1, 2, \ldots, k$ and $\ell = 1, 2, \ldots, m$, where $\boldsymbol{\mu}^{(\ell)}_i = (\mu^{(\ell)}_1, \mu^{(\ell)}_2, \ldots, \mu^{(\ell)}_{p_i})'$ and $\sum_{\ell=1}^m n^{(\ell)}_1 - m \geq p_1$. Let

$$\overline{\boldsymbol{x}}^{(\ell)}_i = \frac{1}{n^{(\ell)}_i} \sum_{j=1}^{n^{(\ell)}_i} \boldsymbol{x}^{(\ell)}_{ij}, \ \boldsymbol{E}^{(\ell)}_i = \sum_{j=1}^{n^{(\ell)}_i} (\boldsymbol{x}^{(\ell)}_{ij} - \overline{\boldsymbol{x}}^{(\ell)}_i)(\boldsymbol{x}^{(\ell)}_{ij} - \overline{\boldsymbol{x}}^{(\ell)}_i)', \ i = 1, 2, \ldots, k.$$

Further, we define

$$\begin{aligned}
&N^{(\ell)}_1 = 0, \ N^{(\ell)}_{i+1} = \sum_{j=1}^{i} n^{(\ell)}_j, \ i = 1, 2, \ldots, k, \\
&\nu_{i\cdot m} = \sum_{\ell=1}^{m} n^{(\ell)}_i, \ i = 1, 2, \ldots, k, \quad M_{i\cdot m} = \sum_{\ell=1}^{m} N^{(\ell)}_i, i = 1, 2, \ldots, k+1, \\
&\boldsymbol{d}^{(\ell)}_1 = \overline{\boldsymbol{x}}^{(\ell)}_1, \ \boldsymbol{d}^{(\ell)}_i = \frac{n^{(\ell)}_i}{N^{(\ell)}_{i+1}} \Big[\overline{\boldsymbol{x}}^{(\ell)}_i - \frac{1}{N^{(\ell)}_i} \sum_{j=1}^{i-1} n^{(\ell)}_j (\overline{\boldsymbol{x}}^{(\ell)}_j)_i \Big], \ i = 2, 3, \ldots, k, \\
&\boldsymbol{f}^{(\ell)}_1 = \boldsymbol{d}^{(\ell)}_1, \quad \boldsymbol{f}^{(\ell)}_i = \boldsymbol{U}_i \boldsymbol{d}^{(\ell)}_i, \quad i = 2, 3, \ldots, k, \\
&\boldsymbol{U}_1 = \boldsymbol{T}_1, \quad \boldsymbol{U}_i = \boldsymbol{U}_{i-1} \boldsymbol{T}_i, \quad i = 2, 3, \ldots, k, \\
&\boldsymbol{T}_1 = \boldsymbol{I}_{p_1}, \quad \boldsymbol{T}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{\Sigma}'_{(i,2)} \boldsymbol{\Sigma}^{-1}_{i+1} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1.
\end{aligned}$$

Using a straightforward manner based on the case of the one-sample problem by Jinadasa and Tracy (1992), we can derive the MLEs of $\boldsymbol{\mu}^{(\ell)}$ $\ell = 1, 2, \ldots, m$ and $\boldsymbol{\Sigma}$. We partially differentiate the loglikelihood function with respect to $\boldsymbol{\mu}^{(\ell)}$ and $\boldsymbol{\Sigma}$, respectively, then solving the likelihood equation, we obtain the following theorem.

Theorem 3.3. (Yagi and Seo, 2017) *Let $\boldsymbol{x}_{ij}^{(\ell)}$ $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i^{(\ell)}$, $\ell = 1, 2, \ldots, m$ be the j-th random vector of the i-th step from the ℓ-th population distributed as $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$. Then, the MLEs of $\boldsymbol{\mu}^{(\ell)}$, $\ell = 1, 2, \ldots, m$ are given by*

$$\widehat{\boldsymbol{\mu}}^{(\ell)} = \sum_{i=1}^{k} \widehat{\boldsymbol{f}}_i^{(\ell)},$$

where

$$\widehat{\boldsymbol{f}}_1^{(\ell)} = \boldsymbol{d}_1^{(\ell)}, \ \widehat{\boldsymbol{f}}_i^{(\ell)} = \widehat{\boldsymbol{U}}_i^{[\mathrm{pl}]} \boldsymbol{d}_i^{(\ell)}, \ i = 2, 3, \ldots, k,$$
$$\widehat{\boldsymbol{U}}_1^{[pl]} = \boldsymbol{T}_1, \quad \widehat{\boldsymbol{U}}_i^{[pl]} = \widehat{\boldsymbol{U}}_{i-1}^{[pl]} \widehat{\boldsymbol{T}}_i^{[pl]}, \quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{T}_1 = \boldsymbol{I}_{p_1}, \quad \widehat{\boldsymbol{T}}_{i+1}^{[pl]} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \widehat{\boldsymbol{\Sigma}}_{(i,2)}^{[pl]\prime} \widehat{\boldsymbol{\Sigma}}_{i+1}^{[pl]-1} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1;$$

then, the MLE of the covariance matrix is given by

$$\widehat{\boldsymbol{\Sigma}}^{[\mathrm{pl}]} = \frac{1}{M_{2\cdot m}} \sum_{\ell=1}^{m} \boldsymbol{H}_1^{(\ell)} + \sum_{\ell=1}^{m} \sum_{i=2}^{k} \frac{1}{M_{i+1\cdot m}} \boldsymbol{F}_i^{[\mathrm{pl}]} \left[\boldsymbol{H}_i^{(\ell)} - \frac{\nu_{i\cdot m}}{M_{i\cdot m}} \boldsymbol{L}_{i-1,1}^{(\ell)} \right] \boldsymbol{F}_i^{[\mathrm{pl}]\prime},$$

where

$$\boldsymbol{H}_1^{(\ell)} = \boldsymbol{E}_1^{(\ell)}, \ \boldsymbol{H}_i^{(\ell)} = \boldsymbol{E}_i^{(\ell)} + \frac{N_i^{(\ell)} N_{i+1}^{(\ell)}}{n_i^{(\ell)}} \boldsymbol{d}_i^{(\ell)} \boldsymbol{d}_i^{(\ell)\prime}, \ i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_1^{(\ell)} = \boldsymbol{H}_1^{(\ell)}, \ \boldsymbol{L}_i^{(\ell)} = (\boldsymbol{L}_{i-1}^{(\ell)})_i + \boldsymbol{H}_i^{(\ell)}, \ i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_{i1}^{(\ell)} = (\boldsymbol{L}_i^{(\ell)})_{i+1}, \ \boldsymbol{L}_i^{(\ell)} = \begin{pmatrix} \boldsymbol{L}_{i1}^{(\ell)} & \boldsymbol{L}_{i2}^{(\ell)} \\ \boldsymbol{L}_{i2}^{(\ell)\prime} & \boldsymbol{L}_{i3}^{(\ell)} \end{pmatrix}, \ i = 1, 2, \ldots, k-1,$$

and

$$\boldsymbol{F}_1^{[\mathrm{pl}]} = \boldsymbol{G}_1, \quad \boldsymbol{F}_i^{[\mathrm{pl}]} = \boldsymbol{F}_{i-1}^{[\mathrm{pl}]} \boldsymbol{G}_i^{[\mathrm{pl}]}, \quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{G}_1 = \boldsymbol{I}_{p_1}, \quad \boldsymbol{G}_{i+1}^{[\mathrm{pl}]} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \left(\sum_{\ell=1}^{m} \boldsymbol{L}_{i2}^{(\ell)} \right)' \left(\sum_{\ell=1}^{m} \boldsymbol{L}_{i1}^{(\ell)} \right)^{-1} \end{pmatrix}, \quad i = 1, 2, \ldots, k-1.$$

We note that, in the case of one-sample problem ($m = 1$), $\widehat{\boldsymbol{\mu}}^{(\ell)}$ and $\widehat{\boldsymbol{\Sigma}}^{[\mathrm{pl}]}$ in Theorem 3.3 are reduced to $\widehat{\boldsymbol{\mu}}(= \sum_{i=1}^{k} \widehat{\boldsymbol{f}}_i)$, and $\widehat{\boldsymbol{\Sigma}}$ in Jinadasa and Tracy (1992), respectively.

3.2.2 Two-sample problem

In this section, we test the equality of two mean vectors with k-step monotone missing data. We give the simplified T^2-type statistic and its approximate upper percentiles. To test the hypothesis $H_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$ vs. $H_1 : \boldsymbol{\mu}^{(1)} \neq \boldsymbol{\mu}^{(2)}$ when two data sets have the same k-step monotone missing data pattern, we adopt the simplified T^2-type statistic given by

$$\widetilde{T}^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \, \widetilde{\boldsymbol{\Gamma}}^{[\mathrm{pl}]^{-1}} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}),$$

where $\widehat{\boldsymbol{\mu}}^{(\ell)} = \sum_{i=1}^{k} \widehat{\boldsymbol{f}}_i^{(\ell)}$, $\ell = 1, 2$, and $\widetilde{\boldsymbol{\Gamma}}^{[\mathrm{pl}]}$ is an estimator of $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}]$, $\widetilde{\boldsymbol{\mu}}^{(\ell)} = \sum_{i=1}^{k} \boldsymbol{f}_i^{(\ell)}$, $\ell = 1, 2$. Then, using Theorem 2.3 for the one-sample problem, we have

$$\widetilde{\boldsymbol{\Gamma}}^{[\mathrm{pl}]} = \frac{n_1^{(1)} + n_1^{(2)}}{n_1^{(1)} n_1^{(2)}} \widehat{\boldsymbol{\Sigma}}_1^{[\mathrm{pl}]} - \sum_{\ell=1}^{2} \sum_{i=2}^{k} \frac{n_i^{(\ell)}}{N_i^{(\ell)} N_{i+1}^{(\ell)}} \widehat{\boldsymbol{U}}_i^{[\mathrm{pl}]} \widehat{\boldsymbol{\Sigma}}_i^{[\mathrm{pl}]} \widehat{\boldsymbol{U}}_i^{[\mathrm{pl}]'},$$

where $\widehat{\boldsymbol{\Sigma}}^{[\mathrm{pl}]}$ is the MLE when $m = 2$ in Theorem 3.3. We note that under H_0, $\widetilde{T}^2$ is asymptotically distributed as a χ^2 distribution with p degrees of freedom when $n_1^{(\ell)}, N_{k+1}^{(\ell)} \to \infty$ with $n_1^{(\ell)}/N_{k+1}^{(\ell)} \to \delta^{(\ell)} \in (0, 1]$, $\ell = 1, 2$. However, as with the one-sample problem, we note that the χ^2 approximation is not a good approximate upper percentile of the $\widetilde{T}^2$ statistic when the sample size is not large. We propose the two approximate upper 100α percentiles of the $\widetilde{T}^2$ statistic given by

$$t^2_{\mathrm{YS\cdot L2}}(\alpha) = (1 - \omega_2) T^2_{\nu_{1\cdot2},\alpha} + \omega_2 T^2_{M_{k+1\cdot2},\alpha},$$

$$t^2_{\mathrm{YS\cdot}F2}(\alpha) = \frac{n_2^* p_1}{n_2^* - p_1 - 1} F_{p_1, n_2^* - p_1 - 1, \alpha},$$

where

$$\omega_2 = \frac{\sum_{i=2}^{k} \nu_{i\cdot2} p_i}{p_1 \sum_{i=2}^{k} \nu_{i\cdot2}}, \; n_2^* = \frac{1}{p_1} \sum_{i=1}^{k} \nu_{i\cdot2} p_i,$$

$$T^2_{\nu_{1\cdot2},\alpha} = \frac{\nu_{1\cdot2} p_1}{\nu_{1\cdot2} - p_1 - 1} F_{p_1, \nu_{1\cdot2} - p_1 - 1, \alpha}, \; T^2_{M_{k+1\cdot2},\alpha} = \frac{M_{k+1\cdot2} p_1}{M_{k+1\cdot2} - p_1 - 1} F_{p_1, M_{k+1\cdot2} - p_1 - 1, \alpha}.$$

Further, we can test the equality of two mean vectors when two data sets have unequal general step monotone missing data patterns. For two-step case, see Section 3.1.3.

Next, under the two-sample problem, we consider the simultaneous confidence intervals when each data set has k-step monotone missing observations.

For any nonnull vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)'$, the simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$ with the confidence level $(1 - \alpha)$ are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) - \sqrt{L} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) + \sqrt{L}, \; \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L = t^2(\alpha) \boldsymbol{c}' \widehat{\boldsymbol{\Gamma}}^{[pl]} \boldsymbol{c}$ and $t^2(\alpha)$ is the upper 100α percentile of the $T^2 (= (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \widetilde{\boldsymbol{\Gamma}}^{[\mathrm{pl}]^{-1}} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}))$ statistic and $\widehat{\boldsymbol{\Gamma}}^{[\mathrm{pl}]}$ is an estimator of $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}]$. However, it is not easy to obtain $t^2(\alpha)$. Therefore, using the approximate upper percentiles of the $\widetilde{T}^2$ statistic, $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ or $t^2_{\mathrm{YS\cdot}F2}(\alpha)$, for any nonnull vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_p)'$, the approximate simultaneous confidence intervals for $\boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$ can be obtained by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) - \sqrt{L_{\mathrm{app}}} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}) + \sqrt{L_{\mathrm{app}}}, \; \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

where $L_{\mathrm{app}} = t^2_{\mathrm{app}}(\alpha) \boldsymbol{c}' \widetilde{\boldsymbol{\Gamma}}^{[pl]} \boldsymbol{c}$ and the value of $t^2_{\mathrm{app}}(\alpha)$ is $t^2_{\mathrm{YS\cdot L2}}(\alpha)$ or $t^2_{\mathrm{YS\cdot}F2}(\alpha)$.

3.2.3 Simultaneous confidence intervals for multiple comparisons among mean vectors

Under the m-sample problem, we consider the simultaneous confidence intervals for pairwise multiple comparisons among mean vectors when each data set has k-step monotone missing observations. As with the three-step case in Section 3.1.4, we adopt Bonferroni's approximation. Let $n_i^{(1)} = n_i^{(2)} = \cdots = n_i^{(m)}$, $i = 1, 2, \ldots, k$; then, the null distributions of T_{ab}^2 and T_{1b}^2 are identical. Their approximate simultaneous confidence intervals for pairwise comparisons and comparisons with a control are given by

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L_{\mathrm{p}}^*} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(a)} - \boldsymbol{\mu}^{(b)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L_{\mathrm{p}}^*},$$
$$1 \le a < b \le m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

and

$$\boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) - \sqrt{L_{\mathrm{c}}^*} \le \boldsymbol{c}'(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(b)}) \le \boldsymbol{c}'(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)}) + \sqrt{L_{\mathrm{c}}^*},$$
$$2 \le b \le m,\ \forall \boldsymbol{c} \in \boldsymbol{R}^p - \{\boldsymbol{0}\},$$

respectively, where

$$L_{\mathrm{p}}^* = t_{\mathrm{Bon}}^2(\alpha_{\mathrm{p}})\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}_{ab}^{[pl]}\boldsymbol{c},\ L_{\mathrm{c}}^* = t_{\mathrm{Bon}}^2(\alpha_{\mathrm{c}})\boldsymbol{c}'\widetilde{\boldsymbol{\Gamma}}_{1b}^{[pl]}\boldsymbol{c}.$$

Further, the value of $t_{\mathrm{Bon}}^2(\alpha_{\mathrm{p}})$ is $t_{\mathrm{YS}\cdot\mathrm{L}m}^2(\alpha_{\mathrm{p}})$ or $t_{\mathrm{YS}\cdot Fm}^2(\alpha_{\mathrm{p}})$, and the value of $t_{\mathrm{Bon}}^2(\alpha_{\mathrm{c}})$ is $t_{\mathrm{YS}\cdot\mathrm{L}m}^2(\alpha_{\mathrm{c}})$ or $t_{\mathrm{YS}\cdot Fm}^2(\alpha_{\mathrm{c}})$, which are given in

$$t_{\mathrm{YS}\cdot\mathrm{L}m}^2(\alpha_{\mathrm{p}}) = (1-\omega_m)T_{\nu_{1\cdot m},\alpha_{\mathrm{p}}}^2 + \omega_m T_{M_{k+1\cdot m},\alpha_{\mathrm{p}}}^2,$$
$$t_{\mathrm{YS}\cdot Fm}^2(\alpha_{\mathrm{p}}) = \frac{n_m^* p_1}{n_m^* - p_1 - (m-1)} F_{p_1, n_m^* - p_1 - (m-1), \alpha_{\mathrm{p}}},$$

and

$$t_{\mathrm{YS}\cdot\mathrm{L}m}^2(\alpha_{\mathrm{c}}) = (1-\omega_m)T_{\nu_{1\cdot m},\alpha_{\mathrm{c}}}^2 + \omega_m T_{M_{k+1\cdot m},\alpha_{\mathrm{c}}}^2,$$
$$t_{\mathrm{YS}\cdot Fm}^2(\alpha_{\mathrm{c}}) = \frac{n_m^* p_1}{n_m^* - p_1 - (m-1)} F_{p_1, n_m^* - p_1 - (m-1), \alpha_{\mathrm{c}}},$$

respectively, where

$$\alpha_{\mathrm{p}} = \frac{2\alpha}{m(m-1)},\ \alpha_{\mathrm{c}} = \frac{\alpha}{m-1},\ \omega_m = \frac{\sum_{i=2}^k \nu_{i\cdot m} p_i}{p_1 \sum_{i=2}^k \nu_{i\cdot m}},\ n_m^* = \frac{1}{p_1}\sum_{i=1}^k \nu_{i\cdot m} p_i,$$
$$T_{\nu_{1\cdot m},\alpha}^2 = \frac{\nu_{1\cdot m} p_1}{\nu_{1\cdot m} - p_1 - (m-1)} F_{p_1, \nu_{1\cdot m} - p_1 - (m-1), \alpha},$$
$$T_{M_{k+1\cdot m},\alpha}^2 = \frac{M_{k+1\cdot m} p_1}{M_{k+1\cdot m} - p_1 - (m-1)} F_{p_1, M_{k+1\cdot m} - p_1 - (m-1), \alpha}.$$

We note that $\widetilde{\boldsymbol{\Gamma}}_{ab}^{[\mathrm{pl}]}$ and $\widetilde{\boldsymbol{\Gamma}}_{1b}^{[\mathrm{pl}]}$ are estimated by the use of $\widehat{\boldsymbol{\Sigma}}^{[\mathrm{pl}]}$ in Theorem 3.3.

3.2.4 Simulation studies

To investigate the accuracy of some of the approximations under the two-sample problem, we compute the upper percentiles of the $\widetilde{T}^2$ statistic by Monte Carlo simulation. As with the one-sample problem in Section 2.2.3, the $\widetilde{T}^2$ statistic are computed 10^6 times for each set $(\alpha, p_i, n_i^{(\ell)})$ of parameters based on the normal random vectors $\boldsymbol{x}_{ij}^{(\ell)}$ generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$, $i = 1, 2, \dots, k$, $\ell = 1, 2$. The simulation results related to the upper percentiles of the $\widetilde{T}^2$ statistic and their approximations in the cases of five-step and ten-step monotone missing data are summarized in Tables 3.7 and 3.8. Computations are carried out for the following two cases:

$$\begin{aligned}
\text{Five-step Case:}\quad & (p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3),\\
& n_1^{(1)} = n_1^{(2)} = 10, 11, 30(10)50, 100, 200, 400, 800,\\
& n_2^{(\ell)} = n_3^{(\ell)} = \cdots = n_5^{(\ell)} = 5, 10,\ \ \ell = 1, 2,\ \ \alpha = 0.05, 0.01.\\
\text{Ten-step Case:}\quad & (p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}) = (20, 18, 16, 14, 12, 10, 8, 6, 4, 2),\\
& n_1^{(1)} = n_1^{(2)} = 12, 13, 30(10)50, 100, 200, 400, 800,\\
& n_2^{(\ell)} = n_3^{(\ell)} = \cdots = n_{10}^{(\ell)} = 5, 10,\ \ \ell = 1, 2,\ \ \alpha = 0.05, 0.01.
\end{aligned}$$

Tables 3.7 and 3.8 give the simulated upper 100α percentiles of the $\widetilde{T}^2$ statistic $(\widetilde{t}^2_{\text{simu}}(\alpha))$, the approximate upper 100α percentiles of $\widetilde{T}^2$ $(t^2_{\text{YS}\cdot\text{L2}}(\alpha)$, $t^2_{\text{YS}\cdot F2}(\alpha)$, and $t^2_{\text{YKP}}(\alpha))$, where $t^2_{\text{YKP}}(\alpha)$ is the approximation in Yu et al. (2006). In the tables, we denote $\widetilde{t}^2_{\text{simu}}(\alpha)$, $t^2_{\text{YS}\cdot\text{L2}}(\alpha)$, $t^2_{\text{YS}\cdot F2}(\alpha)$, and $t^2_{\text{YKP}}(\alpha)$ as $\widetilde{t}^2_{\text{simu}}$, $t^2_{\text{YS}\cdot\text{L2}}$, $t^2_{\text{YS}\cdot F2}$, and t^2_{YKP}, respectively. In addition, we provide the simulated coverage probabilities for the approximate upper 100α percentiles given by

$$\text{CP}(t^2_{\text{YS}\cdot\text{L2}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YS}\cdot\text{L2}}(\alpha)\},\quad \text{CP}(t^2_{\text{YS}\cdot F2}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YS}\cdot F2}(\alpha)\},$$

$$\text{CP}(t^2_{\text{YKP}}(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > t^2_{\text{YKP}}(\alpha)\},\quad \text{CP}(\chi^2_p(\alpha)) = 1 - \Pr\{\widetilde{T}^2 > \chi^2_p(\alpha)\}.$$

It may be noted from Tables 3.7 and 3.8 that the simulated values are not close to $\chi^2_p(\alpha)$ even when the sample size $n_1^{(\ell)}$ is moderately large. However, the proposed approximations are accurate even for cases in which $n_1^{(\ell)}$ is not large. In particular, the values of $t^2_{\text{YS}\cdot\text{L2}}(\alpha)$ are highly accurate when $n_1^{(\ell)} \geq 30$, $\ell = 1, 2$. In other words, the actual coverage probabilities for $t^2_{\text{YS}\cdot\text{L2}}(\alpha)$ are considerably close to the nominal level $1-\alpha$. Thus, it can be concluded that the approximation $t^2_{\text{YS}\cdot\text{L2}}(\alpha)$ is highly accurate for moderately large sample sizes when the data have a k-step monotone pattern of missing observations. As with the one-sample problem, the condition for $t^2_{\text{YS}\cdot\text{L2}}$ and $t^2_{\text{YS}\cdot F2}$ is $\sum_{\ell=1}^{2} n_1^{(\ell)} - 2 \geq p_1$ and the condition for t^2_{YKP} is $\sum_{\ell=1}^{2} n_1^{(\ell)} - 6 \geq p_1$. For example, when $p_1 = 15$ and $n_1^{(1)} = n_1^{(2)} = 10$ in Table 3.7, the value of t^2_{YKP} cannot be computed but that of $t^2_{\text{YS}\cdot\text{L2}}$ or $t^2_{\text{YS}\cdot F2}$ can be computed.

Next, in order to compare the approximate values with the simulated values in the cases of pairwise comparisons and comparisons with a control, we compute for the following case:

$$\begin{aligned}
\text{Five-step Case:}\quad & m = 6, 10,\ \ (p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3),\\
& n_1^{(\ell)} = 25(5)50, 100, 200, 400, 800,\ \ n_2^{(\ell)} = n_3^{(\ell)} = \cdots = n_5^{(\ell)} = 5, 10,\\
& \ell = 1, 2, \dots, m,\ \ \alpha = 0.05, 0.01.
\end{aligned}$$

Table 3.7: Simulated and approximate values and coverage probabilities when $m = 2$ and $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability			
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L2}$	$t^2_{{\rm YS}\cdot F2}$	$t^2_{\rm YKP}$	$\rm CP_{YS\cdot L2}$	$\mathrm{CP}_{\mathrm{YS}\cdot F2}$	$\rm CP_{YKP}$	CP_{χ^2}
				$\alpha = 0.05$					
10	5	374.58	239.10	52.69	–	0.901	0.412	–	0.106
11	5	173.46	127.40	50.20	174.30	0.906	0.563	0.951	0.181
30	5	36.17	35.39	34.23	34.97	0.944	0.933	0.940	0.753
40	5	32.77	32.38	31.91	31.90	0.946	0.942	0.942	0.816
50	5	31.02	30.76	30.52	30.31	0.947	0.945	0.943	0.849
100	5	27.84	27.78	27.75	27.49	0.949	0.949	0.946	0.906
200	5	26.38	26.38	26.37	26.22	0.950	0.950	0.948	0.930
400	5	25.68	25.68	25.68	25.60	0.950	0.950	0.949	0.941
800	5	25.33	25.34	25.34	25.30	0.950	0.950	0.949	0.945
10	10	369.14	235.62	38.87	–	0.902	0.306	–	0.131
11	10	168.55	124.17	38.21	166.62	0.908	0.451	0.949	0.216
30	10	34.75	34.23	31.91	33.89	0.946	0.921	0.942	0.778
40	10	31.91	31.58	30.52	31.23	0.947	0.935	0.943	0.832
50	10	30.46	30.18	29.60	29.85	0.947	0.941	0.944	0.860
100	10	27.67	27.59	27.50	27.36	0.949	0.948	0.946	0.909
200	10	26.36	26.32	26.31	26.18	0.949	0.949	0.948	0.930
400	10	25.64	25.67	25.67	25.59	0.950	0.950	0.949	0.941
800	10	25.33	25.34	25.34	25.29	0.950	0.950	0.950	0.945
				$\alpha = 0.01$					
10	5	971.77	557.74	72.22	–	0.974	0.558	–	0.171
11	5	351.45	232.68	68.21	344.88	0.974	0.717	0.990	0.278
30	5	46.61	45.48	43.72	43.79	0.988	0.984	0.985	0.885
40	5	41.52	41.03	40.35	39.59	0.989	0.988	0.986	0.927
50	5	39.08	38.69	38.35	37.47	0.989	0.988	0.986	0.946
100	5	34.47	34.46	34.42	33.78	0.990	0.990	0.988	0.974
200	5	32.48	32.49	32.49	32.14	0.990	0.990	0.989	0.983
400	5	31.52	31.53	31.53	31.35	0.990	0.990	0.990	0.987
800	5	31.09	31.05	31.05	30.96	0.990	0.990	0.990	0.988
10	10	968.92	552.61	50.62	–	0.974	0.428	–	0.203
11	10	347.78	227.92	49.62	334.80	0.974	0.597	0.989	0.322
30	10	44.49	43.82	40.35	42.41	0.989	0.980	0.986	0.903
40	10	40.23	39.91	38.35	38.73	0.989	0.986	0.987	0.936
50	10	38.22	37.87	37.03	36.88	0.989	0.987	0.987	0.951
100	10	34.26	34.19	34.06	33.61	0.990	0.990	0.988	0.975
200	10	32.42	32.41	32.39	32.09	0.990	0.990	0.989	0.983
400	10	31.51	31.51	31.51	31.34	0.990	0.990	0.989	0.987
800	10	31.00	31.05	31.05	30.96	0.990	0.990	0.990	0.989

Note. $\rm CP_{YS\cdot L2} = CP(t^2_{YS\cdot L2}(\alpha))$, $\mathrm{CP}_{\mathrm{YS}\cdot F2} = \mathrm{CP}(t^2_{\mathrm{YS}\cdot F2}(\alpha))$, $\rm CP_{YKP} = CP(t^2_{YKP}(\alpha))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_p(\alpha))$, $\chi^2_{15}(0.05) = 25.00$, $\chi^2_{15}(0.01) = 30.58$.

Then, we define

$$\widetilde{T}^2_{\rm max\cdot p} = \max_{1\le a<b\le m} \widetilde{T}^2_{ab}, \quad \widetilde{T}^2_{\rm max\cdot c} = \max_{2\le b\le m} \widetilde{T}^2_{1b},$$

where $\widetilde{T}^2_{ab} = (\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})' \widetilde{\boldsymbol{\Gamma}}^{[\rm pl]^{-1}}_{ab} (\widehat{\boldsymbol{\mu}}^{(a)} - \widehat{\boldsymbol{\mu}}^{(b)})$, $\widetilde{T}^2_{1b} = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})' \widetilde{\boldsymbol{\Gamma}}^{[\rm pl]^{-1}}_{1b} (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(b)})$, and $\widetilde{\boldsymbol{\Gamma}}^{[\rm pl]}_{ab}$ and $\widetilde{\boldsymbol{\Gamma}}^{[\rm pl]}_{1b}$ are estimators of $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}^{(a)} - \widetilde{\boldsymbol{\mu}}^{(b)}]$ and $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(b)}]$, respectively. Tables 3.9 and 3.10 give the simulated upper 100α percentiles of the $\widetilde{T}^2_{\rm max\cdot p}$ statistic $(\widetilde{t}^2_{\rm simu\cdot p}(\alpha))$,

Table 3.8: Simulated and approximate values and coverage probabilities when $m = 2$ and $(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}) = (20, 18, 16, 14, 12, 10, 8, 6, 4, 2)$

Sample Size		Upper Percentile				Coverage Probability			
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L2}$	$t^2_{{\rm YS}\cdot F2}$	$t^2_{\rm YKP}$	$\mathrm{CP_{YS\cdot L2}}$	$\mathrm{CP}_{{\rm YS}\cdot F2}$	$\mathrm{CP_{YKP}}$	CP_{χ^2}
				$\alpha = 0.05$					
12	5	1105.49	713.46	51.55	–	0.916	0.203	–	0.060
13	5	326.95	257.56	50.67	316.67	0.925	0.348	0.947	0.117
30	5	47.81	47.61	42.42	46.15	0.949	0.903	0.938	0.670
40	5	42.74	42.47	40.20	41.49	0.948	0.927	0.940	0.759
50	5	40.18	39.94	38.72	39.17	0.948	0.937	0.941	0.807
100	5	35.66	35.56	35.39	35.14	0.949	0.947	0.945	0.888
200	5	33.59	33.51	33.49	33.27	0.949	0.949	0.946	0.921
400	5	32.50	32.48	32.47	32.34	0.950	0.950	0.948	0.936
800	5	31.94	31.95	31.95	31.88	0.950	0.950	0.949	0.943
12	10	1103.64	710.97	41.30	–	0.917	0.172	–	0.091
13	10	315.75	255.15	41.08	298.22	0.929	0.302	0.945	0.168
30	10	44.67	46.14	38.43	43.59	0.959	0.884	0.942	0.726
40	10	40.72	41.31	37.46	39.85	0.955	0.916	0.942	0.795
50	10	38.77	39.00	36.72	38.02	0.952	0.929	0.943	0.832
100	10	35.22	35.14	34.71	34.76	0.949	0.944	0.945	0.895
200	10	33.38	33.36	33.29	33.15	0.950	0.949	0.947	0.923
400	10	32.42	32.43	32.42	32.31	0.950	0.950	0.949	0.937
800	10	31.93	31.94	31.93	31.87	0.950	0.950	0.949	0.944
				$\alpha = 0.01$					
12	5	3839.09	2160.69	65.62	–	0.979	0.296	–	0.101
13	5	774.54	522.10	64.33	701.70	0.979	0.481	0.988	0.191
30	5	60.29	60.18	52.59	56.50	0.990	0.973	0.984	0.828
40	5	52.96	52.74	49.50	50.25	0.990	0.982	0.984	0.893
50	5	49.37	49.17	47.46	47.25	0.990	0.986	0.985	0.923
100	5	43.26	43.14	42.90	42.16	0.990	0.989	0.987	0.966
200	5	40.50	40.37	40.34	39.85	0.990	0.990	0.988	0.980
400	5	38.99	38.99	38.98	38.71	0.990	0.990	0.989	0.985
800	5	38.32	38.28	38.28	38.14	0.990	0.990	0.990	0.988
12	10	3821.46	2157.26	51.02	–	0.978	0.247	–	0.141
13	10	755.81	518.77	50.72	666.97	0.979	0.417	0.987	0.255
30	10	56.04	58.17	47.06	53.34	0.993	0.964	0.985	0.870
40	10	50.07	51.16	45.72	48.22	0.992	0.979	0.986	0.917
50	10	47.38	47.90	44.71	45.83	0.991	0.983	0.987	0.938
100	10	42.67	42.58	41.98	41.70	0.990	0.988	0.988	0.969
200	10	40.19	40.17	40.07	39.71	0.990	0.990	0.989	0.981
400	10	38.92	38.92	38.91	38.67	0.990	0.990	0.989	0.986
800	10	38.27	38.26	38.26	38.13	0.990	0.990	0.990	0.988

Note. $\mathrm{CP_{YS\cdot L2}} = \mathrm{CP}(t^2_{\rm YS\cdot L2}(\alpha))$, $\mathrm{CP}_{{\rm YS}\cdot F2} = \mathrm{CP}(t^2_{{\rm YS}\cdot F2}(\alpha))$, $\mathrm{CP_{YKP}} = \mathrm{CP}(t^2_{\rm YKP}(\alpha))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_p(\alpha))$, $\chi^2_{20}(0.05) = 31.41$, $\chi^2_{20}(0.01) = 37.57$.

the simulated upper $100\alpha_{\rm p}$ percentiles of the $\widetilde{T}^2_{ab}$ statistic ($\widetilde{t}^2_{\rm simu\cdot Bon}(\alpha_{\rm p})$), and the approximate upper $100\alpha_{\rm p}$ percentiles of the $\widetilde{T}^2_{ab}$ statistic ($t^2_{{\rm YS\cdot L}m}(\alpha_{\rm p})$, $t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm p})$). The values of $\widetilde{t}^2_{\rm simu\cdot Bon}(\alpha_{\rm p})$ are simulated values obtained via Monte Carlo simulation. In the tables, we denote $\widetilde{t}^2_{\rm simu\cdot p}(\alpha)$, $\widetilde{t}^2_{\rm simu\cdot Bon}(\alpha_{\rm p})$, $t^2_{{\rm YS\cdot L}m}(\alpha_{\rm p})$, and $t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm p})$ as $\widetilde{t}^2_{\rm simu\cdot p}$, $\widetilde{t}^2_{\rm simu\cdot Bon}$, $t^2_{{\rm YS\cdot L}m}$, and $t^2_{{\rm YS}\cdot Fm}$, respectively. In addition, we provide the actual coverage probabilities given by

$$\mathrm{CP}(t^2_{{\rm YS\cdot L}m}(\alpha_{\rm p})) = 1 - \Pr\{\widetilde{T}^2_{\rm max\cdot p} > t^2_{{\rm YS\cdot L}m}(\alpha_{\rm p})\},$$
$$\mathrm{CP}(t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm p})) = 1 - \Pr\{\widetilde{T}^2_{\rm max\cdot p} > t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm p})\},$$

Table 3.9: Simulated and approximate values and coverage probabilities for pairwise comparisons when $m = 6$ and $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{p}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\text{CP}_{\text{YS}\cdot\text{L}m}$	$\text{CP}_{\text{YS}\cdot Fm}$	CP_{χ^2}
				$\alpha = 0.05$				
25	5	39.86	40.62	40.48	39.78	0.957	0.949	0.828
30	5	38.82	39.40	39.40	38.98	0.957	0.952	0.855
35	5	38.04	38.56	38.65	38.37	0.957	0.954	0.873
40	5	37.52	38.04	38.09	37.89	0.957	0.954	0.886
45	5	37.08	37.65	37.65	37.51	0.957	0.955	0.895
50	5	36.76	37.34	37.31	37.20	0.957	0.955	0.902
100	5	35.24	35.77	35.73	35.71	0.956	0.956	0.931
200	5	34.42	34.87	34.91	34.91	0.957	0.957	0.945
400	5	34.04	34.55	34.49	34.49	0.956	0.956	0.950
800	5	33.78	34.21	34.28	34.28	0.957	0.957	0.954
25	10	38.92	39.50	39.77	38.37	0.959	0.943	0.852
30	10	38.09	38.73	38.82	37.89	0.958	0.947	0.872
35	10	37.48	38.02	38.16	37.51	0.958	0.950	0.886
40	10	37.03	37.62	37.67	37.20	0.958	0.952	0.896
45	10	36.72	37.18	37.30	36.94	0.957	0.953	0.903
50	10	36.44	36.97	37.00	36.72	0.957	0.954	0.909
100	10	35.10	35.65	35.62	35.57	0.957	0.956	0.934
200	10	34.38	34.83	34.88	34.87	0.957	0.956	0.945
400	10	33.98	34.49	34.48	34.48	0.957	0.957	0.951
800	10	33.80	34.30	34.28	34.28	0.957	0.957	0.954
				$\alpha = 0.01$				
25	5	46.58	46.98	46.88	45.99	0.991	0.988	0.938
30	5	45.25	45.41	45.52	44.98	0.991	0.989	0.951
35	5	44.31	44.29	44.57	44.22	0.991	0.990	0.959
40	5	43.62	43.92	43.87	43.63	0.991	0.990	0.965
45	5	43.06	43.36	43.33	43.15	0.991	0.990	0.969
50	5	42.65	43.18	42.90	42.77	0.991	0.990	0.971
100	5	40.73	41.01	40.94	40.92	0.991	0.991	0.982
200	5	39.71	40.08	39.93	39.93	0.991	0.991	0.987
400	5	39.23	39.55	39.41	39.41	0.991	0.991	0.989
800	5	38.89	39.27	39.15	39.15	0.991	0.991	0.990
25	10	45.36	45.85	45.99	44.22	0.992	0.987	0.950
30	10	44.30	44.85	44.79	43.63	0.991	0.988	0.959
35	10	43.54	43.61	43.96	43.15	0.991	0.989	0.965
40	10	42.99	43.30	43.36	42.77	0.991	0.989	0.969
45	10	42.65	42.75	42.89	42.44	0.991	0.989	0.972
50	10	42.21	42.42	42.51	42.17	0.991	0.990	0.974
100	10	40.61	40.76	40.80	40.74	0.991	0.990	0.983
200	10	39.66	39.64	39.89	39.88	0.991	0.991	0.987
400	10	39.14	39.44	39.40	39.40	0.991	0.991	0.989
800	10	38.92	39.01	39.15	39.15	0.991	0.991	0.990

Note. $\text{CP}_{\text{YS}\cdot\text{L}m} = \text{CP}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{p}))$, $\text{CP}_{\text{YS}\cdot Fm} = \text{CP}(t^2_{\text{YS}\cdot Fm}(\alpha_\text{p}))$, $\text{CP}_{\chi^2} = \text{CP}(\chi^2_p(\alpha_\text{p}))$, $\alpha_\text{p} = \alpha/15$, $\chi^2_{15}(0.05/15) = 34.07$, $\chi^2_{15}(0.01/15) = 38.89$.

$$\text{CP}(\chi^2_p(\alpha_\text{p})) = 1 - \Pr\{\widetilde{T}^2_{\max\cdot\text{p}} > \chi^2_p(\alpha_\text{p})\}.$$

As Tables 3.9 and 3.10 show, the simulated values for $\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}(\alpha_\text{p})$ are larger than the simulated values for $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$. It may be seen from the tables that the approximate values of $t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{p})$ and $t^2_{\text{YS}\cdot Fm}(\alpha_\text{p})$ are closer to the simulated values of $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$ when the

Table 3.10: Simulated and approximate values and coverage probabilities for pairwise comparisons when $m = 10$ and $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\text{simu}\cdot\text{p}}$	$\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}$	$t^2_{\text{YS}\cdot\text{L}m}$	$t^2_{\text{YS}\cdot Fm}$	$\text{CP}_{\text{YS}\cdot\text{L}m}$	$\text{CP}_{\text{YS}\cdot Fm}$	CP_{χ^2}
			$\alpha = 0.05$					
25	5	41.42	41.91	42.14	41.66	0.959	0.953	0.865
30	5	40.74	41.44	41.38	41.08	0.958	0.954	0.883
35	5	40.21	40.74	40.84	40.64	0.958	0.955	0.896
40	5	39.78	40.36	40.43	40.29	0.958	0.956	0.905
45	5	39.52	40.26	40.11	40.01	0.958	0.956	0.911
50	5	39.26	39.94	39.85	39.78	0.958	0.957	0.917
100	5	38.14	38.82	38.67	38.66	0.957	0.957	0.938
200	5	37.49	37.94	38.04	38.04	0.958	0.958	0.948
400	5	37.21	37.59	37.72	37.72	0.957	0.957	0.953
800	5	37.06	37.66	37.55	37.55	0.957	0.957	0.955
25	10	40.72	41.30	41.61	40.64	0.960	0.949	0.884
30	10	40.16	40.75	40.94	40.29	0.960	0.952	0.897
35	10	39.73	40.30	40.47	40.01	0.959	0.954	0.907
40	10	39.42	39.97	40.12	39.78	0.959	0.955	0.913
45	10	39.20	39.81	39.84	39.58	0.958	0.955	0.918
50	10	38.98	39.63	39.62	39.42	0.958	0.956	0.922
100	10	38.05	38.43	38.58	38.55	0.957	0.957	0.939
200	10	37.48	38.05	38.02	38.01	0.957	0.957	0.949
400	10	37.18	37.71	37.71	37.71	0.957	0.957	0.953
800	10	37.06	37.63	37.55	37.55	0.957	0.957	0.955
			$\alpha = 0.01$					
25	5	47.37	47.72	47.75	47.18	0.991	0.989	0.957
30	5	46.47	46.93	46.83	46.47	0.991	0.990	0.965
35	5	45.87	46.13	46.18	45.94	0.991	0.990	0.970
40	5	45.36	45.40	45.69	45.52	0.991	0.990	0.973
45	5	45.05	45.41	45.30	45.18	0.991	0.990	0.975
50	5	44.75	45.20	44.99	44.90	0.991	0.990	0.977
100	5	43.34	43.74	43.57	43.55	0.991	0.991	0.985
200	5	42.60	42.50	42.82	42.82	0.991	0.991	0.988
400	5	42.18	42.55	42.43	42.43	0.991	0.991	0.989
800	5	42.00	42.48	42.23	42.23	0.991	0.991	0.990
25	10	46.47	46.82	47.12	45.94	0.992	0.988	0.965
30	10	45.77	46.13	46.31	45.52	0.992	0.989	0.970
35	10	45.30	45.62	45.74	45.18	0.991	0.990	0.974
40	10	44.91	45.17	45.31	44.90	0.991	0.990	0.976
45	10	44.61	44.81	44.98	44.67	0.991	0.990	0.978
50	10	44.43	44.87	44.71	44.47	0.991	0.990	0.979
100	10	43.24	43.29	43.46	43.42	0.991	0.991	0.985
200	10	42.59	43.27	42.79	42.78	0.991	0.991	0.988
400	10	42.17	42.29	42.42	42.42	0.991	0.991	0.990
800	10	41.96	42.11	42.23	42.23	0.991	0.991	0.990

Note. $\text{CP}_{\text{YS}\cdot\text{L}m} = \text{CP}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{p}))$, $\text{CP}_{\text{YS}\cdot Fm} = \text{CP}(t^2_{\text{YS}\cdot Fm}(\alpha_\text{p}))$, $\text{CP}_{\chi^2} = \text{CP}(\chi^2_p(\alpha_\text{p}))$, $\alpha_\text{p} = \alpha/45$, $\chi^2_{15}(0.05/45) = 37.39$, $\chi^2_{15}(0.01/45) = 42.03$.

sample size becomes large. The simulation studies show that $t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{p})$ is close to $\widetilde{t}^2_{\text{simu}\cdot\text{p}}(\alpha)$ and is a conservative approximation. Tables 3.11 and 3.12 list the results for the case of comparisons with a control. We provide $\widetilde{t}^2_{\text{simu}\cdot\text{c}}(\alpha)$, $\widetilde{t}^2_{\text{simu}\cdot\text{Bon}}(\alpha_\text{c})$, $t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{c})$, and $t^2_{\text{YS}\cdot Fm}(\alpha_\text{c})$ as well as $\text{CP}(t^2_{\text{YS}\cdot\text{L}m}(\alpha_\text{c}))$, $\text{CP}(t^2_{\text{YS}\cdot Fm}(\alpha_\text{c}))$, and $\text{CP}(\chi^2_p(\alpha_\text{c}))$. The accuracy of the approximations is similar to that in the case of pairwise comparisons.

Table 3.11: Simulated and approximate values and coverage probabilities for comparisons with a control when $m = 6$ and $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\rm simu\cdot c}$	$\widetilde{t}^2_{\rm simu\cdot Bon}$	$t^2_{{\rm YS}\cdot Lm}$	$t^2_{{\rm YS}\cdot Fm}$	${\rm CP}_{{\rm YS}\cdot Lm}$	${\rm CP}_{{\rm YS}\cdot Fm}$	${\rm CP}_{\chi^2}$
				$\alpha = 0.05$				
25	5	35.54	36.06	35.96	35.38	0.955	0.948	0.857
30	5	34.59	35.15	35.06	34.71	0.955	0.951	0.878
35	5	33.96	34.47	34.43	34.20	0.955	0.953	0.891
40	5	33.54	33.99	33.96	33.80	0.955	0.953	0.900
45	5	33.19	33.66	33.60	33.48	0.955	0.954	0.906
50	5	32.89	33.35	33.31	33.22	0.955	0.954	0.912
100	5	31.60	32.03	31.98	31.97	0.955	0.955	0.935
200	5	30.86	31.29	31.29	31.29	0.955	0.955	0.946
400	5	30.54	30.89	30.94	30.94	0.955	0.955	0.951
800	5	30.34	30.74	30.76	30.76	0.956	0.956	0.953
25	10	34.72	35.20	35.36	34.20	0.957	0.944	0.875
30	10	34.04	34.55	34.57	33.80	0.956	0.947	0.890
35	10	33.52	33.97	34.02	33.48	0.956	0.949	0.900
40	10	33.11	33.62	33.61	33.22	0.956	0.951	0.908
45	10	32.82	33.30	33.30	33.00	0.956	0.952	0.913
50	10	32.62	33.06	33.05	32.82	0.955	0.953	0.917
100	10	31.46	31.85	31.89	31.85	0.955	0.955	0.937
200	10	30.85	31.29	31.26	31.26	0.955	0.955	0.946
400	10	30.53	30.94	30.93	30.93	0.955	0.955	0.951
800	10	30.37	30.75	30.76	30.76	0.955	0.955	0.953
				$\alpha = 0.01$				
25	5	42.34	42.64	42.53	41.78	0.990	0.989	0.951
30	5	41.16	41.50	41.37	40.91	0.991	0.989	0.961
35	5	40.31	40.55	40.55	40.25	0.991	0.990	0.967
40	5	39.74	39.96	39.95	39.74	0.991	0.990	0.970
45	5	39.35	39.44	39.48	39.33	0.990	0.990	0.973
50	5	38.91	39.17	39.11	39.00	0.991	0.990	0.976
100	5	37.31	37.42	37.41	37.40	0.990	0.990	0.984
200	5	36.43	36.62	36.54	36.53	0.990	0.990	0.987
400	5	35.93	36.04	36.09	36.09	0.991	0.991	0.989
800	5	35.63	35.84	35.86	35.86	0.991	0.991	0.990
25	10	41.37	41.54	41.76	40.25	0.991	0.987	0.960
30	10	40.45	40.70	40.74	39.74	0.991	0.988	0.966
35	10	39.73	39.82	40.03	39.33	0.991	0.989	0.971
40	10	39.17	39.34	39.50	39.00	0.991	0.989	0.974
45	10	38.84	39.08	39.10	38.72	0.991	0.990	0.976
50	10	38.60	38.79	38.77	38.48	0.991	0.990	0.977
100	10	37.13	37.30	37.29	37.24	0.991	0.990	0.984
200	10	36.34	36.46	36.50	36.49	0.990	0.990	0.987
400	10	35.95	36.13	36.08	36.08	0.990	0.990	0.989
800	10	35.65	35.82	35.86	35.86	0.991	0.991	0.990

Note. ${\rm CP}_{{\rm YS}\cdot Lm} = {\rm CP}(t^2_{{\rm YS}\cdot Lm}(\alpha_{\rm c}))$, ${\rm CP}_{{\rm YS}\cdot Fm} = {\rm CP}(t^2_{{\rm YS}\cdot Fm}(\alpha_{\rm c}))$, ${\rm CP}_{\chi^2} = {\rm CP}(\chi^2_p(\alpha_{\rm c}))$, $\alpha_{\rm c} = \alpha/5$, $\chi^2_{15}(0.05/5) = 30.58$, $\chi^2_{15}(0.01/5) = 35.63$.

3.3 Conclusions

In conclusion, we have developed the approximate upper percentiles of the simplified T^2-type statistic for testing the equaliy of mean vectors with k-step monotone missing data under the two-sample problem. As the additional results to obtain the simplified T^2-type statistic, we derived the MLEs of the mean vectors and the covariance matrix for the m-

Table 3.12: Simulated and approximate values and coverage probabilities for comparisons with a control when $m = 10$ and $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

Sample Size		Upper Percentile				Coverage Probability		
$n_1^{(\ell)}$	$n_2^{(\ell)} = \cdots = n_5^{(\ell)}$	$\widetilde{t}^2_{\mathrm{simu\cdot c}}$	$\widetilde{t}^2_{\mathrm{simu\cdot Bon}}$	$t^2_{\mathrm{YS}\cdot Lm}$	$t^2_{\mathrm{YS}\cdot Fm}$	$\mathrm{CP}_{\mathrm{YS}\cdot Lm}$	$\mathrm{CP}_{\mathrm{YS}\cdot Fm}$	CP_{χ^2}
				$\alpha = 0.05$				
25	5	35.64	36.24	36.29	35.91	0.957	0.953	0.894
30	5	35.07	35.71	35.68	35.44	0.957	0.954	0.906
35	5	34.65	35.28	35.25	35.09	0.957	0.955	0.914
40	5	34.33	34.88	34.92	34.81	0.957	0.956	0.920
45	5	34.07	34.66	34.66	34.58	0.957	0.956	0.925
50	5	33.92	34.54	34.46	34.40	0.957	0.956	0.927
100	5	32.96	33.52	33.50	33.49	0.957	0.957	0.943
200	5	32.43	32.98	33.00	33.00	0.958	0.958	0.951
400	5	32.19	32.72	32.74	32.74	0.957	0.957	0.954
800	5	32.10	32.65	32.60	32.60	0.957	0.957	0.955
25	10	35.06	35.63	35.86	35.09	0.959	0.950	0.906
30	10	34.59	35.23	35.33	34.81	0.959	0.953	0.915
35	10	34.32	34.91	34.95	34.58	0.958	0.953	0.920
40	10	34.05	34.67	34.67	34.40	0.957	0.954	0.925
45	10	33.84	34.43	34.45	34.24	0.957	0.955	0.928
50	10	33.67	34.25	34.27	34.11	0.957	0.956	0.931
100	10	32.87	33.40	33.43	33.41	0.957	0.957	0.944
200	10	32.44	32.99	32.98	32.97	0.957	0.957	0.950
400	10	32.23	32.76	32.73	32.73	0.957	0.957	0.953
800	10	32.07	32.64	32.60	32.60	0.957	0.957	0.955
				$\alpha = 0.01$				
25	5	41.78	41.96	42.14	41.66	0.991	0.990	0.968
30	5	41.19	41.60	41.38	41.08	0.991	0.990	0.972
35	5	40.57	40.91	40.84	40.64	0.991	0.990	0.976
40	5	40.12	40.33	40.43	40.29	0.991	0.990	0.978
45	5	39.85	39.97	40.11	40.01	0.991	0.990	0.980
50	5	39.60	39.88	39.85	39.78	0.991	0.990	0.981
100	5	38.45	38.67	38.67	38.66	0.991	0.991	0.986
200	5	37.78	38.02	38.04	38.04	0.991	0.991	0.989
400	5	37.45	37.82	37.72	37.72	0.991	0.991	0.990
800	5	37.34	37.44	37.55	37.55	0.991	0.991	0.990
25	10	41.07	41.21	41.61	40.64	0.991	0.989	0.973
30	10	40.48	40.76	40.94	40.29	0.991	0.989	0.976
35	10	40.12	40.32	40.47	40.01	0.991	0.990	0.978
40	10	39.82	39.98	40.12	39.78	0.991	0.990	0.980
45	10	39.52	39.76	39.84	39.58	0.991	0.990	0.981
50	10	39.32	39.43	39.62	39.42	0.991	0.990	0.982
100	10	38.36	38.64	38.58	38.55	0.991	0.991	0.987
200	10	37.79	38.04	38.02	38.01	0.991	0.991	0.989
400	10	37.51	37.69	37.71	37.71	0.991	0.991	0.990
800	10	37.30	37.62	37.55	37.55	0.991	0.991	0.990

Note. $\mathrm{CP}_{\mathrm{YS\cdot L}m} = \mathrm{CP}(t^2_{\mathrm{YS\cdot L}m}(\alpha_{\mathrm{c}}))$, $\mathrm{CP}_{\mathrm{YS}\cdot Fm} = \mathrm{CP}(t^2_{\mathrm{YS}\cdot Fm}(\alpha_{\mathrm{c}}))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_p(\alpha_{\mathrm{c}}))$, $\alpha_{\mathrm{c}} = \alpha/9$, $\chi^2_{15}(0.05/9) = 32.47$, $\chi^2_{15}(0.01/9) = 37.39$.

sample problem in the case of k-step monotone missing data. Further, we presented the approximate simultaneous confidence intervals for pairwise comparisons among mean vectors and comparisons with a control using Bonferroni's approximation. The proposed approximate values can be easily calculated, and the accuracy of the approximations is considerably higher than that of the χ^2 approximations in almost all cases.

Chapter 4

Likelihood ratio test in one-sample problem

In this chapter, we consider the LRT for a normal mean vector when the data have a monotone pattern of missing observations. We derive MT and MLRT statistics by using decomposition of the LR. Further, we investigate the accuracy of the upper percentiles of this test statistic by Monte Carlo simulation.

In statistical data analyses, testing hypotheses with missing data is an important problem. In this chapter, we consider the one-sample test for a normal mean vector with monotone missing data. The one-sample problem of the test for the mean vector with monotone missing data has been discussed by many authors. For a discussion of the LRT statistic, see Krishnamoorthy and Pannala (1998) and Seko et al. (2012). For a general missing data pattern, Srivastava (1985) discussed the LRT for mean vectors, and Seo and Srivastava (2000) gave a test of equality of means and the simultaneous confidence intervals. In addition, for non-missing and multivariate normality, the asymptotic expansion for LR-criterion was discussed by Muirhead (1982), Siotani et al. (1985), Kakizawa (1996), and Anderson (2003), among others.

In this chapter, for the one-sample test of the mean vector, we give the LRT statistic for k-step monotone missing data and derive MLRT statistic by using the decomposition (see Bhargava, 1962 and Krishnamoorthy and Pannala, 1998). In the process deriving the MLRT statistic, we give asymptotic expansions of the LR of the test for a mean vector and those of subvector. For the test for a subvector under multivariate normality, see e.g., Siotani et al. (1985). This chapter is organized in the following way, which is a summary of Yagi et al. (2017a). In Section 4.1, we present the assumptions and notation. In Section 4.2, we derive the LRT statistic, MT and MLRT statistics, which converge to the χ^2 distribution faster than the LRT statistic as the sample size tends to infinity. That is, we derive transformations with Bartlett adjustments. Indeed, it is well known that Bartlett adjustment yields an improvement on the chi-squared approximation to the LRT statistic. In Section 4.3, some simulation results for three- and five-step monotone missing data cases are presented to investigate the accuracy of the upper percentiles of the null distributions of MT and MLRT statistics. Finally, in Section 4.4, we state our conclusions.

4.1 Assumptions and notation

We consider the one-sample problem of testing for a mean vector with a k-step monotone missing data pattern. As with the simplified T^2-type test statistic case in Section 2.2, let $\boldsymbol{x}_i$ be a $p_i \times 1$ normal random vector with the mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$, where $\boldsymbol{\mu}_i = (\boldsymbol{\mu})_i = (\mu_1, \mu_2, \ldots, \mu_{p_i})'$, and $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ principal submatrix of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}_1)$ with $p = p_1 > p_2 > \cdots > p_k > 0$. Further, let $\boldsymbol{x}_i$, $i = 1, 2, \ldots, k$ be mutually independent. Suppose that $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{in_i}$ are independent and identically distributed samples from $\boldsymbol{x}_i$, $i = 1, 2, \ldots, k$, where $n_1 > p$. As for the partitions of $\boldsymbol{\Sigma}$, for $1 \le i < j \le k$, let $(\boldsymbol{\Sigma}_i)_j$ be the principal submatrix of $\boldsymbol{\Sigma}_i$ of order $p_j \times p_j$; we define

$$\boldsymbol{\Sigma}_i = (\boldsymbol{\Sigma}_1)_i, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{i2} \\ \boldsymbol{\Sigma}_{i2}' & \boldsymbol{\Sigma}_{i3} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{i-1} = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{(i-1,2)} \\ \boldsymbol{\Sigma}_{(i-1,2)}' & \boldsymbol{\Sigma}_{(i-1,3)} \end{pmatrix},$$

and

$$\boldsymbol{\Sigma}_{(i-1,3)\cdot i} = \boldsymbol{\Sigma}_{(i-1,3)} - \boldsymbol{\Sigma}_{(i-1,2)}' \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Sigma}_{(i-1,2)}, \quad i = 2, 3, \ldots, k.$$

For example, when $k = 3$, we can express $\boldsymbol{\Sigma}_1$ as

$$\boldsymbol{\Sigma}_1 = \left(\begin{array}{cc|c} \overbrace{\boldsymbol{\Sigma}_3}^{p_3} & \overbrace{\boldsymbol{\Sigma}_{(2,2)}}^{p_2-p_3} & \overbrace{\phantom{\boldsymbol{\Sigma}_{(1,2)}}}^{p_1-p_2} \\ \boldsymbol{\Sigma}_{(2,2)}' & \boldsymbol{\Sigma}_{(2,3)} & \boldsymbol{\Sigma}_{(1,2)} \\ \hline \multicolumn{2}{c|}{\boldsymbol{\Sigma}_{(1,2)}'} & \boldsymbol{\Sigma}_{(1,3)} \end{array}\right) \begin{array}{l} \}p_3 \\ \}p_2 - p_3 \\ \}p_1 - p_2 \end{array}.$$

4.2 LRT, MT and MLRT statistics

Consider the null hypothesis

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$$

against the alternative $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, where $\boldsymbol{\mu}_0$ is known. Without loss of generality, we can assume that $\boldsymbol{\mu}_0 = \boldsymbol{0}$. Then, the LR is given by

$$\lambda = \prod_{i=1}^{k} \left(\frac{|\widehat{\boldsymbol{\Sigma}}_i|}{|\widetilde{\boldsymbol{\Sigma}}_i|} \right)^{\frac{1}{2}n_i},$$

where $\widehat{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_1, and $\widetilde{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_0. Let

$$\boldsymbol{E}_i = \sum_{j=1}^{n_i} (\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}}_i)', \quad \overline{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}, \quad i = 1, 2, \ldots, k,$$

$$\boldsymbol{d}_1 = \overline{\boldsymbol{x}}_1, \quad \boldsymbol{d}_i = \frac{n_i}{N_{i+1}} \Big[\overline{\boldsymbol{x}}_i - \frac{1}{N_i} \sum_{j=1}^{i-1} n_j (\overline{\boldsymbol{x}}_j)_i \Big], \quad i = 2, 3, \ldots, k,$$

$$N_1 = 0, \quad N_{i+1} = N_i + n_i \left(= \sum_{j=1}^{i} n_j \right), \quad i = 1, 2, \ldots, k.$$

Table 4.1: The upper percentiles of Q and the actual type I error

	$(p_1, p_2, p_3) = (8, 4, 2)$		$(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$	
n_1	$q(\alpha)$	α_Q	$q(\alpha)$	α_Q
		$\alpha = 0.05$		
20	20.74	16.71	46.86	47.84
30	18.53	11.24	33.86	23.03
40	17.70	9.25	30.91	16.13
50	17.18	8.17	29.52	13.05
100	16.29	6.42	27.15	8.37
200	15.88	5.65	26.03	6.53
400	15.68	5.31	25.53	5.74
∞	15.51	5.00	25.00	5.00

Note. $n_2 = n_3 = n_4 = n_5 = 10$, $\chi^2_8(0.05) = 15.51$, $\chi^2_{15}(0.05) = 25.00$.

Then, we can express $\widehat{\boldsymbol{\Sigma}}$ concretely (see Theorem 2.2) as

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n_1}\boldsymbol{H}_1 + \sum_{i=2}^{k} \frac{1}{N_{i+1}} \boldsymbol{F}_i \left[\boldsymbol{H}_i - \frac{n_i}{N_i}\boldsymbol{L}_{i-1,1}\right] \boldsymbol{F}_i',$$

where

$$\begin{aligned}
&\boldsymbol{H}_1 = \boldsymbol{E}_1,\ \boldsymbol{H}_i = \boldsymbol{E}_i + \frac{N_i N_{i+1}}{n_i}\boldsymbol{d}_i\boldsymbol{d}_i',\ i = 2, 3, \ldots, k, \\
&\boldsymbol{L}_1 = \boldsymbol{H}_1,\ \boldsymbol{L}_i = (\boldsymbol{L}_{i-1})_i + \boldsymbol{H}_i,\ i = 2, 3, \ldots, k, \\
&\boldsymbol{L}_{i1} = (\boldsymbol{L}_i)_{i+1},\ \boldsymbol{L}_i = \begin{pmatrix} \boldsymbol{L}_{i1} & \boldsymbol{L}_{i2} \\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3} \end{pmatrix},\ i = 1, 2, \ldots, k-1, \\
&\boldsymbol{G}_1 = \boldsymbol{I}_{p_1},\ \boldsymbol{G}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{L}_{i2}'\boldsymbol{L}_{i1}^{-1} \end{pmatrix},\ i = 1, 2, \ldots, k-1, \\
&\boldsymbol{F}_1 = \boldsymbol{G}_1,\ \boldsymbol{F}_i = \boldsymbol{F}_{i-1}\boldsymbol{G}_i,\ i = 2, 3, \ldots, k.
\end{aligned}$$

By the same derivation in Jinadasa and Tracy (1992), it holds that the MLE of $\boldsymbol{\Sigma}$ under H_0, $\widetilde{\boldsymbol{\Sigma}}$ is equal to $\widehat{\boldsymbol{\Sigma}}$ with $\overline{\boldsymbol{x}}_i = \boldsymbol{0}$, $i = 1, 2, \ldots, k$. That is, $\widetilde{\boldsymbol{\Sigma}}$ can be obtained as $\widehat{\boldsymbol{\Sigma}}$ in the case that

$$\boldsymbol{H}_i = \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}\boldsymbol{x}_{ij}', \quad i = 1, 2, \ldots, k.$$

We note that the null distribution of the LRT statistic $Q(= -2\log\lambda)$ is asymptotically a χ^2 distribution with p degrees of freedom. However, it may be noted that the upper percentiles of the χ^2 distribution are not a good approximation to those of the LRT statistic when the sample size is not large. For example, Table 4.1 gives the simulated values of the upper 100α percentiles of Q, $q(\alpha)$ and the actual type I error rates, $\alpha_Q/100 = \Pr\{Q > \chi_p^2(\alpha)\}$ for the three- and five-step monotone missing data cases, where $\chi_p^2(\alpha)$ is the upper 100α percentile of the χ^2 distribution with p degrees of freedom. It may be seen that the upper percentiles of the χ^2 distribution are useful as an approximation to the upper percentiles of Q for cases in which the sample size is considerably large. Therefore, we consider the MLRT statistic whose null distribution is closer to the χ^2 distribution than that of the LRT statistic even when the sample size is small. In particular, we derive an asymptotic expansion for the distribution function of the LRT statistic in a situation when $n_1 \to \infty$ with $q_i = n_i/n_1 \to \delta_i \in [0, \infty)$, $i = 2, 3, \ldots, k$. Using the notation in Section 2, we can write λ as

$$\lambda = \prod_{i=1}^{k} \lambda_i,$$

where

$$\lambda_1 = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_k|}{|\widetilde{\boldsymbol{\Sigma}}_k|} \right)^{\frac{N_{k+1}}{2}}, \quad \lambda_i = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}|}{|\widetilde{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}|} \right)^{\frac{N_{k-i+2}}{2}}, \quad i = 2, 3, \ldots, k,$$

$$N_{k+1} = \sum_{j=1}^{k} n_j, \quad N_{k-i+2} = \sum_{j=1}^{k-i+1} n_j,$$

$\widehat{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}$ and $\widetilde{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}$ are given by

$$\widehat{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2} = \widehat{\boldsymbol{\Sigma}}_{(k-i+1,3)} - \widehat{\boldsymbol{\Sigma}}'_{(k-i+1,2)} \widehat{\boldsymbol{\Sigma}}_{k-i+2}^{-1} \widehat{\boldsymbol{\Sigma}}_{(k-i+1,2)}$$

and

$$\widetilde{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2} = \widetilde{\boldsymbol{\Sigma}}_{(k-i+1,3)} - \widetilde{\boldsymbol{\Sigma}}'_{(k-i+1,2)} \widetilde{\boldsymbol{\Sigma}}_{k-i+2}^{-1} \widetilde{\boldsymbol{\Sigma}}_{(k-i+1,2)},$$

respectively. We note that the values of λ_i, $i = 1, 2, \ldots, k$ are mutually independent.

Further, we consider the following hypotheses:

$$\begin{aligned} H_{01} :& \boldsymbol{\mu}_k = \mathbf{0} \text{ vs. } H_{11} : \boldsymbol{\mu}_k \neq \mathbf{0}, \\ H_{0i} :& \boldsymbol{A}_{k-i+1}\boldsymbol{\mu}_{k-i+1} = \mathbf{0} \text{ given } \boldsymbol{\mu}_{k-i+2} = \mathbf{0} \\ & \quad \text{vs. } H_{1i} : \boldsymbol{A}_{k-i+1}\boldsymbol{\mu}_{k-i+1} \neq \mathbf{0} \text{ given } \boldsymbol{\mu}_{k-i+2} = \mathbf{0}, \quad i = 2, 3, \ldots, k, \end{aligned}$$

where $\boldsymbol{A}_{k-i+1} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_{p_{k-i+1}-p_{k-i+2}} \end{pmatrix}$ is a $(p_{k-i+1} - p_{k-i+2}) \times p_{k-i+1}$ matrix.

Let the parameter spaces of Ω_0, Ω_i, $i = 1, 2, \ldots, k-1$ and Ω_k be

$$\begin{aligned} \Omega_0 =& \{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : -\infty < \mu_j < \infty,\ j = 1, 2, \ldots, p,\ \boldsymbol{\Sigma} > \boldsymbol{O}\}, \\ \Omega_i =& \{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : \boldsymbol{\mu}_{k-i+1} = \mathbf{0},\ -\infty < \mu_j < \infty,\ j = p_{k-i+1}+1, p_{k-i+1}+2, \ldots, p, \\ & \quad \boldsymbol{\Sigma} > \boldsymbol{O}\},\ i = 1, 2, \ldots, k-1, \\ \Omega_k =& \{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : \boldsymbol{\mu} = \mathbf{0},\ \boldsymbol{\Sigma} > \boldsymbol{O}\}, \end{aligned}$$

respectively. Then, the LR for the hypothesis H_{0i} is given by

$$\lambda_{0i} = \frac{\max_{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \Omega_i} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}{\max_{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \Omega_{i-1}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}, \quad i = 1, 2, \ldots, k.$$

Therefore, since λ_i is equal to λ_{0i}, we have

$$\lambda = \prod_{i=1}^{k} \lambda_{0i}.$$

That is, we note $-2\log\lambda_1$ is the usual LRT statistic of the test for p_k dimensional mean vector, and $-2\log\lambda_i$, $i = 2, 3, \ldots, k$ is LRT statistic of the test for a subvector. The above result is obtained by Krishnamoorthy and Pannala (1998). Then, we have the following theorem.

Theorem 4.1 (Yagi et al., 2017a) *Suppose that $\boldsymbol{x}_{ij}$ is distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i$, where $p = p_1 > p_2 > \cdots > p_k > p_{k+1} = 0$ and $n_1 > p$. Then, when the null hypothesis H_{0i} is true, the cumulative distribution function of $Q_i^*(= -2\rho_i \log\lambda_i)$ can be expressed for large N_{k-i+2} as*

$$\Pr(Q_i^* \le x) = G_{p_{k-i+1}-p_{k-i+2}}(x) + O(N_{k-i+2}^{-2}), \quad i = 1, 2, \ldots, k,$$

where

$$\rho_i = 1 - \frac{1}{2N_{k-i+2}}(p_{k-i+1} + p_{k-i+2} + 2), \quad i = 1, 2, \ldots, k,$$

$G_p(x)$ is the distribution function of a χ^2-variate with p degrees of freedom.

Proof. First we derive an asymptotic expansion of the characteristic function of $Q_1(= -2\log\lambda_1)$. We use the following notation to simplify setting. Let $\boldsymbol{y}_1, \ \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N_{k+1}}$ be distributed as p_k dimensional multivariate normal distribution. Then λ_1 can be written as

$$\lambda_1 = \left(\frac{|\boldsymbol{U}_k|}{|\boldsymbol{U}_k + N_{k+1}\overline{\boldsymbol{y}}_k\overline{\boldsymbol{y}}_k'|} \right)^{\frac{N_{k+1}}{2}},$$

where

$$\overline{\boldsymbol{y}}_k = \frac{1}{N_{k+1}} \sum_{j=1}^{N_{k+1}} \boldsymbol{y}_j, \quad \boldsymbol{U}_k = \sum_{j=1}^{N_{k+1}} (\boldsymbol{y}_j - \overline{\boldsymbol{y}}_k)(\boldsymbol{y}_j - \overline{\boldsymbol{y}}_k)'.$$

Therefore, expanding $Q_1(= -2\log\lambda_1)$ by the perturbation method and calculating the characteristic function, we obtain

$$\mathrm{E}[\exp(itQ_1)] = (1-2it)^{-\frac{p_k}{2}} \left[1 + \frac{\beta_1}{N_{k+1}} \left\{ 1 - (1-2it)^{-1} \right\} \right] + O(N_{k+1}^{-2}),$$

where

$$\beta_1 = -\frac{1}{4} p_k(p_k + 2).$$

Inverting the characteristic function, we have

$$\Pr(Q_1 \le x) = G_{p_k}(x) + \frac{\beta_1}{N_{k+1}} \left[G_{p_k}(x) - G_{p_k+2}(x) \right] + O(N_{k+1}^{-2}).$$

Therefore, if $\rho_1 = 1 - (p_k + 2)/(2N_{k+1})$, then the cumulative distribution function of $Q_1^*(= -2\rho_1 \log \lambda_1)$ is given by

$$\Pr(Q_1^* \le x) = G_{p_k}(x) + O(N_{k+1}^{-2}).$$

Similar to the case of Q_1, we consider the cumulative distribution function of $Q_i(= -2\log \lambda_i)$, $i = 2, 3, \dots, k$. Let $\boldsymbol{y}_1,\ \boldsymbol{y}_2, \dots, \boldsymbol{y}_{N_{k-i+2}}$ be distributed as $N_{p_{k-i+1}}(\boldsymbol{\eta}, \boldsymbol{\Delta})$, where $\boldsymbol{\eta}$ is a $p_{k-i+1} \times 1$ mean vector and $\boldsymbol{\Delta}$ is a $p_{k-i+1} \times p_{k-i+1}$ covariance matrix. In order to be the notation shorter, we omit the index i of $\boldsymbol{\eta}$ and $\boldsymbol{\Delta}$. Further let $\boldsymbol{\eta}$ and $\boldsymbol{\Delta}$ be partitioned as

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} \begin{matrix} \}r_i \\ \}s_i \end{matrix}, \quad \boldsymbol{\Delta} = \left(\begin{array}{c|c} \overbrace{\boldsymbol{\Delta}_{11}}^{r_i} & \overbrace{\boldsymbol{\Delta}_{12}}^{s_i} \\ \hline \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} \end{array}\right) \begin{matrix} \}r_i \\ \}s_i \end{matrix},$$

where $r_i = p_{k-i+2}$, $s_i = p_{k-i+1} - p_{k-i+2}$. Then, since λ_i is the LR for testing

$$H_i : \boldsymbol{\eta}_2 = \boldsymbol{0} \text{ given } \boldsymbol{\eta}_1 = \boldsymbol{0} \text{ vs. } K_i : \boldsymbol{\eta}_2 \ne \boldsymbol{0} \text{ given } \boldsymbol{\eta}_1 = \boldsymbol{0}, \quad i = 2, 3, \dots, k,$$

we can write

$$\lambda_i = \left(\frac{1 + N_{k-i+2}\overline{\boldsymbol{y}}_1' \boldsymbol{U}_{11}^{-1} \overline{\boldsymbol{y}}_1}{1 + N_{k-i+2}\overline{\boldsymbol{y}}' \boldsymbol{U}^{-1} \overline{\boldsymbol{y}}} \right)^{\frac{N_{k-i+2}}{2}},$$

where

$$\overline{\boldsymbol{y}} = \frac{1}{N_{k-i+2}} \sum_{j=1}^{N_{k-i+2}} \boldsymbol{y}_j, \quad \boldsymbol{U} = \sum_{j=1}^{N_{k-i+2}} (\boldsymbol{y}_j - \overline{\boldsymbol{y}})(\boldsymbol{y}_j - \overline{\boldsymbol{y}})',$$

and

$$\overline{\boldsymbol{y}} = \begin{pmatrix} \overline{\boldsymbol{y}}_1 \\ \overline{\boldsymbol{y}}_2 \end{pmatrix} \begin{matrix} \}r_i \\ \}s_i \end{matrix}, \quad \boldsymbol{U} = \left(\begin{array}{c|c} \overbrace{\boldsymbol{U}_{11}}^{r_i} & \overbrace{\boldsymbol{U}_{12}}^{s_i} \\ \hline \boldsymbol{U}_{21} & \boldsymbol{U}_{22} \end{array}\right) \begin{matrix} \}r_i \\ \}s_i \end{matrix}.$$

Without loss of generality, we can assume that $\boldsymbol{\eta} = \boldsymbol{0}$, and $\boldsymbol{\Delta} = \boldsymbol{I}$. Let

$$\overline{\boldsymbol{y}} = \frac{1}{\sqrt{N_{k-i+2}}} \boldsymbol{z}, \quad \frac{1}{N_{k-i+2} - 1} \boldsymbol{U} = \boldsymbol{I} + \frac{1}{\sqrt{N_{k-i+2}}} \boldsymbol{V}.$$

We use partitions of $\boldsymbol{z}$ and $\boldsymbol{V}$ as

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{pmatrix} \begin{matrix} \}r_i \\ \}s_i \end{matrix}, \quad \boldsymbol{V} = \left(\begin{array}{c|c} \overbrace{\boldsymbol{V}_{11}}^{r_i} & \overbrace{\boldsymbol{V}_{12}}^{s_i} \\ \hline \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{array}\right) \begin{matrix} \}r_i \\ \}s_i \end{matrix}.$$

Then, we can expand Q_i as

$$\begin{aligned} Q_i = {} & \boldsymbol{z}'\boldsymbol{z} - \boldsymbol{z}_1'\boldsymbol{z}_1 - \frac{1}{\sqrt{N_{k-i+2}}} (\boldsymbol{z}'\boldsymbol{V}\boldsymbol{z} - \boldsymbol{z}_1'\boldsymbol{V}_{11}\boldsymbol{z}_1) \\ & + \frac{1}{N_{k-i+2}} \Big\{ \boldsymbol{z}'\boldsymbol{V}^2\boldsymbol{z} - \boldsymbol{z}_1'\boldsymbol{V}_{11}^2\boldsymbol{z}_1 - \frac{1}{2}(\boldsymbol{z}'\boldsymbol{z} - \boldsymbol{z}_1'\boldsymbol{z}_1)(\boldsymbol{z}'\boldsymbol{z} + \boldsymbol{z}_1'\boldsymbol{z}_1 - 2) \Big\} + O_p(N_{k-i+2}^{-\frac{3}{2}}). \end{aligned}$$

Hence, for $j = 2, 3, \ldots, k$,

$$E[\exp\{it(Q_j)\}] = E[\exp\{it(\boldsymbol{z}_2'\boldsymbol{z}_2)\}] + E[(X + \frac{1}{2}X^2)\exp\{it(\boldsymbol{z}_2'\boldsymbol{z}_2)\}] + O(N_{k-j+2}^{-\frac{3}{2}}),$$

where $i = \sqrt{-1}$ and

$$\begin{aligned} X = it\Big[&-\frac{1}{\sqrt{N_{k-j+2}}}(\boldsymbol{z}'\boldsymbol{V}\boldsymbol{z} - \boldsymbol{z}_1'\boldsymbol{V}_{11}\boldsymbol{z}_1) \\ &+ \frac{1}{N_{k-j+2}}\Big\{\boldsymbol{z}_1'\boldsymbol{V}_{12}\boldsymbol{V}_{21}\boldsymbol{z}_1 + 2\boldsymbol{z}_1'(\boldsymbol{V}_{11}\boldsymbol{V}_{12} + \boldsymbol{V}_{12}\boldsymbol{V}_{22})\boldsymbol{z}_2 \\ &+ \boldsymbol{z}_2'(\boldsymbol{V}_{21}\boldsymbol{V}_{12} + \boldsymbol{V}_{22}^2)\boldsymbol{z}_2 - \frac{1}{2}\boldsymbol{z}_2'\boldsymbol{z}_2(2\boldsymbol{z}_1'\boldsymbol{z}_1 + \boldsymbol{z}_2'\boldsymbol{z}_2 - 2)\Big\}\Big]. \end{aligned}$$

Therefore, after calculating the expectation, we obtain

$$\mathrm{E}[\exp(itQ_j)] = (1-2it)^{-\frac{s_j}{2}}\Big[1 + \frac{\beta_j}{N_{k-j+2}}\left\{1 - (1-2it)^{-1}\right\}\Big] + O(N_{k-j+2}^{-2}),$$

where

$$\beta_j = -\frac{1}{4}s_j(2r_j + s_j + 2),$$

and hence

$$\Pr(Q_j \leq x) = G_{s_j}(x) + \frac{\beta_j}{N_{k-j+2}}\left[G_{s_j}(x) - G_{s_j+2}(x)\right] + O(N_{k-j+2}^{-2}).$$

Therefore, if $\rho_i = 1 - (2r_i + s_i + 2)/(2N_{k-i+2})$, then the cumulative distribution function of $Q_i^*(= -2\rho_i \log \lambda_i)$ is given by

$$\Pr(Q_i^* \leq x) = G_{s_i}(x) + O(N_{k-i+2}^{-2}),$$

and the proof is complete. □

Using Theorem 4.1, we can give the MT statistic

$$Q^* = \sum_{i=1}^{k} Q_i^*$$

with an improved chi-squared approximation. However, this transformation statistic Q^* is not always monotone. For the monotone transformation, see Fujikoshi (2000). On the other hand, by gathering up the expanded results for the characteristic functions of Q_i, $i = 1, 2, \ldots, k$, we obtain the following theorem.

Theorem 4.2 (Yagi et al., 2017a)
Under H_0, the cumulative distribution function of $Q^\dagger(= -2\rho \log \lambda)$ can be expressed for large n_1 as

$$\Pr(Q^\dagger \leq x) = G_p(x) + O(n_1^{-2}),$$

where

$$\rho = 1 - \frac{1}{2n_1 p}\sum_{i=1}^{k}\frac{1}{m_{k-i+1}}(p_{k-i+1} - p_{k-i+2})(p_{k-i+1} + p_{k-i+2} + 2),$$

$m_{k-i+1} = 1 + \sum_{j=2}^{k-i+1} q_j$, $q_j = n_j/n_1$, *and* $p_{k+1} = 0$.

We note that the value of ρ coincides with that of Krishnamoorthy and Pannala (1998) when $k = 2$.

4.3 Simulation studies

In this section, we study the numerical accuracy of the upper percentiles of the MT and MLRT statistics using the actual type I error rates. In order to investigate the accuracy of the approximation, we compute the upper percentiles of Q, Q^* and $Q^\dagger$ with monotone missing data by Monte Carlo simulation. For each parameter, the simulation was executed 10^6 times using normal random vectors generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$, $i = 1, 2, \ldots, k$.

In Tables 4.2–4.5, we provide the simulated upper 100α percentiles of Q, Q^* and $Q^\dagger$ for the three-step and five-step cases. Further, we provide the actual type I error rates, α_Q, α_{Q^*} and $\alpha_{Q^\dagger}$ given by

$$\frac{\alpha_Q}{100} = \Pr\{Q > \chi_p^2(\alpha)\}, \quad \frac{\alpha_{Q^*}}{100} = \Pr\{Q^* > \chi_p^2(\alpha)\},$$

and

$$\frac{\alpha_{Q^\dagger}}{100} = \Pr\{Q^\dagger > \chi_p^2(\alpha)\},$$

respectively, where $\chi_p^2(\alpha)$ is the upper 100α percentile of the χ^2 distribution with p degrees of freedom.

It may be noted from Tables 4.2–4.5 that each value of $q(\alpha)$, $q^*(\alpha)$ and $q^\dagger(\alpha)$ is closer to the upper percentiles of the χ^2 distribution with p degrees of freedom, $\chi_p^2(\alpha)$, when n_1 becomes large. It is seen from Tables 4.2–4.4 that $q^*(\alpha)$ for $(p_1, p_2, p_3) = (8, 4, 2)$ and $(15, 12, 9)$ is a considerably good approximate value when n_1 is greater than 20. Similarly, it is seen from Table 4.5 that $q^*(\alpha)$ for $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$ is a considerably good approximate value when n_1 is greater than 20 without regard to the sample size of n_i, $i \geq 2$. As for $q^\dagger(\alpha)$, it is seen from Tables 4.2 and 4.3 that $q^\dagger(\alpha)$ for $(p_1, p_2, p_3) = (8, 4, 2)$ is a good approximate value when n_1 is greater than 30. Similarly, it is seen from Table 4.4 that $q^\dagger(\alpha)$ for $(p_1, p_2, p_3) = (15, 12, 9)$ is a good approximate value when n_1 is greater than 50. For the case of five-step monotone missing data in Table 4.5, $q^\dagger(\alpha)$ for $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$ is a good approximation to $\chi_p^2(\alpha)$ when n_1 is greater than 50. It may be noted from the simulation results that the MT statistic Q^* converges to the χ^2 distribution faster than the MLRT statistic $Q^\dagger$ in almost all cases, including the case of unbalanced sample sizes.

Table 4.2: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates when $(p_1, p_2, p_3) = (8, 4, 2)$ and $\alpha = 0.05$

n_1	n_2	n_3	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_Q	α_{Q^*}	$\alpha_{Q^\dagger}$
10	10	10	41.98	16.77	24.49	56.66	**7.20**	22.06
20	10	10	20.74	15.58	16.16	16.71	**5.12**	6.15
30	10	10	18.53	15.52	15.72	11.24	**5.02**	5.37
40	10	10	17.70	15.55	15.65	9.25	**5.07**	5.24
50	10	10	17.18	15.52	15.57	8.17	**5.02**	5.11
100	10	10	16.29	15.50	15.51	6.42	4.98	**5.00**
200	10	10	15.88	15.48	15.49	5.65	4.96	**4.97**
400	10	10	15.68	15.49	15.49	5.31	**4.97**	**4.97**
10	20	20	41.70	16.77	25.30	55.54	**7.21**	23.63
20	20	20	20.55	15.57	16.27	16.00	**5.11**	6.29
30	20	20	18.43	15.55	15.78	10.88	**5.07**	5.45
40	20	20	17.54	15.50	15.60	8.97	**4.99**	5.16
50	20	20	17.12	15.52	15.58	8.03	**5.01**	5.12
100	20	20	16.29	15.52	15.53	6.37	**5.01**	5.03
200	20	20	15.89	15.51	15.51	5.67	**5.00**	**5.00**
400	20	20	15.69	15.49	15.50	5.30	4.97	**4.98**
10	50	50	41.40	16.77	26.03	54.61	**7.19**	25.01
20	50	50	20.29	15.56	16.36	15.40	**5.09**	6.46
30	50	50	18.21	15.51	15.79	10.45	**5.00**	5.47
40	50	50	17.44	15.53	15.66	8.74	**5.04**	5.26
50	50	50	17.01	15.52	15.60	7.83	**5.02**	5.15
100	50	50	16.24	15.51	15.53	6.30	**5.01**	5.03
200	50	50	15.91	15.53	15.54	5.68	**5.04**	5.05
400	50	50	15.69	15.50	15.50	5.30	**4.99**	**4.99**
30	30	30	18.30	15.52	15.76	10.65	**5.01**	5.43
40	40	40	17.47	15.53	15.65	8.78	**5.03**	5.24
100	100	100	16.18	15.49	15.51	6.22	4.98	**5.00**
200	200	200	15.83	15.50	15.50	5.57	4.98	**4.99**
400	400	400	15.68	15.51	15.51	5.28	**5.01**	**5.01**

Note. $\chi^2_8(0.05) = 15.51$. The closest to $\alpha(= 0.05)$ in the values α_Q, α_{Q^*} and $\alpha_{Q^\dagger}$ of each row is boldface.

Table 4.3: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates when $(p_1, p_2, p_3) = (8, 4, 2)$, $\alpha = 0.05$, and $n_2 \neq n_3$

n_1	n_2	n_3	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_Q	α_{Q^*}	$\alpha_{Q^\dagger}$
10	10	20	42.00	16.77	24.67	56.44	**7.22**	22.36
20	10	20	20.69	15.56	16.18	16.50	**5.08**	6.15
30	10	20	18.52	15.53	15.75	11.17	**5.05**	5.39
40	10	20	17.64	15.51	15.62	9.17	**5.01**	5.17
50	10	20	17.18	15.53	15.59	8.18	**5.03**	5.14
10	10	50	41.84	16.77	24.80	56.29	**7.21**	22.73
20	10	50	20.65	15.58	16.22	16.46	**5.12**	6.22
30	10	50	18.46	15.52	15.75	11.06	**5.02**	5.41
40	10	50	17.57	15.49	15.60	9.04	**4.98**	5.15
50	10	50	17.17	15.54	15.60	8.13	**5.05**	5.15
10	20	10	41.74	16.77	25.22	55.65	**7.22**	23.39
20	20	10	20.55	15.57	16.24	16.14	**5.10**	6.25
30	20	10	18.42	15.53	15.75	10.94	**5.03**	5.41
40	20	10	17.62	15.55	15.66	9.07	**5.08**	5.25
50	20	10	17.13	15.52	15.58	8.06	**5.02**	5.13
10	20	50	41.63	16.78	25.41	55.36	**7.25**	23.83
20	20	50	20.44	15.53	16.24	15.89	**5.05**	6.27
30	20	50	18.35	15.51	15.75	10.75	**5.00**	5.41
40	20	50	17.50	15.48	15.59	8.89	**4.95**	5.14
50	20	50	17.09	15.52	15.58	8.01	**5.01**	5.13
10	50	10	41.45	16.76	25.95	54.71	**7.19**	24.80
20	50	10	20.30	15.56	16.33	15.48	**5.09**	6.43
30	50	10	18.21	15.49	15.76	10.50	**4.97**	5.42
40	50	10	17.48	15.55	15.67	8.79	**5.06**	5.28
50	50	10	16.98	15.47	15.55	7.79	**4.94**	5.06
10	50	20	41.47	16.77	26.00	54.77	**7.21**	24.94
20	50	20	20.34	15.60	16.38	15.49	**5.15**	6.49
30	50	20	18.23	15.52	15.79	10.48	**5.02**	5.46
40	50	20	17.44	15.52	15.64	8.73	**5.01**	5.23
50	50	20	17.01	15.51	15.58	7.84	**5.00**	5.11

Note. $\chi^2_8(0.05) = 15.51$. The closest to $\alpha(= 0.05)$ in the values α_Q, α_{Q^*} and $\alpha_{Q^\dagger}$ of each row is boldface.

Table 4.4: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates when $(p_1, p_2, p_3) = (15, 12, 9)$ and $\alpha = 0.05$

n_1	n_2	n_3	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_Q	α_{Q^*}	$\alpha_{Q^\dagger}$
20	10	10	47.04	25.17	32.73	48.54	**5.23**	17.38
30	10	10	34.05	25.08	26.56	23.47	**5.11**	7.23
40	10	10	31.07	25.06	25.68	16.51	**5.09**	5.97
50	10	10	29.58	25.00	25.34	13.31	**5.01**	5.46
100	10	10	27.17	25.00	25.06	8.43	**5.00**	5.09
200	10	10	26.09	25.02	25.03	6.60	**5.03**	5.05
400	10	10	25.51	24.98	24.98	5.74	**4.98**	**4.98**
20	20	20	45.50	25.14	33.78	43.71	**5.20**	18.81
30	20	20	32.98	25.02	26.72	20.62	**5.04**	7.48
40	20	20	30.36	25.04	25.74	14.79	**5.05**	6.04
50	20	20	29.10	25.01	25.39	12.16	**5.02**	5.54
100	20	20	26.99	24.99	25.06	8.08	**4.99**	5.08
200	20	20	26.02	25.00	25.01	6.50	**5.00**	5.02
400	20	20	25.50	24.98	24.99	5.71	4.98	**4.99**
20	50	50	43.91	25.09	34.89	39.21	**5.13**	20.48
30	50	50	31.77	25.03	26.98	17.59	**5.04**	7.88
40	50	50	29.40	25.00	25.82	12.70	**5.00**	6.15
50	50	50	28.38	25.01	25.46	10.63	**5.02**	5.64
100	50	50	26.68	24.98	25.06	7.58	**4.98**	5.09
200	50	50	25.91	24.99	25.01	6.34	**4.99**	5.02
400	50	50	25.47	24.98	24.98	5.66	**4.98**	**4.98**
30	30	30	32.40	25.02	26.84	19.12	**5.04**	7.67
40	40	40	29.60	24.99	25.79	13.17	**4.99**	6.11
100	100	100	26.43	24.97	25.06	7.13	**4.97**	5.09
200	200	200	25.71	25.03	25.05	6.01	**5.04**	5.07
400	400	400	25.34	25.00	25.01	5.46	**5.01**	5.02

Note. $\chi^2_{15}(0.05) = 25.00$. The closest to $\alpha(= 0.05)$ in the values α_Q, α_{Q^*} and $\alpha_{Q^\dagger}$ of each row is boldface.

4.4 Conclusions

We have developed the MLRT statistic $Q^\dagger$ and the MT statistic Q^* with general monotone missing data in one-sample problem, where Q^* is not always monotone. Further, we presented that the LR for the one-sample test of the mean vector with monotone missing

Table 4.5: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates when $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$ and $\alpha = 0.05$

n_1	$n_2 = \cdots = n_5$	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_Q	α_{Q^*}	$\alpha_{Q^\dagger}$
20	10	46.86	25.08	33.06	47.84	**5.13**	17.89
30	10	33.86	24.99	26.63	23.03	**4.99**	7.40
40	10	30.91	24.98	25.69	16.13	**4.98**	5.97
50	10	29.52	24.99	25.39	13.05	**4.99**	5.53
100	10	27.15	25.02	25.08	8.37	**5.03**	5.11
200	10	26.03	24.98	24.99	6.53	4.98	**4.99**
400	10	25.53	25.00	25.01	5.74	**5.01**	**5.01**
20	20	45.27	25.11	33.97	43.11	**5.16**	19.09
30	20	32.82	25.01	26.79	20.24	**5.01**	7.60
40	20	30.21	24.99	25.76	14.49	**4.99**	6.07
50	20	29.01	25.00	25.43	11.96	**5.00**	5.58
100	20	26.96	24.99	25.07	8.03	**4.99**	5.10
200	20	25.98	24.97	24.99	6.44	4.97	**4.99**
400	20	25.53	25.01	25.01	5.75	**5.02**	5.03
20	50	43.83	25.07	35.03	38.92	**5.10**	20.66
30	50	31.68	25.03	27.03	17.32	**5.05**	7.94
40	50	29.34	25.01	25.87	12.51	**5.02**	6.23
50	50	28.31	25.01	25.48	10.47	**5.02**	5.67
100	50	26.65	25.00	25.08	7.50	**5.00**	5.11
200	50	25.83	24.94	24.95	6.20	4.92	**4.94**
400	50	25.49	25.01	25.02	5.67	**5.02**	5.03
30	30	32.24	25.00	26.88	18.84	**5.01**	7.75
40	40	29.52	25.00	25.84	12.93	**5.00**	6.18
100	100	26.41	25.01	25.10	7.15	**5.02**	5.14
200	200	25.65	24.98	25.01	5.91	4.97	**5.01**
400	400	25.32	25.00	25.01	5.44	**5.01**	5.02

Note. $\chi^2_{15}(0.05) = 25.00$. The closest to $\alpha(= 0.05)$ in the values α_Q, α_{Q^*} and $\alpha_{Q^\dagger}$ of each row is boldface.

data can be expressed as the products of the LR of the test for a mean vector and those of subvector, and derived the asymptotic expansion by the perturbation method. The null distribution of MLRT or MT statistic is considerably closer to the χ^2 distribution than that of the LRT statistic, even for small samples.

Chapter 5

Likelihood ratio test in a one-way MANOVA

In this chapter, testing the equality of mean vectors in a one-way MANOVA is considered when each dataset has a monotone pattern of missing observations. The LRT statistic in a one-way MANOVA with monotone missing data is given. Furthermore, the MT statistic based on LR and the MLRT statistic with monotone missing data are proposed using the decomposition of the LR and an asymptotic expansion for each decomposed LR. The accuracy of the approximation for the chi-square distribution is investigated using a Monte Carlo simulation. Finally, an example is given to illustrate the methods.

For the one-sample problem, as for the LRT, Krishnamoorthy and Pannala (1998) gave the decomposition of LR and provided comparisons with several approximation procedures. Then, Seko et al. (2012) discussed the LRT statistic and the linear interpolation approximation to the null distribution in the two-step monotone missing case. Recently, the MT and MLRT statistics of the one-sample test for a normal mean vector with monotone missing data are obtained by Yagi et al. (2017a). For two-sample problem, Seko et al. (2011) gave the LRT statistic with two-step monotone missing data, and the approximate upper percentiles for the null distribution. Recently, Seko (2012) gave the LRT statistic with two-step monotone missing data for the m-sample problem.

In this chapter, we consider the LRT, MT and MLRT statistics in a one-way MANOVA with monotone missing data. In the case of a one-way MANOVA with non-missing data, it is well known that Wilks' Λ statistic is the LRT statistic, and its MLRT statistic is given (see, e.g., Srivastava, 2002; Fujikoshi, Ulyanov and Shimizu, 2010). The main purpose of this chapter is to propose the LRT, MT and MLRT statistics with monotone missing data. We first give the LRT statistic for general monotone missing data. In order to establish the purpose, we decompose the LRT statistic and derive an asymptotic expansion of the characteristic function of each decomposed LRT statistic. In particular, we consider the decomposition of the LR as the products of independent LRs for a one-way MANOVA of the reduced dimension and those of the remaining subvectors with complete data. For the non-normal case with complete data, Gupta et al. (2006) derived the asymptotic expansion of the distribution of generalized U-statistic in the m-sample problem.

This chapter is summarized based on Yagi et al. (2017b) and organized as follows. In Section 5.1, we give the LR for a one-way MANOVA, using the MLEs of the covariance matrix. In Section 5.2, the MT and MLRT statistics are derived. In Section 5.3, simulation results are presented to investigate the accuracy of the approximation to the chi-square distribution for the null distribution of the test statistics. The methods are illustrated using an example in Section 5.4. Finally, some concluding remarks are given in Section 5.5.

5.1 LRT statistic

Suppose that $\boldsymbol{x}_{ij}^{(\ell)}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i^{(\ell)}$, $\ell = 1, 2, \ldots, m$ are independent and identically distributed as $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)}, \boldsymbol{\Sigma}_i)$, where $\boldsymbol{\mu}_i^{(\ell)} = (\boldsymbol{\mu}^{(\ell)})_i = (\mu_1^{(\ell)}, \mu_2^{(\ell)}, \ldots, \mu_{p_i}^{(\ell)})'$, and $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ principal submatrix of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}_1)$, with $\sum_{\ell=1}^m n_1^{(\ell)} - m > p(= p_1) > p_2 > \cdots > p_k$. Note that k denotes the number of steps. Then, the data from the samples $\boldsymbol{x}_{ij}^{(\ell)}$ are referred to as k-step monotone missing data. For this notation and assumption, see Section 3.2.1. With regard to the partitions of $\boldsymbol{\Sigma}$, for $1 \le i < j \le k$, let $(\boldsymbol{\Sigma}_i)_j$ be the principal submatrix of $\boldsymbol{\Sigma}_i$ of order $p_j \times p_j$; then, we define

$$\boldsymbol{\Sigma}_i = (\boldsymbol{\Sigma}_1)_i, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{i2} \\ \boldsymbol{\Sigma}_{i2}' & \boldsymbol{\Sigma}_{i3} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{i-1} = \begin{pmatrix} \boldsymbol{\Sigma}_i & \boldsymbol{\Sigma}_{(i-1,2)} \\ \boldsymbol{\Sigma}_{(i-1,2)}' & \boldsymbol{\Sigma}_{(i-1,3)} \end{pmatrix},$$

and $\boldsymbol{\Sigma}_{(i-1,3)\cdot i} = \boldsymbol{\Sigma}_{(i-1,3)} - \boldsymbol{\Sigma}_{(i-1,2)}' \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Sigma}_{(i-1,2)}$, $i = 2, 3, \ldots, k$. Then, the LR for

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

can be obtained as

$$\lambda = \prod_{i=1}^{k} \left(\frac{|\widehat{\boldsymbol{\Sigma}}_i|}{|\widetilde{\boldsymbol{\Sigma}}_i|} \right)^{\frac{\nu_i}{2}},$$

where $\nu_i = \sum_{\ell=1}^m n_i^{(\ell)}$, $i = 1, 2, \ldots, k$, $\widehat{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_1, and $\widetilde{\boldsymbol{\Sigma}}_i$ is the MLE of $\boldsymbol{\Sigma}_i$ under H_0. Note that $\widehat{\boldsymbol{\Sigma}}$ is obtained by Theorem 3.3. That is, setting

$$\overline{\boldsymbol{x}}_i^{(\ell)} = \frac{1}{n_i^{(\ell)}} \sum_{j=1}^{n_i^{(\ell)}} \boldsymbol{x}_{ij}^{(\ell)}, \quad \boldsymbol{E}_i^{(\ell)} = \sum_{j=1}^{n_i^{(\ell)}} (\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i^{(\ell)})(\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i^{(\ell)})', \quad i = 1, 2, \ldots, k,$$

$$\boldsymbol{d}_1^{(\ell)} = \overline{\boldsymbol{x}}_1^{(\ell)}, \quad \boldsymbol{d}_i^{(\ell)} = \frac{n_i^{(\ell)}}{N_{i+1}^{(\ell)}} \left[\overline{\boldsymbol{x}}_i^{(\ell)} - \frac{1}{N_i^{(\ell)}} \sum_{j=1}^{i-1} n_j^{(\ell)} (\overline{\boldsymbol{x}}_j^{(\ell)})_i \right], \quad i = 2, 3, \ldots, k,$$

$$N_1^{(\ell)} = 0, \; N_{i+1}^{(\ell)} = N_i^{(\ell)} + n_i^{(\ell)} \; \left(= \sum_{j=1}^{i} n_j^{(\ell)} \right), \quad i = 1, 2, \ldots, k,$$

$$M_i = \sum_{\ell=1}^{m} N_i^{(\ell)}, \quad i = 1, 2, \ldots, k+1,$$

the MLE of $\boldsymbol{\Sigma}$ under H_1 is given by

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{M_2} \sum_{\ell=1}^{m} \boldsymbol{H}_1^{(\ell)} + \sum_{\ell=1}^{m} \sum_{i=2}^{k} \frac{1}{M_{i+1}} \boldsymbol{F}_i^{[\mathrm{pl}]} \left[\boldsymbol{H}_i^{(\ell)} - \frac{\nu_i}{M_i} \boldsymbol{L}_{i-1,1}^{(\ell)} \right] \boldsymbol{F}_i^{[\mathrm{pl}]\prime},$$

where

$$\boldsymbol{H}_1^{(\ell)} = \boldsymbol{E}_1^{(\ell)},\ \boldsymbol{H}_i^{(\ell)} = \boldsymbol{E}_i^{(\ell)} + \frac{N_i^{(\ell)} N_{i+1}^{(\ell)}}{n_i^{(\ell)}} \boldsymbol{d}_i^{(\ell)} \boldsymbol{d}_i^{(\ell)\prime},\ i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_1^{(\ell)} = \boldsymbol{H}_1^{(\ell)},\ \boldsymbol{L}_i^{(\ell)} = (\boldsymbol{L}_{i-1}^{(\ell)})_i + \boldsymbol{H}_i^{(\ell)},\ i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_{i1}^{(\ell)} = (\boldsymbol{L}_i^{(\ell)})_{i+1},\ \boldsymbol{L}_i^{(\ell)} = \begin{pmatrix} \boldsymbol{L}_{i1}^{(\ell)} & \boldsymbol{L}_{i2}^{(\ell)} \\ \boldsymbol{L}_{i2}^{(\ell)\prime} & \boldsymbol{L}_{i3}^{(\ell)} \end{pmatrix},\ i = 1, 2, \ldots, k-1,$$
$$\boldsymbol{F}_1^{[\mathrm{pl}]} = \boldsymbol{G}_1,\quad \boldsymbol{F}_i^{[\mathrm{pl}]} = \boldsymbol{F}_{i-1}^{[\mathrm{pl}]} \boldsymbol{G}_i^{[\mathrm{pl}]},\quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{G}_1 = \boldsymbol{I}_{p_1},\quad \boldsymbol{G}_{i+1}^{[\mathrm{pl}]} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \left(\sum_{\ell=1}^{m} \boldsymbol{L}_{i2}^{(\ell)}\right)' \left(\sum_{\ell=1}^{m} \boldsymbol{L}_{i1}^{(\ell)}\right)^{-1} \end{pmatrix},\quad i = 1, 2, \ldots, k-1.$$

On the other hand, we define

$$\overline{\boldsymbol{x}}_i = \frac{1}{\nu_i} \sum_{\ell=1}^{m} \sum_{j=1}^{n_i^{(\ell)}} \boldsymbol{x}_{ij}^{(\ell)},\quad \boldsymbol{E}_i = \sum_{\ell=1}^{m} \sum_{j=1}^{n_i^{(\ell)}} (\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij}^{(\ell)} - \overline{\boldsymbol{x}}_i)',\quad i = 1, 2, \ldots, k,$$
$$\boldsymbol{d}_1 = \overline{\boldsymbol{x}}_1,\quad \boldsymbol{d}_i = \frac{\nu_i}{M_{i+1}} \Big[\overline{\boldsymbol{x}}_i - \frac{1}{M_i} \sum_{j=1}^{i-1} \nu_j (\overline{\boldsymbol{x}}_j)_i \Big],\quad i = 2, 3, \ldots, k.$$

Then, in the derivation analogous to the MLE under H_1, the MLE of $\boldsymbol{\Sigma}$ under H_0 can be obtained as

$$\widetilde{\boldsymbol{\Sigma}} = \frac{1}{M_2} \boldsymbol{H}_1 + \sum_{i=2}^{k} \frac{1}{M_{i+1}} \boldsymbol{F}_i \left[\boldsymbol{H}_i - \frac{\nu_i}{M_i} \boldsymbol{L}_{i-1,1} \right] \boldsymbol{F}_i',$$

where

$$\boldsymbol{H}_1 = \boldsymbol{E}_1,\quad \boldsymbol{H}_i = \boldsymbol{E}_i + \frac{M_i M_{i+1}}{\nu_i} \boldsymbol{d}_i \boldsymbol{d}_i',\quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_1 = \boldsymbol{H}_1,\ \boldsymbol{L}_i = (\boldsymbol{L}_{i-1})_i + \boldsymbol{H}_i,\ i = 2, 3, \ldots, k,$$
$$\boldsymbol{L}_{i1} = (\boldsymbol{L}_i)_{i+1},\ \boldsymbol{L}_i = \begin{pmatrix} \boldsymbol{L}_{i1} & \boldsymbol{L}_{i2} \\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3} \end{pmatrix},\ i = 1, 2, \ldots, k-1,$$
$$\boldsymbol{F}_1 = \boldsymbol{G}_1,\ \boldsymbol{F}_i = \boldsymbol{F}_{i-1} \boldsymbol{G}_i,\ i = 2, 3, \ldots, k,$$
$$\boldsymbol{G}_1 = \boldsymbol{I}_{p_1},\ \boldsymbol{G}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \boldsymbol{L}_{i2}' \boldsymbol{L}_{i1}^{-1} \end{pmatrix},\ i = 1, 2, \ldots, k-1.$$

In addition, the MLE of $\boldsymbol{\mu}^{(\ell)}$, $\widehat{\boldsymbol{\mu}}^{(\ell)}$ $(\ell = 1, 2, \ldots m)$ is also given in Theorem 3.3, and the MLE of $\boldsymbol{\mu}$ under H_0 can be obtained as

$$\widetilde{\boldsymbol{\mu}} = \sum_{i=1}^{k} \widetilde{\boldsymbol{f}}_i,$$

where

$$\widetilde{\boldsymbol{f}}_1 = \boldsymbol{d}_1,\ \widetilde{\boldsymbol{f}}_i = \widetilde{\boldsymbol{U}}_i \boldsymbol{d}_i,\ i = 2, 3, \ldots, k,$$
$$\widetilde{\boldsymbol{U}}_1 = \boldsymbol{T}_1,\quad \widetilde{\boldsymbol{U}}_i = \widetilde{\boldsymbol{U}}_{i-1}\widetilde{\boldsymbol{T}}_i,\quad i = 2, 3, \ldots, k,$$
$$\boldsymbol{T}_1 = \boldsymbol{I}_{p_1},\quad \widetilde{\boldsymbol{T}}_{i+1} = \begin{pmatrix} \boldsymbol{I}_{p_{i+1}} \\ \widetilde{\boldsymbol{\Sigma}}'_{(i,2)}\widetilde{\boldsymbol{\Sigma}}^{-1}_{i+1} \end{pmatrix},\quad i = 1, 2, \ldots, k-1.$$

Note that the above results of $\widetilde{\boldsymbol{\Sigma}}$ and $\widetilde{\boldsymbol{\mu}}$ essentially coincide with the one-sample case of Jinadasa and Tracy (1992).

Using the above results, we can obtain the LRT statistic $Q(= -2\log\lambda)$, the null distribution of which is asymptotically distributed as a χ^2 distribution with $(m-1)p$ degrees of freedom when the sample size is large.

5.2 MT and MLRT statistics and their null distributions

In this section, we derive the MT statistic and the MLRT statistic using a decomposition of the LR and an asymptotic expansion procedure. Let

$$\lambda_1 = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_k|}{|\widetilde{\boldsymbol{\Sigma}}_k|}\right)^{\frac{M_{k+1}}{2}},\quad \lambda_i = \left(\frac{|\widehat{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}|}{|\widetilde{\boldsymbol{\Sigma}}_{(k-i+1,3)\cdot k-i+2}|}\right)^{\frac{M_{k-i+2}}{2}},\quad i = 2, 3, \ldots, k.$$

Then, λ in Section 5.1 can be expressed as

$$\lambda = \prod_{i=1}^{k} \lambda_i,$$

where λ_i, $i = 1, 2, \ldots, k$ are mutually independent. Furthermore, let $Q_i^* = -2\rho_i \log\lambda_i$, where

$$\rho_i = 1 - \frac{1}{2M_{k-i+2}}(p_{k-i+1} + p_{k-i+2} + m + 2),\quad i = 1, 2, \ldots, k,\quad p_{k+1} = 0.$$

Then, we obtain

$$\Pr(Q_i^* \le x) = G_{(m-1)(p_{k-i+1} - p_{k-i+2})}(x) + O(M_{k-i+2}^{-2}),\quad i = 1, 2, \ldots, k,$$

where $G_f(x)$ is the distribution function of a χ^2-variate with f degrees of freedom.

The derivation of the above result of the Bartlett correction factor ρ_i $(i = 1, 2, \ldots, k)$ is as follows. For simplicity, we consider the case of $k = 2$. That is, we derive ρ_1 and ρ_2. Here, we consider the following hypotheses:

$$H_{01} : \boldsymbol{\mu}_2^{(1)} = \boldsymbol{\mu}_2^{(2)} = \cdots = \boldsymbol{\mu}_2^{(m)} \text{ vs. } H_{11} : \text{ not } H_{01},$$

and

$$H_{02}: \boldsymbol{A}\boldsymbol{\mu}_1^{(1)} = \boldsymbol{A}\boldsymbol{\mu}_1^{(2)} = \cdots = \boldsymbol{A}\boldsymbol{\mu}_1^{(m)} \text{ given } \boldsymbol{\mu}_2^{(1)} = \boldsymbol{\mu}_2^{(2)} = \cdots = \boldsymbol{\mu}_2^{(m)}$$
$$\text{vs. } H_{12}: \boldsymbol{A}\boldsymbol{\mu}_1^{(i)} \neq \boldsymbol{A}\boldsymbol{\mu}_1^{(j)} \text{ for some } i \neq j \text{ given } \boldsymbol{\mu}_2^{(1)} = \boldsymbol{\mu}_2^{(2)} = \cdots = \boldsymbol{\mu}_2^{(m)},$$

where

$$\boldsymbol{\mu}_1^{(\ell)} = \begin{pmatrix} \boldsymbol{\mu}_2^{(\ell)} \\ \boldsymbol{A}\boldsymbol{\mu}_1^{(\ell)} \end{pmatrix} \begin{matrix} \}p_2 \\ \}p_1 - p_2 \end{matrix},$$

and $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_{p_1 - p_2} \end{pmatrix}$ is a $(p_1 - p_2) \times p_1$ matrix. Then, λ_1 and λ_2 are equal to the LRs for H_{01} and H_{02}, respectively (see Figure 5.1). That is, $Q_1 (= -2 \log \lambda_1)$ and $Q_2 (= -2 \log \lambda_2)$ are the LRT statistics of the tests for a p_2-dimensional mean vector and a $(p_1 - p_2)$-dimensional subvector, respectively. The one-sample case is discussed by Krishnamoorthy and Pannala (1998) and Yagi et al. (2017a) (see Chapter 4).

We first consider the null distribution of Q_1. Then, Q_1 is the LRT statistic for the test of the mean vector in a one-way MANOVA, where the data consist of complete data sets ($N_3^{(\ell)} \times p_2$, $\ell = 1, 2, \ldots, m$). The Wilks' Λ statistic is given by

$$\Lambda = \frac{|\boldsymbol{S}_w|}{|\boldsymbol{S}_w + \boldsymbol{S}_b|},$$

where $\boldsymbol{S}_b$ and $\boldsymbol{S}_w$ are matrices of the sums of squares and the products (SSP matrices) from treatments (between groups) and errors (within groups), respectively (see Fujikoshi et al., 2010). Therefore, the MLRT statistic Q_1^* is given by

$$Q_1^* = -2\rho_1 \log \Lambda^{\frac{M_3}{2}},$$

where

$$\rho_1 = 1 - \frac{1}{2M_3}(p_2 + m + 2).$$

Furthermore, we have

$$\Pr(Q_1^* \leq x) = G_{(m-1)p_2}(x) + O(M_3^{-2}).$$

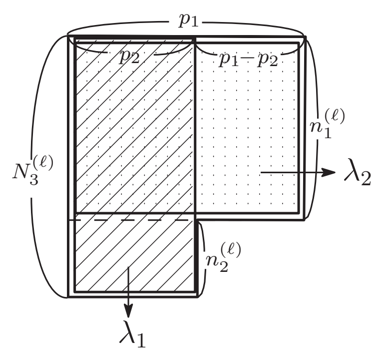

Figure 5.1: The LRs λ_1 and λ_2 in the case of two-step monotone missing data

Second, we derive an asymptotic expansion of the null distribution of Q_2. For convenience, let $\boldsymbol{y}_1^{(\ell)},\ \boldsymbol{y}_2^{(\ell)},\dots,\boldsymbol{y}_{n_1^{(\ell)}}^{(\ell)} \overset{i.i.d.}{\sim} N_p(\boldsymbol{\eta}^{(\ell)},\boldsymbol{\Delta}),\ \ \ell=1,2,\dots,m$, where

$$\boldsymbol{\eta}^{(\ell)}=\begin{pmatrix}\boldsymbol{\eta}_1^{(\ell)}\\ \boldsymbol{\eta}_2^{(\ell)}\end{pmatrix}\begin{matrix}\}r\\ \}s\end{matrix},\quad \boldsymbol{\Delta}=\left(\begin{array}{c|c}\overbrace{\boldsymbol{\Delta}_{11}}^{r} & \overbrace{\boldsymbol{\Delta}_{12}}^{s}\\ \hline \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22}\end{array}\right)\begin{matrix}\}r\\ \}s\end{matrix},$$

with $r=p_2,\ s=p_1-p_2$. Then, because λ_2 is equal to the LR when testing

$$\begin{aligned}&H_{02}:\boldsymbol{\eta}_2^{(1)}=\boldsymbol{\eta}_2^{(2)}=\cdots=\boldsymbol{\eta}_2^{(m)}\ \text{given}\ \boldsymbol{\eta}_1^{(1)}=\boldsymbol{\eta}_1^{(2)}=\cdots=\boldsymbol{\eta}_1^{(m)}\\ &\text{vs. } H_{12}:\boldsymbol{\eta}_2^{(i)}\neq\boldsymbol{\eta}_2^{(j)}\ \text{for some}\ i\neq j\ \text{given}\ \boldsymbol{\eta}_1^{(1)}=\boldsymbol{\eta}_1^{(2)}=\cdots=\boldsymbol{\eta}_1^{(m)},\end{aligned}$$

we can write

$$\lambda_2=\left(\frac{|\boldsymbol{I}_p+\boldsymbol{W}^{-1}\boldsymbol{B}|}{|\boldsymbol{I}_r+\boldsymbol{W}_{11}^{-1}\boldsymbol{B}_{11}|}\right)^{-\frac{M_2}{2}},$$

where

$$\begin{aligned}&\boldsymbol{W}=\sum_{\ell=1}^{m}\boldsymbol{W}^{(\ell)},\quad \boldsymbol{W}^{(\ell)}=\sum_{j=1}^{n_1^{(\ell)}}(\boldsymbol{y}_j^{(\ell)}-\overline{\boldsymbol{y}}^{(\ell)})(\boldsymbol{y}_j^{(\ell)}-\overline{\boldsymbol{y}}^{(\ell)})',\quad \overline{\boldsymbol{y}}^{(\ell)}=\frac{1}{n_1^{(\ell)}}\sum_{j=1}^{n_1^{(\ell)}}\boldsymbol{y}_j^{(\ell)},\\ &\boldsymbol{B}=\sum_{\ell=1}^{m}n_1^{(\ell)}(\overline{\boldsymbol{y}}^{(\ell)}-\overline{\overline{\boldsymbol{y}}})(\overline{\boldsymbol{y}}^{(\ell)}-\overline{\overline{\boldsymbol{y}}})',\quad \overline{\overline{\boldsymbol{y}}}=\frac{1}{M_2}\sum_{\ell=1}^{m}n_1^{(\ell)}\overline{\boldsymbol{y}}^{(\ell)},\end{aligned}$$

and

$$\overline{\boldsymbol{y}}^{(\ell)}=\begin{pmatrix}\overline{\boldsymbol{y}}_1^{(\ell)}\\ \overline{\boldsymbol{y}}_2^{(\ell)}\end{pmatrix}\begin{matrix}\}r\\ \}s\end{matrix},\quad \boldsymbol{W}=\left(\begin{array}{c|c}\overbrace{\boldsymbol{W}_{11}}^{r} & \overbrace{\boldsymbol{W}_{12}}^{s}\\ \hline \boldsymbol{W}_{21} & \boldsymbol{W}_{22}\end{array}\right)\begin{matrix}\}r\\ \}s\end{matrix},\quad \boldsymbol{B}=\left(\begin{array}{c|c}\overbrace{\boldsymbol{B}_{11}}^{r} & \overbrace{\boldsymbol{B}_{12}}^{s}\\ \hline \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{array}\right)\begin{matrix}\}r\\ \}s\end{matrix}.$$

Without loss of generality, we may assume that $\boldsymbol{\Delta}=\boldsymbol{I}$. Therefore, under H_{02}, let $\boldsymbol{\eta}^{(1)}=\boldsymbol{\eta}^{(2)}=\cdots=\boldsymbol{\eta}^{(m)}=\boldsymbol{\eta}$, and let

$$\overline{\boldsymbol{y}}^{(\ell)}=\boldsymbol{\eta}+\frac{1}{\sqrt{n_1^{(\ell)}}}\boldsymbol{z}^{(\ell)},\quad \frac{1}{n_1^{(\ell)}-1}\boldsymbol{W}^{(\ell)}=\boldsymbol{I}+\frac{1}{\sqrt{n_1^{(\ell)}}}\boldsymbol{V}^{(\ell)},\quad \ell=1,2,\dots,m.$$

Furthermore, using the partitions of $\boldsymbol{z}^{(\ell)}$ and $\boldsymbol{V}^{(\ell)}$ as

$$\boldsymbol{z}^{(\ell)}=\begin{pmatrix}\boldsymbol{z}_1^{(\ell)}\\ \boldsymbol{z}_2^{(\ell)}\end{pmatrix}\begin{matrix}\}r\\ \}s\end{matrix},\quad \boldsymbol{V}^{(\ell)}=\left(\begin{array}{c|c}\overbrace{\boldsymbol{V}_{11}^{(\ell)}}^{r} & \overbrace{\boldsymbol{V}_{12}^{(\ell)}}^{s}\\ \hline \boldsymbol{V}_{21}^{(\ell)} & \boldsymbol{V}_{22}^{(\ell)}\end{array}\right)\begin{matrix}\}r\\ \}s\end{matrix},\quad \ell=1,2,\dots,m,$$

we can expand Q_2 as

$$Q_2=\mathrm{tr}\boldsymbol{B}_{22}+\frac{1}{\sqrt{n_1^{(1)}}}C_1+\frac{1}{n_1^{(1)}}C_2+O_p((n_1^{(1)})^{-\frac{3}{2}}),$$

where

$$C_1 = -\frac{1}{q}\mathrm{tr}\Big(\sum_{\ell=1}^{m}\sqrt{q_\ell}\boldsymbol{V}^{(\ell)}\Big)\boldsymbol{B} + \frac{1}{q}\mathrm{tr}\Big(\sum_{\ell=1}^{m}\sqrt{q_\ell}\boldsymbol{V}_{11}^{(\ell)}\Big)\boldsymbol{B}_{11},$$

$$C_2 = \frac{m}{q}\mathrm{tr}\boldsymbol{B}_{22} + \frac{1}{q^2}\Big\{\mathrm{tr}\Big(\sum_{\ell=1}^{m}\sqrt{q_\ell}\boldsymbol{V}^{(\ell)}\Big)^2\boldsymbol{B} - \mathrm{tr}\Big(\sum_{\ell=1}^{m}\sqrt{q_\ell}\boldsymbol{V}_{11}^{(\ell)}\Big)^2\boldsymbol{B}_{11}\Big\} - \frac{1}{2q}(\mathrm{tr}\boldsymbol{B}^2 - \mathrm{tr}\boldsymbol{B}_{11}^2),$$

$$\boldsymbol{B} = \sum_{\ell=1}^{m}\Big(1-\frac{q_\ell}{q}\Big)\boldsymbol{z}^{(\ell)}\boldsymbol{z}^{(\ell)\prime} - \frac{1}{q}\sum_{\substack{i=1\\ i\neq j}}^{m}\sum_{j=1}^{m}\sqrt{q_iq_j}\boldsymbol{z}^{(i)}\boldsymbol{z}^{(j)\prime},$$

$$q = \sum_{\ell=1}^{m} q_\ell, \quad q_\ell = \frac{n_1^{(\ell)}}{n_1^{(1)}}, \quad \ell = 1,2,\ldots,m.$$

Hence

$$E[\exp(itQ_2)] = E[\exp(it\,\mathrm{tr}\boldsymbol{B}_{22})] + E\Big[\Big(D + \frac{1}{2}D^2\Big)\exp(it\,\mathrm{tr}\boldsymbol{B}_{22})\Big] + O((n_1^{(1)})^{-\frac{3}{2}}),$$

where

$$D = it\Bigg(\frac{1}{\sqrt{n_1^{(1)}}}C_1 + \frac{1}{n_1^{(1)}}C_2\Bigg).$$

After calculating the expectation and inverting the characteristic function, we obtain the following result:

$$\Pr(Q_2 \le x) = G_{(m-1)s}(x) + \frac{\beta_2}{M_2}\left[G_{(m-1)s}(x) - G_{(m-1)s+2}(x)\right] + O(M_2^{-2}),$$

where $\beta_2 = -(m-1)s(2r+s+m+2)/4$. In addition, letting

$$\rho_2 = 1 - \frac{1}{2M_2}(p_1 + p_2 + m + 2),$$

the distribution function of $Q_2^*(= -2\rho_2 \log \lambda_2)$ is given by

$$\Pr(Q_2^* \le x) = G_{(m-1)(p_1-p_2)}(x) + O(M_2^{-2}).$$

Therefore, we can give the MT statistic

$$Q^* = \sum_{i=1}^{k} Q_i^*$$

such that $\Pr(Q^* \le x) = G_{(m-1)p}(x) + O(M_2^{-2})$. Note that Q^* converges to the χ^2 distribution much faster than the LRT statistic does. On the other hand, since an asymptotic expansion of the characteristic function of Q_1 is given by

$$E[\exp(itQ_1)] = (1-2it)^{-\frac{1}{2}(m-1)r} + \frac{\beta_1}{M_3}(1-2it)^{-\frac{1}{2}(m-1)r}\left[1-(1-2it)^{-1}\right] + O(M_3^{-2}),$$

where $\beta_1 = -(m-1)r(r+m+2)/4$, gathering up the expanded results for the characteristic function of Q_i, we can propose

$$Q^\dagger = -2\rho \log \lambda$$

as an MLRT statistic, where

$$\rho = 1 - \frac{1}{2\nu_1 p} \sum_{i=1}^{k} \frac{1}{w_{k-i+1}} (p_{k-i+1} - p_{k-i+2})(p_{k-i+1} + p_{k-i+2} + m + 2),$$

$w_{k-i+1} = \sum_{j=1}^{k-i+1} u_j$, $u_j = \nu_j/\nu_1$, and $p_{k+1} = 0$. Note that the null distribution of $Q^\dagger$ can be expressed as

$$\Pr(Q^\dagger \le x) = G_{(m-1)p}(x) + O(\nu_1^{-2}).$$

This means that the error from using the $\chi^2_{(m-1)p}$ distribution is of order ν_1^{-2}. We note that $Q^\dagger$ is a monotone transformation, but Q^* is not always monotone.

5.3 Simulation studies

In this section, we investigate the numerical accuracy and the asymptotic behavior of the upper percentiles of the test statistics, Q, Q^*, and $Q^\dagger$, using the actual type I error rates. We compute the upper percentiles of the null distribution of the test statistics in a one-way MANOVA using a Monte Carlo simulation (10^6 runs). That is, the LRT statistic Q, the MT statistic Q^*, and the MLRT statistic $Q^\dagger$ are computed 10^6 times, based on the normal random vectors generated from $N_{p_i}(\mathbf{0}, \boldsymbol{I}_{p_i})$ $i = 1, 2, \ldots, k$. In Table 5.1, we provide the simulated upper 100α percentiles of Q, Q^* and $Q^\dagger$, and their actual type I error rates for the two-step case. Note that the actual type I error rates are defined as

$$\frac{\alpha_q}{100} = \Pr\{Q > \chi^2_{(m-1)p}(\alpha)\}, \quad \frac{\alpha_{q^*}}{100} = \Pr\{Q^* > \chi^2_{(m-1)p}(\alpha)\},$$

and

$$\frac{\alpha_{q^\dagger}}{100} = \Pr\{Q^\dagger > \chi^2_{(m-1)p}(\alpha)\},$$

where $\chi^2_{(m-1)p}(\alpha)$ is the upper 100α percentile of the χ^2 distribution with $(m-1)p$ degrees of freedom. Computations are carried out for the following parameter sets for $m = 3$, where $n_i^{(1)} = n_i^{(2)} = n_i^{(3)} = n_i$, $i = 1, 2, \ldots, k$:

$$\text{(I) } k = 2;\ (p_1, p_2) = (15, 3), (15, 8), (15, 12);$$
$$n_1 = 20, 50, 100, 200;\ n_2 = 5, 10, 50;\ \alpha = 0.05.$$

We note from Table 5.1 that the value of $q(\alpha)$ converges very slowly to that of $\chi^2_{2p}(\alpha)$. However, the values of $q^*(\alpha)$ and $q^\dagger(\alpha)$ are close to that of $\chi^2_{2p}(\alpha)$ even when the sample size n_1 is not large. In addition, from Table 5.1, the value of $q^*(\alpha)$ may be closer to $\chi^2_{2p}(\alpha)$ when the value of p_2 is large. For example, comparing the cases of $(p_1, p_2) = (15, 3)$ and $(p_1, p_2) = (15, 12)$, the simulated values of $(p_1, p_2) = (15, 12)$ converge to the χ^2 distribution much faster than those of $(p_1, p_2) = (15, 3)$ do. Tables 5.2 and 5.3 give the

Table 5.1: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates (two-step case)

n_1	n_2	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_q	α_{q^*}	$\alpha_{q^\dagger}$
			$(p_1, p_2) = (15, 3)$				
20	5	52.88	43.99	44.21	19.88	5.23	5.46
50	5	46.95	43.81	43.84	9.00	5.04	5.07
100	5	45.30	43.79	43.80	6.78	5.02	5.03
200	5	44.51	43.77	43.77	5.80	5.00	5.00
20	10	52.78	43.97	44.22	19.78	5.20	5.46
50	10	46.92	43.81	43.83	8.96	5.04	5.07
100	10	45.29	43.78	43.79	6.71	5.00	5.02
200	10	44.51	43.77	43.77	5.81	5.00	5.00
20	50	52.61	44.03	44.34	19.30	5.27	5.60
50	50	46.78	43.75	43.78	8.80	4.97	5.01
100	50	45.27	43.80	43.80	6.69	5.03	5.03
200	50	44.50	43.77	43.77	5.78	4.99	5.00
			$(p_1, p_2) = (15, 8)$				
20	5	52.38	43.87	44.26	18.91	5.10	5.53
50	5	46.88	43.81	43.85	8.92	5.04	5.08
100	5	45.26	43.77	43.78	6.72	4.99	5.00
200	5	44.53	43.80	43.80	5.81	5.02	5.02
20	10	51.99	43.86	44.32	17.98	5.08	5.59
50	10	46.74	43.75	43.81	8.73	4.98	5.04
100	10	45.25	43.79	43.79	6.70	5.01	5.01
200	10	44.50	43.78	43.77	5.81	5.00	5.00
20	50	50.94	43.82	44.55	15.79	5.05	5.84
50	50	46.39	43.77	43.84	8.16	5.00	5.07
100	50	45.10	43.76	43.77	6.49	4.98	5.00
200	50	44.45	43.76	43.76	5.75	4.98	4.99
			$(p_1, p_2) = (15, 12)$				
20	5	51.63	43.90	44.19	17.29	5.13	5.44
50	5	46.73	43.78	43.81	8.70	5.00	5.04
100	5	45.25	43.79	43.79	6.69	5.02	5.03
200	5	44.51	43.78	43.78	5.83	5.01	5.01
20	10	50.79	43.87	44.25	15.55	5.09	5.50
50	10	46.61	43.81	43.85	8.49	5.04	5.08
100	10	45.22	43.79	43.80	6.65	5.02	5.03
200	10	44.48	43.77	43.77	5.76	5.00	4.99
20	50	48.54	43.78	44.38	11.48	5.00	5.65
50	50	45.84	43.77	43.82	7.47	4.99	5.05
100	50	44.93	43.76	43.77	6.28	4.99	4.99
200	50	44.39	43.76	43.75	5.68	4.98	4.98

Note. $\chi^2_{30}(0.05) = 43.77$.

Table 5.2: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates (three-step case)

n_1	$n_2 = n_3$	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_q	α_{q^*}	$\alpha_{q^\dagger}$
			$(p_1, p_2, p_3) = (15, 10, 5)$				
20	5	51.83	43.77	44.25	17.75	5.00	5.51
50	5	46.80	43.81	43.86	8.74	5.04	5.09
100	5	45.27	43.79	43.81	6.71	5.02	5.03
200	5	44.48	43.75	43.75	5.78	4.98	4.98
20	10	51.26	43.82	44.37	16.46	5.04	5.65
50	10	46.62	43.78	43.84	8.51	5.01	5.07
100	10	45.20	43.77	43.78	6.62	5.00	5.01
200	10	44.51	43.78	43.79	5.80	5.01	5.01
20	50	49.78	43.83	44.61	13.61	5.06	5.91
50	50	46.05	43.75	43.83	7.73	4.98	5.06
100	50	45.00	43.77	43.79	6.37	5.00	5.02
200	50	44.42	43.77	43.77	5.70	5.00	5.00
			$(p_1, p_2, p_3) = (15, 12, 9)$				
20	5	51.12	43.80	44.23	16.24	5.03	5.50
50	5	46.63	43.77	43.81	8.52	4.99	5.04
100	5	45.21	43.77	43.78	6.65	5.00	5.01
200	5	44.49	43.77	43.77	5.79	4.99	5.00
20	10	50.19	43.82	44.31	14.41	5.05	5.57
50	10	46.41	43.77	43.82	8.22	5.00	5.05
100	10	45.14	43.77	43.77	6.54	5.00	5.00
200	10	44.48	43.77	43.77	5.76	5.00	5.00
20	50	48.19	43.77	44.46	10.87	5.00	5.74
50	50	45.67	43.80	43.87	7.21	5.02	5.10
100	50	44.88	43.81	43.83	6.22	5.04	5.05
200	50	44.39	43.79	43.79	5.66	5.02	5.02

Note. $\chi^2_{30}(0.05) = 43.77$.

simulated results for the three-step case (II) and the five-step case (III), respectively:

(II) $k = 3$; $(p_1, p_2, p_3) = (15, 10, 5), (15, 12, 9)$;
$n_1 = 20, 50, 100, 200$; $n_2 = n_3 = 5, 10, 50$; $\alpha = 0.05$.

(III) $k = 5$; $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$;
$n_1 = 20, 50, 100, 200$, $n_2 = n_3 = \cdots = n_5 = 5, 10, 50$; $\alpha = 0.05$.

From Tables 5.2 and 5.3, the asymptotic behavior of the approximations of the χ^2 distribution in the case of three-step or five-step monotone missing data show the same tendencies as in the two-step case.

Table 5.3: The upper percentiles of Q, Q^*, $Q^\dagger$ and the actual type I error rates when $(p_1, p_2, p_3, p_4, p_5) = (15, 12, 9, 6, 3)$

n_1	$n_2 = \cdots = n_5$	$q(\alpha)$	$q^*(\alpha)$	$q^\dagger(\alpha)$	α_q	α_{q^*}	$\alpha_{q^\dagger}$
20	5	50.91	43.76	44.28	15.87	4.99	5.55
50	5	46.60	43.78	43.84	8.50	5.01	5.07
100	5	45.22	43.79	43.80	6.61	5.02	5.03
200	5	44.50	43.78	43.78	5.77	5.01	5.01
20	10	49.93	43.76	44.30	13.94	4.99	5.57
50	10	46.31	43.74	43.81	8.11	4.97	5.04
100	10	45.12	43.77	43.79	6.52	5.00	5.01
200	10	44.47	43.77	43.78	5.77	5.00	5.00
20	50	48.10	43.79	44.49	10.70	5.02	5.77
50	50	45.61	43.81	43.89	7.13	5.04	5.12
100	50	44.78	43.76	43.77	6.12	4.99	5.00
200	50	44.36	43.78	43.78	5.63	5.00	5.01

Note. $\chi^2_{30}(0.05) = 43.77$.

5.4 A numerical illustration

In this section, we illustrate the results of this paper using an example given in well known "Fisher's Iris Data" which presents measurements of the sepal length and width, and pedal length and width in centimeters of 50 plants for Iris virginica. For illustration purpose and reproduction of the Iris Data with the equality of covariance matrices, three datasets with three-step monotone missing pattern ($(p_1, p_2, p_3) = (4, 3, 2)$, $n_1^{(\ell)} = 10, n_2^{(\ell)} = n_3^{(\ell)} = 5$, $\ell = 1, 2, 3$) were generated from the dataset of Iris virginica by Bootstrap method, where each value of the sepal length in the first group is added to the value of 0.5. These data are presented in Table 5.4, where the entries in parentheses were discarded to make a monotone pattern. According to the standard procedure of the MLRT (e.g. Srivastava 2002, p.490) to check the equality of covariance matrices, we obtain the value of the test statistic as 23.17 with p value 0.28, and so the assumption of equality of covariance matrices is actually tenable. Then, we want to test $H_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} = \boldsymbol{\mu}^{(3)}$ vs. H_1 : not H_0. Under the three-step monotone missing data in Table 5.4, we computed the MLEs of $\boldsymbol{\mu}^{(\ell)}$ and $\boldsymbol{\Sigma}$ as

$$\widehat{\boldsymbol{\mu}}^{(1)} = \begin{pmatrix} 7.02 \\ 2.95 \\ 5.50 \\ 2.03 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(2)} = \begin{pmatrix} 6.40 \\ 3.06 \\ 5.43 \\ 2.07 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(3)} = \begin{pmatrix} 6.76 \\ 2.98 \\ 5.69 \\ 1.98 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.2997 & 0.0687 & 0.2273 & 0.0691 \\ 0.0687 & 0.0870 & 0.0463 & 0.0427 \\ 0.2273 & 0.0463 & 0.2819 & 0.0529 \\ 0.0691 & 0.0427 & 0.0529 & 0.0752 \end{pmatrix},$$

Table 5.4: Artificial bootstrap dataset from Fisher's Iris data

Group 1				Group 2				Group 3			
Sepal		Petal		Sepal		Petal		Sepal		Petal	
Length	Width	Length	Width	Length	Width	Length	Width	Length	Width	Length	Width
7.8	2.9	6.3	1.8	6.8	3.0	5.5	2.1	6.3	2.8	5.1	1.5
7.4	3.1	5.1	2.3	6.3	2.8	5.1	1.5	6.3	2.5	5.0	1.9
8.2	2.6	6.9	2.3	6.9	3.1	5.1	2.3	6.9	3.1	5.1	2.3
7.2	3.1	5.6	2.4	6.2	3.4	5.4	2.3	6.3	2.9	5.6	1.8
5.4	2.5	4.5	1.7	6.4	2.7	5.3	1.9	7.7	2.6	6.9	2.3
6.3	2.7	5.1	1.9	6.7	3.3	5.7	2.1	6.9	3.2	5.7	2.3
7.8	2.9	6.3	1.8	6.3	3.4	5.6	2.4	6.8	3.0	5.5	2.1
6.7	2.8	4.8	1.8	6.4	3.1	5.5	1.8	7.7	3.8	6.7	2.2
6.7	3.4	5.4	2.3	6.2	3.4	5.4	2.3	6.3	2.8	5.1	1.5
6.6	3.0	4.9	1.8	5.7	2.5	5.0	2.0	7.7	2.8	6.7	2.0
7.3	3.2	5.9	(2.3)	6.3	3.3	6.0	(2.5)	6.1	3.0	4.9	(1.8)
7.7	3.2	6.0	(1.8)	7.2	3.2	6.0	(1.8)	6.3	2.9	5.6	(1.8)
6.3	2.7	5.1	(1.9)	6.5	3.0	5.5	(1.8)	6.5	3.0	5.8	(2.2)
6.6	2.6	5.6	(1.4)	6.1	3.0	4.9	(1.8)	7.9	3.8	6.4	(2.0)
7.4	3.1	5.1	(2.3)	6.3	3.4	5.6	(2.4)	6.1	2.6	5.6	(1.4)
6.8	3.4	(5.6)	(2.4)	6.5	3.0	(5.8)	(2.2)	7.1	3.0	(5.9)	(2.1)
6.4	3.0	(5.1)	(1.8)	6.2	2.8	(4.8)	(1.8)	6.4	3.2	(5.3)	(2.3)
7.6	3.0	(5.9)	(2.1)	5.7	2.5	(5.0)	(2.0)	6.4	2.8	(5.6)	(2.1)
7.4	3.1	(5.1)	(2.3)	6.8	3.2	(5.9)	(2.3)	6.3	2.5	(5.0)	(1.9)
6.8	2.7	(4.9)	(1.8)	6.5	3.0	(5.2)	(2.0)	7.2	3.2	(6.0)	(1.8)

Note. The entries in parentheses were discarded to make a monotone pattern.

respectively. Further, under the null hypothesis H_0, we computed

$$\widetilde{\boldsymbol{\mu}} = \begin{pmatrix} 6.73 \\ 2.99 \\ 5.54 \\ 2.03 \end{pmatrix}, \quad \widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.3643 & 0.0575 & 0.2446 & 0.0650 \\ 0.0575 & 0.0890 & 0.0474 & 0.0436 \\ 0.2446 & 0.0474 & 0.3042 & 0.0524 \\ 0.0650 & 0.0436 & 0.0524 & 0.0764 \end{pmatrix},$$

respectively. These MLEs are obtained by using the results of Section 5.1 when $k = 3$ and $m = 3$. Therefore, the observed values of the LRT, MT and MLRT statistics are given by $Q = 30.56$ with p-value 1.68×10^{-4}, $Q^* = 28.06$ with p value 4.63×10^{-4} and $Q^\dagger = 27.29$ with p value 6.30×10^{-4}, respectively. Using only the first 10 complete observations of each group (partially complete), we computed

$$\widehat{\boldsymbol{\mu}}^{(1)}_{\rm pc} = \begin{pmatrix} 7.01 \\ 2.90 \\ 5.49 \\ 2.01 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(2)}_{\rm pc} = \begin{pmatrix} 6.39 \\ 3.07 \\ 5.36 \\ 2.07 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(3)}_{\rm pc} = \begin{pmatrix} 6.89 \\ 2.95 \\ 5.74 \\ 1.99 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}_{\rm pc} = \begin{pmatrix} 0.3589 & 0.0484 & 0.3077 & 0.0708 \\ 0.0484 & 0.0922 & 0.0369 & 0.0419 \\ 0.3077 & 0.0369 & 0.3652 & 0.0614 \\ 0.0708 & 0.0419 & 0.0614 & 0.0753 \end{pmatrix},$$

and

$$\widetilde{\boldsymbol{\mu}}_{\rm pc} = \begin{pmatrix} 6.76 \\ 2.97 \\ 5.53 \\ 2.02 \end{pmatrix}, \quad \widetilde{\boldsymbol{\Sigma}}_{\rm pc} = \begin{pmatrix} 0.4310 & 0.0294 & 0.3344 & 0.0625 \\ 0.0294 & 0.0973 & 0.0308 & 0.0440 \\ 0.3344 & 0.0308 & 0.3901 & 0.0566 \\ 0.0625 & 0.0440 & 0.0566 & 0.0765 \end{pmatrix},$$

which give $Q^{\dagger}_{\rm pc} = 14.57$ with p value 0.0680, where $Q^{\dagger}_{\rm pc}$ is the value of the usual MLRT statistic under the partially complete data. If we test the hypothesis at the level of $\alpha = 0.05$, the null hypothesis is rejected using the three-step monotone missing data. On the other hand, the null hypothesis is not rejected using the partially complete data. In the case of this example, since we have the whole sample of Table 5.4, we obtained the MLEs and the value of MLRT statistic under the complete data given by

$$\widehat{\boldsymbol{\mu}}^{(1)}_{\rm c} = \begin{pmatrix} 7.02 \\ 2.95 \\ 5.46 \\ 2.01 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(2)}_{\rm c} = \begin{pmatrix} 6.40 \\ 3.06 \\ 5.42 \\ 2.07 \end{pmatrix}, \quad \widehat{\boldsymbol{\mu}}^{(3)}_{\rm c} = \begin{pmatrix} 6.76 \\ 2.98 \\ 5.68 \\ 1.97 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}_{\rm c} = \begin{pmatrix} 0.2997 & 0.0687 & 0.2261 & 0.0521 \\ 0.0687 & 0.0870 & 0.0527 & 0.0412 \\ 0.2261 & 0.0527 & 0.2745 & 0.0445 \\ 0.0521 & 0.0412 & 0.0445 & 0.0745 \end{pmatrix},$$

and

$$\widetilde{\boldsymbol{\mu}}_{\rm c} = \begin{pmatrix} 6.73 \\ 2.99 \\ 5.52 \\ 2.01 \end{pmatrix}, \quad \widetilde{\boldsymbol{\Sigma}}_{\rm c} = \begin{pmatrix} 0.3643 & 0.0575 & 0.2334 & 0.0456 \\ 0.0575 & 0.0890 & 0.0504 & 0.0426 \\ 0.2334 & 0.0504 & 0.2874 & 0.0403 \\ 0.0456 & 0.0426 & 0.0403 & 0.0762 \end{pmatrix},$$

and $Q^{\dagger}_{\rm c} = 33.39$ with p value 5.24×10^{-5}, where $Q^{\dagger}_{\rm c}$ is the value of the usual MLRT statistic under the complete data. We note that this p value is closer to the respective values 1.68×10^{-4}, 4.63×10^{-4} and 6.30×10^{-4} of the tests proposed in this paper, than the similar one obtained by using partially complete data. In this way, we can test the equality of mean vectors in a one-way MANOVA under the monotone missing data.

5.5 Conclusions

We have developed the LRT, MT and MLRT statistics with general monotone missing data in a one-way MANOVA. Furthermore, we showed that the LR for this test can be decomposed into the LR for the test for a one-way MANOVA of reduced dimension and those of the remaining subvectors with complete data. We note that these LRs are mutually independent. We also derived an asymptotic expansion for the distribution function of each LR using the perturbation procedure. Indeed, the results include the two-sample problem. From the simulation results, the null distribution of MT statistic as well as MLRT statistic is considerably closer to the χ^2 distribution than that of the LRT statistic even if the sample size is moderately small. We recommend the use of MT or MLRT statistic proposed in this paper if the missing data are of monotone pattern in a one-way MANOVA.

Bibliography

[1] Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association*, **52**, 200–203.

[2] Anderson, T. W. (2003). *Introduction to Multivariate Statistical Analysis, 3rd ed.* Hoboken, NJ: Wiley.

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

[4] Batsidis, A. and Zografos, K. (2006). Statistical inference for location and scale of elliptically contoured models with monotone missing data. *Journal of Statistical Planning and Inference*, **136**, 2606–2629.

[5] Bhargava, R. (1962). Multivariate tests of hypotheses with incomplete data. *Technical report No.3, Applied Mathematics and Statistics Laboratories, Stanford University*, Stanford, California.

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

[7] Fujikoshi, Y. (2000). Transformations with improved chi-squared approximations. *Journal of Multivariate Analysis*, **72**, 249–263.

[8] Fujikoshi, Y., Ulyanov, V. V. and Shimizu, R. (2010). *Multivariate Statistics: High-Dimensional and Large-Sample Approximations*. Hoboken, NJ: Wiley.

[9] Gupta, A. K., Xu, J. and Fujikoshi, Y. (2006). An asymptotic expansion of the distribution of Rao's U-statistic under a general condition. *Journal of Multivariate Analysis*, **97**, 492–513.

[10] 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.

[11] 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.

[12] 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.

[13] Jinadasa, K. G. and Tracy, D. S. (1992). Maximum likelihood estimation for multivariate normal distribution with monotone sample. *Communications in Statistics – Theory and Methods*, **21**, 41–50.

[14] Kakizawa, Y. (1996). Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika*, **83**, 923–927.

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

[16] Kawasaki, T. and Seo, T. (2016). Bias correction for T^2 type statistic with two-step monotone missing data. *Statistics*, **50**, 76–88.

[17] Krishnamoorthy, K. and Pannala, M. K. (1998). Some simple test procedures for normal mean vector with incomplete data. *Annals of the Institute of Statistical Mathematics*, **50**, 531–542.

[18] Krishnamoorthy, K. and Pannala, M. K. (1999). Confidence estimation of a normal mean vector with incomplete data. *The Canadian Journal of Statistics*, **27**, 395–407.

[19] Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data, 2nd ed.* Wiley, New York.

[20] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Hoboken, NJ: Wiley.

[21] Rao, C. R. (1956). Analysis of dispersion with incomplete observations on one of the characters. *Journal of the Royal Statistical Society: Series B*, **18**, 259–264.

[22] Romer, M. M. and Richards, D. St. P. (2013). Finite-sample inference with monotone incomplete multivariate normal data, III: Hotelling's T^2-statistic. *Statistical Modelling*, **13**, 431–457.

[23] Seko, N. (2012). Tests for mean vectors with two-step monotone missing data for the k-sample problem. *SUT Journal of Mathematics*, **48**, 213–229.

[24] Seko, N., Kawasaki, T. and Seo, T. (2011). Testing equality of two mean vectors with two-step monotone missing data. *American Journal of Mathematical and Management Sciences*, **31**, 117–135.

[25] Seko, N., Yamazaki, A. and Seo, T. (2012). Tests for mean vector with two-step monotone missing data. *SUT Journal of Mathematics*, **48**, 13–36.

[26] Seo, T., Mano, S. and Fujikoshi, Y. (1994). A generalized Tukey conjecture for multiple comparisons among mean vectors. *Journal of the American Statistical Association*, **89**, 676–679.

[27] Seo, T. and Siotani, M. (1992). The multivariate Studentized range and its upper percentiles. *Journal of the Japan Statistical Society*, **22**, 123–137.

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

[29] Siotani M., Hayakawa T. and Fujikoshi Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*. Columbus, OH: American Science Press.

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

[31] Srivastava, M. S. (2002). *Methods of Multivariate Statistics*. New York: Wiley.

[32] Srivastava, M. S. and Carter, E. M. (1983). *An Introduction to Applied Multivariate Statistics*. North-Holland, New York.

[33] Tsukada S. (2014). Equivalence testing of mean vector and covariance matrix for multi-populations under a two-step monotone incomplete sample. *Journal of Multivariate Analysis*, **132**, 183–196.

[34] Yagi, A. and Seo, T. (2014). A test for mean vector and simultaneous confidence intervals with three-step monotone missing data. *American Journal of Mathematical and Management Sciences*, **33**, 161–175.

[35] Yagi, A. and Seo, T. (2015a). Tests for mean vectors with two-step and three-step monotone samples. *Josai Mathematical Monographs*, **8**, 49–71.

[36] Yagi, A. and Seo, T. (2015b). Tests for equality of mean vectors and simultaneous confidence intervals with two-step or three-step monotone missing data patterns. *American Journal of Mathematical and Management Sciences*, **34**, 213–233.

[37] Yagi, A. and Seo, T. (2017). Tests for normal mean vectors with monotone incomplete data. *American Journal of Mathematical and Management Sciences*, **36**, 1–20.

[38] Yagi, A., Seo, T. and Srivastava, M. S. (2017a). A modified likelihood ratio test for a mean vector with monotone missing data. *Discussiones Mathematicae Probability and Statistics*, **37**, 37–53.

[39] Yagi, A., Seo, T. and Srivastava, M. S. (2017b). Testing equality of mean vectors in a one-way MANOVA with monotone missing data. to appear in *Communications in Statistics – Theory and Methods*.

[40] Yu, J., Krishnamoorthy, K. and Pannala, M. K. (2006). Two-sample inference for normal mean vectors based on monotone missing data. *Journal of Multivariate Analysis*, **97**, 2162–2176.

[41] Wei, L. J. and Lachin, J. M. (1984). Two-sample asymptotically distribution-free tests for incomplete multivariate observations. *Journal of the American Statistical Association*, **79**, 653–661.