Josai Mathematical Monographs
vol. 8 (2015), pp. 49-71

Tests for mean vectors with two-step and three-step monotone samples

Ayaka YAGI and Takashi SEO

Abstract. In this paper, we consider tests for mean vectors when the data have a monotone pattern of missing observations. In particular, we focus on the one-sample problem of testing for mean vector with two-step or three-step monotone missing data, and the two-sample problem of testing the equality of two mean vectors with two-step monotone missing data, where two data sets have the same monotone missing data pattern. To test these problems, we propose two kinds of approximate upper percentiles of the Hotelling's T^2-type statistics. Finally, the accuracy and asymptotic behavior of the approximation are investigated by Monte Carlo simulation.

1. Introduction

The one-sample and the two-sample problems of testing for mean vectors with two-step or three-step monotone missing data are considered in this study. The maximum likelihood estimators (MLEs) for the mean vector and the covariance matrix in the case of the general k-step monotone missing data pattern have been obtained as closed form expressions by Jinadasa and Tracy (1992). Kanda and Fujikoshi (1998) discussed the distributions of the MLEs using different notations and approach. These results were derived for the one-sample problem. For the two-sample problem, the MLEs in the case of two-step and three-step monotone missing data are given in Seko, Kawasaki and Seo (2011) and Yagi and Seo (2014b), respectively. In this paper, we first consider the one-sample problem of the test for the mean vector when the data have two-step or three-step monotone missing data. In the case of a two-step monotone missing data pattern, the usual Hotelling's T^2 statistic and some properties were derived by Chang and Richards (2009) and Seko, Yamazaki and Seo (2012), among others. Further, for the case of a three-step monotone missing data pattern, Krishnamoorthy and Pannala (1999) derived the Hotelling's T^2 statistic and F approximation, and recently Yagi and Seo (2014a) gave a simplified Hotelling's T^2-type statistic and its approximation to the upper percentiles. In this paper, using the notations in Jinadasa and Tracy (1992), we give the usual Hotelling's T^2 statistic for two-step monotone sample, and the simplified Hotelling's T^2-type statistic for two-step and three-step monotone samples. Further, we propose two kinds of approximation to the upper percentiles of the null distributions.

Second, we test the equality of two mean vectors with two-step monotone missing data, where two data sets have the same monotone missing pattern. For related discussions on this issue, see Seko et al. (2011) and Yagi and Seo (2014b). In this paper, as with the one-sample problem, we give the simplified Hotelling's T^2-type statistic and two kinds of approximation to the upper percentile in the case of two-step monotone missing data. Through this paper, we assume that the data are missing completely at random (MCAR). The related discussion is given by Hao and Krishnamoorthy (2001), Little and Rubin (2002), and Chang and Richards (2009), among others.

The remainder of this paper is organized as follows. In Section 2, we first present some notations as preliminaries. Then, the MLEs of the mean vector and the covariance matrix for one-sample problem are given under two-step and three-step monotone missing data. Further, in order to give the usual and the simplified Hotelling's T^2-type statistics, the covariance of the MLE of the mean vector is also given. We give the approximate upper percentiles using linear interpolation based on complete data sets and adjusting the degrees of freedom of the F distribution. In Section 3, for the two-sample problem, we discuss the testing equality of two mean vectors with two-step monotone missing data where two data sets have the same missing pattern. Finally, in Section 4, we give some simulation results and state our conclusions.

2. One-sample case

In this section, we consider the one-sample problem of the test for the mean vector with two-step and three-step monotone missing data. We present some notations and the MLEs, and the covariance of the MLE of the mean vector in order to obtain the Hotelling's T^2-type test statistics.

2.1. Assumptions and notations

We consider the one-sample problem of testing for a mean vector with a three-step monotone missing data pattern. If $\boldsymbol{x}_{ij}$ denotes the jth observation on $\boldsymbol{x}_i$, $i = 1, 2, 3, j = 1, 2, \ldots, n_i$, then the three-step monotone missing data set is of the form given in Figure 1, where $p = p_1 > p_2 > p_3 > 0$, $n_1 > p$, and " $*$ " indicates a missing observation. Such a data set is called a three-step monotone missing data pattern. That is, $(\boldsymbol{x}_{11}, \boldsymbol{x}_{12}, \ldots, \boldsymbol{x}_{1n_1})'$ is an $n_1 \times p_1$ complete data set, and $(\boldsymbol{x}_{21}, \boldsymbol{x}_{22}, \ldots, \boldsymbol{x}_{2n_2})'$ and $(\boldsymbol{x}_{31}, \boldsymbol{x}_{32}, \ldots, \boldsymbol{x}_{3n_3})'$ are $n_2 \times p_2$ and $n_3 \times p_3$ incomplete data sets, respectively. 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

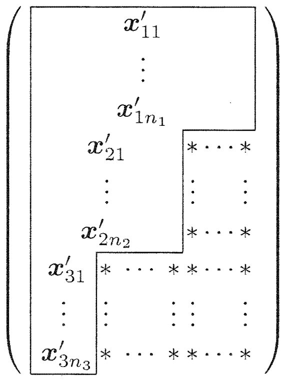

Figure 1: Three-step monotone missing data set

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

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

Note that $\boldsymbol{\Sigma}_{i+1}$ is an upper left submatrix of $\boldsymbol{\Sigma}_i$ and the above gives a partition of $\boldsymbol{\Sigma}_i$. For example, we can express

$$\boldsymbol{\Sigma}_1 = \left(\begin{array}{c|c} \overbrace{\boldsymbol{\Sigma}_2}^{p_2} & \overbrace{\boldsymbol{\Sigma}_{22}}^{p_1-p_2} \\ \hline \boldsymbol{\Sigma}'_{22} & \boldsymbol{\Sigma}_{23} \end{array} \right) \begin{array}{l} \}p_2 \\ \}p_1 - p_2 \end{array} = \left(\begin{array}{c|c} \overbrace{\boldsymbol{\Sigma}_2}^{p_2} & \overbrace{\boldsymbol{\Sigma}_{(1,2)}}^{p_1-p_2} \\ \hline \boldsymbol{\Sigma}'_{(1,2)} & \boldsymbol{\Sigma}_{(1,3)} \end{array} \right) \begin{array}{l} \}p_2 \\ \}p_1 - p_2 \end{array},$$

$$\boldsymbol{\Sigma}_2 = \left(\begin{array}{c|c} \overbrace{\boldsymbol{\Sigma}_3}^{p_3} & \overbrace{\boldsymbol{\Sigma}_{(2,2)}}^{p_2-p_3} \\ \hline \boldsymbol{\Sigma}'_{(2,2)} & \boldsymbol{\Sigma}_{(2,3)} \end{array} \right) \begin{array}{l} \}p_3 \\ \}p_2 - p_3 \end{array}.$$

Also, we have

$$\boldsymbol{\Sigma}_1 = \left(\begin{array}{cc|c} \overbrace{\boldsymbol{\Sigma}_3}^{p_3} & \overbrace{\boldsymbol{\Sigma}_{(2,2)}}^{p_2-p_3} & \overbrace{}^{p_1-p_2} \\ \boldsymbol{\Sigma}'_{(2,2)} & \boldsymbol{\Sigma}_{(2,3)} & \boldsymbol{\Sigma}_{22} \\ \hline \multicolumn{2}{c|}{\boldsymbol{\Sigma}'_{22}} & \boldsymbol{\Sigma}_{23} \end{array} \right) \begin{array}{l} \}p_3 \\ \}p_2 - p_3 \\ \}p_1 - p_2 \end{array}.$$

Since the assumptions and notations of two-step monotone missing data are almost the same as those of three-step monotone missing data, we omit them for the two-step monotone case.

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), and Srivastava (2002), among others.

2.2. Two-step and three-step monotone missing data

In this subsection, since any discussion over a two-step monotone missing data pattern is the same as that over a three-step monotone missing data pattern, we first consider the one-sample problem of testing for the mean vector with three-step monotone missing data. In the case of two-step monotone missing data, Anderson and Olkin (1985) derived the MLEs of the mean vector and the covariance matrix. Kanda and Fujikoshi (1998) discussed the distribution of the MLEs in the case of general k-step monotone missing data including the two-step and three-step cases. Further, as another approach, Jinadasa and Tracy (1992) gave the closed form expressions for the MLEs in the case of k-step monotone missing data. In this paper, we present the MLEs with three-step and two-step monotone missing data. These results are simple and useful to derive the Hotelling's T^2-type statistics for the one-sample problem of testing for the mean vector. Using the notations by Jinadasa and Tracy (1992), we have the following theorem.

THEOREM 1. *Suppose that $\boldsymbol{x}_{ij}$ is distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2, 3$, $j = 1, 2, \ldots, n_i$, where $p = p_1 > p_2 > p_3 > 0$ and $n_1 > p$. 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

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

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

$$\widehat{\boldsymbol{T}}_2 = \begin{pmatrix} \boldsymbol{I}_{p_2} \\ \widehat{\boldsymbol{\Sigma}}'_{(1,2)} \widehat{\boldsymbol{\Sigma}}_2^{-1} \end{pmatrix}, \quad \widehat{\boldsymbol{T}}_3 = \begin{pmatrix} \boldsymbol{I}_{p_3} \\ \widehat{\boldsymbol{\Sigma}}'_{(2,2)} \widehat{\boldsymbol{\Sigma}}_3^{-1} \end{pmatrix},$$

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

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

where

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

$$\boldsymbol{G}_2 = \begin{pmatrix} \boldsymbol{I}_{p_2} \\ \boldsymbol{L}_{12}'\boldsymbol{L}_{11}^{-1} \end{pmatrix}, \quad \boldsymbol{G}_3 = \begin{pmatrix} \boldsymbol{I}_{p_3} \\ \boldsymbol{L}_{22}'\boldsymbol{L}_{21}^{-1} \end{pmatrix},$$

$$\boldsymbol{L}_1 = \boldsymbol{E}_1, \quad \boldsymbol{L}_2 = \boldsymbol{L}_{11} + \boldsymbol{E}_2 + \frac{N_2N_3}{n_2}\boldsymbol{d}_2\boldsymbol{d}_2',$$

$$\boldsymbol{L}_i = \begin{pmatrix} \boldsymbol{L}_{i1} & \boldsymbol{L}_{i2} \\ \boldsymbol{L}_{i2}' & \boldsymbol{L}_{i3} \end{pmatrix}, \quad i = 1, 2.$$

The results in Theorem 1 follow from the results in Yagi and Seo (2014a). If $\boldsymbol{x}_{ij}$ is distributed as $N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, 2$, $j = 1, 2, \ldots, n_i$, then such a data set $\{\boldsymbol{x}_{ij}\}$ is called a two-step monotone missing data pattern, where $p = p_1 > p_2 > 0$ and $n_1 > p$. For the two-step case, we have the following results.

COROLLARY 2. *If the data have a two-step monotone pattern of missing observations, then the MLEs of the mean vector and the covariance matrix are given by*

$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \widehat{\boldsymbol{T}}_2\boldsymbol{d}_2,$$
$$\widehat{\Sigma} = \frac{1}{N_2}\boldsymbol{E}_1 + \frac{1}{N_3}\boldsymbol{G}_2\left[\boldsymbol{E}_2 + \frac{N_2N_3}{n_2}\boldsymbol{d}_2\boldsymbol{d}_2' - \frac{n_2}{N_2}\boldsymbol{L}_{11}\right]\boldsymbol{G}_2',$$

respectively.

Next, we consider the following hypothesis test:

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

where $\boldsymbol{\mu}_0$ is known. We assume that the data set has a three-step or a two-step monotone missing data pattern. Without loss of generality, we can assume that $\boldsymbol{\mu}_0 = \mathbf{0}$. To test H_0, we consider the usual Hotelling's T^2 statistic given by

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

where $\widehat{\boldsymbol{\Gamma}}(= \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}])$ is an estimator of $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}]$. For three-step monotone missing data, $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}]$ has been derived using other notations (see Kanda and Fujikoshi (1998)). However, since it is very complicated, we use the simplified Hotelling's T^2-type statistic with $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}]$. That is, we consider

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

where $\widetilde{\boldsymbol{\Gamma}}(= \widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}])$ is an estimator of $\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}]$ and $\widetilde{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \boldsymbol{T}_2\boldsymbol{d}_2 + \boldsymbol{T}_2\boldsymbol{T}_3\boldsymbol{d}_3$,

$$\boldsymbol{T}_2 = \begin{pmatrix} \boldsymbol{I}_{p_2} \\ \boldsymbol{\Sigma}'_{(1,2)}\boldsymbol{\Sigma}_2^{-1} \end{pmatrix}, \quad \boldsymbol{T}_3 = \begin{pmatrix} \boldsymbol{I}_{p_3} \\ \boldsymbol{\Sigma}'_{(2,2)}\boldsymbol{\Sigma}_3^{-1} \end{pmatrix}.$$

Then, we have the following theorem.

THEOREM 3. *If the data have a three-step monotone pattern of missing observations, then the covariance matrix of* $\widetilde{\boldsymbol{\mu}}$ *is given by*

$$\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \frac{1}{N_2}\boldsymbol{\Sigma}_1 - \frac{n_2}{N_2 N_3}\boldsymbol{U}_2 - \frac{n_3}{N_3 N_4}\boldsymbol{U}_3,$$

where

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

This result was derived by Yagi and Seo (2014a). Further, from Theorem 3, we have the following corollary.

COROLLARY 4. *If the data have a two-step monotone pattern of missing observations, then the covariance matrix of* $\widetilde{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}_1 + \boldsymbol{T}_2\boldsymbol{d}_2$ *is given by*

$$\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \frac{1}{N_2}\boldsymbol{\Sigma}_1 - \frac{n_2}{N_2 N_3}\boldsymbol{U}_2.$$

Further, in order to give the covariance of $\widehat{\mu}$ for a two-step monotone missing data pattern, we have the following lemma.

LEMMA 5. *Suppose that* $\boldsymbol{E}_1$ *is distributed as the Wishart distribution with the covariance matrix* $\boldsymbol{\Sigma}_1$ *and* $n_1 - 1$ *degrees of freedom,* $W_{p_1}(\boldsymbol{\Sigma}_1, n_1 - 1)$, *where* $n_1 > p_1 = p$, *and let*

$$\boldsymbol{E}_1 = \begin{pmatrix} (\boldsymbol{E}_1)_2 & \boldsymbol{E}_{(1,2)} \\ \boldsymbol{E}'_{(1,2)} & \boldsymbol{E}_{(1,3)} \end{pmatrix},$$

where $(\boldsymbol{E}_1)_2$ *is the principal submatrix of* $\boldsymbol{E}_1$ *of order* $p_2 \times p_2$. *Let* $\boldsymbol{E}_{(1,3)\cdot 2} = \boldsymbol{E}_{(1,3)} - \boldsymbol{E}'_{(1,2)}(\boldsymbol{E}_1)_2^{-1}\boldsymbol{E}_{(1,2)}$. *Then,*

(i) $\boldsymbol{E}_{(1,3)\cdot 2} \sim W_{p_1-p_2}(\boldsymbol{\Sigma}_{(1,3)\cdot 2}, n_1 - p_2 - 1)$ *and is independent of* $(\boldsymbol{E}_1)_2$ *and* $\boldsymbol{E}_{(1,2)}$, *where* $\boldsymbol{\Sigma}_{(1,3)\cdot 2} = \boldsymbol{\Sigma}_{(1,3)} - \boldsymbol{\Sigma}'_{(1,2)}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{(1,2)}$;

(ii) *The conditional distribution of* $\mathrm{Vec}[\boldsymbol{E}_{(1,2)}]$ *given* $(\boldsymbol{E}_1)_2$ *is*

$$N_{(p_1-p_2)p_2}\left(\mathrm{Vec}\left[(\boldsymbol{E}_1)_2\Sigma_2^{-1}\boldsymbol{\Sigma}_{(1,2)}\right], \boldsymbol{\Sigma}_{(1,3)\cdot 2}\otimes(\boldsymbol{E}_1)_2\right),$$

and, in particular, $\mathrm{E}[(\boldsymbol{E}_1)_2^{-1}\boldsymbol{E}_{(1,2)}] = \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{(1,2)}$;

(iii) $(\boldsymbol{E}_1)_2 \sim W_{p_2}(\boldsymbol{\Sigma}_2, n_1 - 1)$;

(iv) *if* $n_1 - p_2 - 2 > 0$,

$$\mathrm{E}[(\boldsymbol{E}_1)_2^{-1}] = \frac{1}{n_1 - p_2 - 2}\boldsymbol{\Sigma}_2^{-1};$$

(v) *if* $n_1 - p_2 - 2 > 0$,

$$\begin{aligned}&\mathrm{E}[\boldsymbol{E}'_{(1,2)}(\boldsymbol{E}_1)_2^{-1}\boldsymbol{C}(\boldsymbol{E}_1)_2^{-1}\boldsymbol{E}_{(1,2)}]\\ &\quad = \mathrm{E}[\mathrm{tr}\{(\boldsymbol{E}_1)_2^{-1}\boldsymbol{C}\}]\boldsymbol{\Sigma}_{(1,3)\cdot 2} + \boldsymbol{\Sigma}'_{(1,2)}\boldsymbol{\Sigma}_2^{-1}\mathrm{E}[\boldsymbol{C}]\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{(1,2)},\end{aligned}$$

where $\boldsymbol{C}$ *is a random matrix depending on* $(\boldsymbol{E}_1)_2$.

Note that Lemma 5 is obtained by rewriting the result in Kanda and Fujikoshi (1998) using the notations in Jinadasa and Tracy (1992) and Yagi and Seo (2014a) in order to obtain the next theorem. Further, note that $\boldsymbol{E}_2 \neq (\boldsymbol{E}_1)_2$ but $\boldsymbol{E}_2 = \sum_{j=1}^{n_2}(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_2)(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_2)'$.

THEOREM 6. *If the data have a two-step monotone pattern of missing observations, then the expectation and the covariance matrix of* $\widehat{\boldsymbol{\mu}}$ *are given by*

(i) $\mathrm{E}[\widehat{\boldsymbol{\mu}}] = \boldsymbol{\mu}$,

(ii) $$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \begin{pmatrix} \frac{1}{N_3}\boldsymbol{\Sigma}_2 & \frac{1}{N_3}\boldsymbol{\Sigma}_{22} \\ \frac{1}{N_3}\boldsymbol{\Sigma}_{22}' & \frac{1}{N_2}\left(\boldsymbol{\Sigma}_{23} - \frac{n_2}{N_3}\boldsymbol{\Sigma}_{22}'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{22}\right) + \frac{n_2 p_2}{N_2 N_3(N_2 - p_2 - 2)}\boldsymbol{\Sigma}_{(1,3)\cdot 2} \end{pmatrix},$$

respectively, where $N_2 > p_2 + 2$. *Further, we can reduce* (ii) *as follows.*

$$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \frac{1}{N_2}\boldsymbol{\Sigma}_1 - \frac{n_2}{N_2 N_3}\boldsymbol{U}_2 + \boldsymbol{R},$$

where

$$\boldsymbol{U}_2 = \begin{pmatrix} \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_{22}' \end{pmatrix} \boldsymbol{T}_2', \quad \boldsymbol{R} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \frac{n_2 p_2}{N_2 N_3 (N_2 - p_2 - 2)}\boldsymbol{\Sigma}_{(1,3)\cdot 2} \end{pmatrix}.$$

PROOF. Result (i) can be easily obtained. Since $\mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \mathrm{E}[\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'] - \boldsymbol{\mu}\boldsymbol{\mu}'$, we can expand as

$$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \mathrm{E}[\overline{\boldsymbol{x}}_1\overline{\boldsymbol{x}}_1'] + \mathrm{E}[\overline{\boldsymbol{x}}_1\boldsymbol{d}_2'\widehat{\boldsymbol{T}}_2'] + \mathrm{E}[\widehat{\boldsymbol{T}}_2\boldsymbol{d}_2\overline{\boldsymbol{x}}_1'] + \mathrm{E}[\widehat{\boldsymbol{T}}_2\boldsymbol{d}_2\boldsymbol{d}_2'\widehat{\boldsymbol{T}}_2'] - \boldsymbol{\mu}\boldsymbol{\mu}'.$$

Then, using Lemma 5, each term is given by

$$\mathrm{E}[\overline{\boldsymbol{x}}_1\overline{\boldsymbol{x}}_1'] = \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \boldsymbol{\mu}_1\boldsymbol{\mu}_1',$$

$$\mathrm{E}[\overline{\boldsymbol{x}}_1\boldsymbol{d}_2'\widehat{\boldsymbol{T}}_2'] = -\frac{n_2}{n_1 N_3}\begin{pmatrix} \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_{(1,2)} \\ \boldsymbol{\Sigma}_{22}' & \boldsymbol{\Sigma}_{22}'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{(1,2)} \end{pmatrix},$$

$$\mathrm{E}[\widehat{\boldsymbol{T}}_2\boldsymbol{d}_2\overline{\boldsymbol{x}}_1'] = -\frac{n_2}{n_1 N_3}\begin{pmatrix} \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_{22} \\ \boldsymbol{\Sigma}_{(1,2)}' & \boldsymbol{\Sigma}_{(1,2)}'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{22} \end{pmatrix},$$

$$\mathrm{E}[\widehat{\boldsymbol{T}}_2\boldsymbol{d}_2\boldsymbol{d}_2'\widehat{\boldsymbol{T}}_2'] = \frac{n_2}{n_1N_3}\begin{pmatrix} \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_{(1,2)} \\ \boldsymbol{\Sigma}_{(1,2)}' & \dfrac{p_2}{n_1-p_2-2}\boldsymbol{\Sigma}_{(1,3)\cdot 2} + \boldsymbol{\Sigma}_{(1,2)}'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{(1,2)} \end{pmatrix}.$$

By $\boldsymbol{\Sigma}_{(1,2)} = \boldsymbol{\Sigma}_{22}$, and $N_2 = n_1$, the proof is complete. □

From Corollary 4 and Theorem 6, we have

$$\mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] = \begin{pmatrix} \dfrac{1}{N_3}\boldsymbol{\Sigma}_2 & \dfrac{1}{N_3}\boldsymbol{\Sigma}_{22} \\ \dfrac{1}{N_3}\boldsymbol{\Sigma}_{22}' & \dfrac{1}{N_2}\left(\boldsymbol{\Sigma}_{23} - \dfrac{n_2}{N_3}\boldsymbol{\Sigma}_{22}'\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_{22}\right) \end{pmatrix},$$

$$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \mathrm{Cov}[\widetilde{\boldsymbol{\mu}}] + \boldsymbol{R}.$$

Therefore, in particular, for two-step monotone sample, T^2 and $\widetilde{T}^2$ statistics are reduced as $T^2 = \widehat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Gamma}}^{-1}\widehat{\boldsymbol{\mu}}$ and $\widetilde{T}^2 = \widehat{\boldsymbol{\mu}}'\widetilde{\boldsymbol{\Gamma}}^{-1}\widehat{\boldsymbol{\mu}}$, respectively, where

$$\widehat{\boldsymbol{\Gamma}} = \widetilde{\boldsymbol{\Gamma}} + \widehat{\boldsymbol{R}}, \quad \widetilde{\boldsymbol{\Gamma}} = \frac{1}{N_2}\widehat{\boldsymbol{\Sigma}}_1 - \frac{n_2}{N_2N_3}\widehat{\boldsymbol{U}}_2.$$

However, it is not easy to find the exact null distributions of the Hotelling's T^2-type statistics. In addition, the upper percentiles of the χ^2 distribution are not good approximations for small sample sizes, although the χ^2 distribution is the asymptotic distribution of Hotelling's T^2-type statistics. Indeed, under H_0, the T^2-type statistics are asymptotically distributed as a χ^2 distribution with p degrees of freedom; when $n_1, N_3 \to \infty$ with $n_1/N_3 \to \delta \in (0,1]$ under the two-step case, and when $n_1, N_4 \to \infty$ with $n_1/N_4 \to \delta \in (0,1]$ under the three-step case.

Therefore, we give their approximate upper percentiles using the linear interpolation based on the complete data set and adjusting the degrees of freedom of the F distribution.

THEOREM 7. *If the data have a three-step monotone pattern of missing observations, then the two kinds of approximate upper* 100α *percentiles of the* $\widetilde{T}^2$ *statistic are given by*

$$t^2_{\mathrm{YS\cdot L}}(\alpha) = (1-c)T^2_{n_1,\alpha} + cT^2_{N_4,\alpha},$$

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

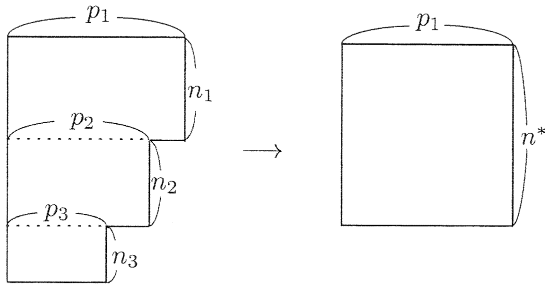

Figure 2: Approximation adjusting the degrees of freedom of the F distribution

where

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

$$c = \frac{n_2 p_2 + n_3 p_3}{(n_2 + n_3) p_1}, \quad n^* = \frac{1}{p_1} \sum_{i=1}^{3} n_i p_i,$$

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

Figure 2 shows that n^* is the solution to the equation $\sum_{i=1}^{3} n_i p_i = n^* p_1$. We note that $T^2_{n_1} = n_1 \overline{\boldsymbol{x}}_1' \widehat{\Sigma}_{\rm ML}^{-1} \overline{\boldsymbol{x}}_1$ is distributed as $n_1 p_1/(n_1 - p_1) F_{p_1, n_1 - p_1}$ since $\widehat{\Sigma}_{\rm ML} = (1/n_1) \sum_{j=1}^{n_1} (\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_1)(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_1)'$. For two-step monotone missing data, we have the following corollary.

COROLLARY 8. *If the data have a two-step monotone pattern of missing observations, then the two kinds of approximate upper percentiles of the T^2 (or $\widetilde{T}^2$) statistic are given by*

$$t^2_{\rm YS\cdot L}(\alpha) = \left(1 - \frac{p_2}{p_1}\right) T^2_{n_1,\alpha} + \frac{p_2}{p_1} T^2_{N_3,\alpha},$$

$$t^2_{{\rm YS}\cdot F}(\alpha) = \frac{n^* p_1}{n^* - p_1} F_{p_1, n^* - p_1, \alpha},$$

where

$$T^2_{n_1,\alpha} = \frac{n_1 p_1}{n_1 - p_1} F_{p_1, n_1 - p_1, \alpha}, \quad T^2_{N_3,\alpha} = \frac{N_3 p_1}{N_3 - p_1} F_{p_1, N_3 - p_1, \alpha},$$

$$n^* = \frac{1}{p_1}(n_1 p_1 + n_2 p_2).$$

Note that the values of $t^2_{\mathrm{YS\cdot L}}(\alpha)$ and $t^2_{\mathrm{YS\cdot}F}(\alpha)$ can be used as approximations to the upper percentiles of both T^2 and $\widetilde{T}^2$ statistics.

3. Two-sample case

In this section, we consider the testing equality of two mean vectors with two-step monotone missing data where two data sets have the same missing pattern and a common covariance matrix.

Let $\widehat{\boldsymbol{\mu}}^{(\ell)}$, $\ell = 1, 2$, and $\widehat{\boldsymbol{\Sigma}}^{[pl]}$ be the MLEs of $\boldsymbol{\mu}^{(\ell)}$, $\ell = 1, 2$, and $\boldsymbol{\Sigma}$, respectively. Then, the MLEs of $\boldsymbol{\mu}^{(\ell)}$, $\ell = 1, 2$, and $\boldsymbol{\Sigma}$ are derived using other notations (see Seko et al. (2011)). In this section, we present the MLEs using the notations in Yagi and Seo (2014b). To test the hypothesis:

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

we also present the Hotelling's T^2-type statistics,

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

and

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

where

$$\widehat{\boldsymbol{\Gamma}}^{[pl]} = \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}], \quad \widetilde{\boldsymbol{\Gamma}}^{[pl]} = \widehat{\mathrm{Cov}}[\widetilde{\boldsymbol{\mu}}^{(1)} - \widetilde{\boldsymbol{\mu}}^{(2)}].$$

First, we give the MLEs of the mean vectors and the covariance matrix when the two data sets have the same pattern of two-step monotone missing data.

THEOREM 9. *Let* $\boldsymbol{x}_{ij}^{(\ell)}$, $i=1,2$, $j=1,2,\ldots,n_i^{(\ell)}$, $\ell=1,2$, *be the* j*-th random vector from the* ℓ*-th population distributed as* $N_{p_i}(\boldsymbol{\mu}_i^{(\ell)},\boldsymbol{\Sigma}_i)$, *where* $p=p_1>p_2>0$ *and* $n_1^{(\ell)}>p$, $\ell=1,2$. *Then, the MLEs of* $\boldsymbol{\mu}^{(\ell)}$, $\ell=1,2$, *and* $\boldsymbol{\Sigma}$ *are given by*

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

$$\widehat{\boldsymbol{\Sigma}}^{[pl]}=\frac{1}{N_2^{(1)}+N_2^{(2)}}\sum_{\ell=1}^{2}\boldsymbol{E}_1^{(\ell)}+\frac{1}{N_3^{(1)}+N_3^{(2)}}\sum_{\ell=1}^{2}\boldsymbol{G}_2^{(\ell)}\left[\boldsymbol{E}_2^{(\ell)}+\frac{N_2^{(\ell)}N_3^{(\ell)}}{n_2^{(\ell)}}\boldsymbol{d}_2^{(\ell)}\boldsymbol{d}_2^{(\ell)\prime}-\frac{n_2^{(1)}+n_2^{(2)}}{N_2^{(1)}+N_2^{(2)}}\boldsymbol{L}_{11}^{(\ell)}\right]\boldsymbol{G}_2^{(\ell)\prime},$$

respectively, where

$$\boldsymbol{d}_2^{(\ell)}=\frac{n_2^{(\ell)}}{N_3^{(\ell)}}\left[\overline{\boldsymbol{x}}_2^{(\ell)}-(\overline{\boldsymbol{x}}_1^{(\ell)})_2\right],\quad N_{i+1}^{(\ell)}=\sum_{j=1}^{i}n_j^{(\ell)},\quad i=1,2,$$

$$\widehat{\boldsymbol{T}}_2^{[pl]}=\begin{pmatrix}\boldsymbol{I}_{p_2}\\ \widehat{\boldsymbol{\Sigma}}_{(1,2)}^{[pl]\prime}\widehat{\boldsymbol{\Sigma}}_2^{[pl]^{-1}}\end{pmatrix},\quad \boldsymbol{G}_2^{(\ell)}=\begin{pmatrix}\boldsymbol{I}_{p_2}\\ \left(\boldsymbol{L}_{12}^{(1)}+\boldsymbol{L}_{12}^{(2)}\right)^{\prime}\left(\boldsymbol{L}_{11}^{(1)}+\boldsymbol{L}_{11}^{(2)}\right)^{-1}\end{pmatrix},$$

$$\boldsymbol{L}_1^{(\ell)}=\boldsymbol{E}_1^{(\ell)}=\begin{pmatrix}\boldsymbol{L}_{11}^{(\ell)} & \boldsymbol{L}_{12}^{(\ell)}\\ \boldsymbol{L}_{12}^{(\ell)\prime} & \boldsymbol{L}_{13}^{(\ell)}\end{pmatrix}.$$

Further, as with the one-sample problem, we can easily obtain the covariance matrices of $\widehat{\boldsymbol{\Gamma}}^{[pl]}$ and $\widetilde{\boldsymbol{\Gamma}}^{[pl]}$,

$$\widehat{\boldsymbol{\Gamma}}^{[pl]}=\widetilde{\boldsymbol{\Gamma}}^{[pl]}+\widehat{\boldsymbol{R}}^{[pl]}$$

and

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

respectively, where $\widetilde{\boldsymbol{\mu}}^{(\ell)}=\overline{\boldsymbol{x}}_1^{(\ell)}+\boldsymbol{T}_2\boldsymbol{d}_2^{(\ell)}$, $\ell=1,2$,

$$\widehat{\boldsymbol{U}}_2^{[pl]}=\begin{pmatrix}\widehat{\boldsymbol{\Sigma}}_2^{[pl]}\\ \widehat{\boldsymbol{\Sigma}}_{22}^{[pl]\prime}\end{pmatrix}\widehat{\boldsymbol{T}}_2^{[pl]\prime},\quad \widehat{\boldsymbol{R}}^{[pl]}=\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & \displaystyle\sum_{\ell=1}^{2}\frac{n_2^{(\ell)}p_2}{N_2^{(\ell)}N_3^{(\ell)}(N_2^{(\ell)}-p_2-2)}\widehat{\boldsymbol{\Sigma}}_{(1,3)\cdot 2}^{[pl]}\end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}^{[pl]}_{(1,3)\cdot 2} = \widehat{\boldsymbol{\Sigma}}^{[pl]}_{(1,3)} - \widehat{\boldsymbol{\Sigma}}^{[pl]'}_{(1,2)} \widehat{\boldsymbol{\Sigma}}^{[pl]^{-1}}_{2} \widehat{\boldsymbol{\Sigma}}^{[pl]}_{(1,2)}$$

and $N_2^{(\ell)} > p_2 + 2$, $\ell = 1, 2$. In addition, we can express $\widetilde{\boldsymbol{\Gamma}}^{[pl]}$ as

$$\widetilde{\boldsymbol{\Gamma}}^{[pl]} = \begin{pmatrix} \left(\sum_{\ell=1}^{2} \frac{1}{N_3^{(\ell)}}\right) \widehat{\boldsymbol{\Sigma}}^{[pl]}_{2} & \left(\sum_{\ell=1}^{2} \frac{1}{N_3^{(\ell)}}\right) \widehat{\boldsymbol{\Sigma}}^{[pl]}_{22} \\ \left(\sum_{\ell=1}^{2} \frac{1}{N_3^{(\ell)}}\right) \widehat{\boldsymbol{\Sigma}}^{[pl]'}_{22} & \left(\sum_{\ell=1}^{2} \frac{1}{N_2^{(\ell)}}\right) \widehat{\boldsymbol{\Sigma}}^{[pl]}_{23} - \left(\sum_{\ell=1}^{2} \frac{n_2^{(\ell)}}{N_2^{(\ell)} N_3^{(\ell)}}\right) \widehat{\boldsymbol{\Sigma}}^{[pl]'}_{22} \widehat{\boldsymbol{\Sigma}}^{[pl]^{-1}}_{2} \widehat{\boldsymbol{\Sigma}}^{[pl]}_{22} \end{pmatrix}.$$

From the above results, we can obtain the T^2 and $\widetilde{T}^2$ statistics, whose null distributions are asymptotically the χ^2 distribution with p degrees of freedom when $n_1^{(\ell)}, N_3^{(\ell)} \to \infty$ with $n_1^{(\ell)}/N_3^{(\ell)} \to \delta^{(\ell)} \in (0, 1], \ell = 1, 2$. Then, we can propose the approximation to the upper percentiles of the T^2 and $\widetilde{T}^2$ statistics.

THEOREM 10. *If two data sets have the same type of two-step monotone missing pattern, then the two kinds of approximate upper* 100α *percentiles of the* T^2 *(or* $\widetilde{T}^2$*) statistic are given by*

$$t^2_{\mathrm{YS\cdot L}}(\alpha) = \left(1 - \frac{p_2}{p_1}\right) T^2_{n,\alpha} + \frac{p_2}{p_1} T^2_{N,\alpha},$$

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

where

$$T^2_{n,\alpha} = \frac{n p_1}{n - p_1 - 1} F_{p_1, n - p_1 - 1, \alpha}, \quad T^2_{N,\alpha} = \frac{N p_1}{N - p_1 - 1} F_{p_1, N - p_1 - 1, \alpha},$$

$$n = n_1^{(1)} + n_1^{(2)}, \quad N = N_3^{(1)} + N_3^{(2)},$$

$$n^* = \frac{1}{p_1} \left[(n_1^{(1)} + n_1^{(2)}) p_1 + (n_2^{(1)} + n_2^{(2)}) p_2 \right].$$

4. Simulation studies

We compute the upper percentiles of the Hotelling's T^2-type statistics with two-step or three-step monotone missing data using the Monte Carlo simulation (10^6 runs). That is, the T^2 (or $\widetilde{T}^2$) statistic are computed 10^6 times based on the normal random vectors generated from $N_p(\mathbf{0}, \boldsymbol{I}_p)$. Note that the Hotelling's T^2-type statistics are asymptotically invariant under the nonsingular transformation. In particular, we evaluate the accuracy of the proposed approximations in Theorem

7 and Corollary 8 for the one-sample problem, and that of the proposed approximations in Theorem 10 for the two-sample problem.

4.1. One-sample problem

Tables 1 and 2 give the simulated upper 100α percentiles of the $\widetilde{T}^2$ statistic with a three-step monotone missing data pattern. That is, we provide $\tilde{t}^2_{\rm simu}(=\tilde{t}^2_{\rm simu}(\alpha))$ for $(p_1,p_2,p_3)=(8,4,2)$; $\alpha=0.05,0.01$; and $(n_1,n_2,n_3)=(m_1,m_2,m_3)$, where $m_1=20,50,100,200,400,800$, $m_2=10,20,30$, and $m_3=10,20,30$, and where the sets of (n_1,n_2,n_3) are combinations of m_1, m_2, and m_3. These tables also give the approximations to the upper percentiles of the $\widetilde{T}^2$ statistic, that is, $t^2_{\rm YS\cdot L}(=t^2_{\rm YS\cdot L}(\alpha))$ and $t^2_{{\rm YS}\cdot F}(=t^2_{{\rm YS}\cdot F}(\alpha))$ in Theorem 7. In addition, we provide the simulated coverage probabilities for the approximate upper percentiles in Tables 1 and 2, which are given by

$$\widetilde{\rm CP}(t^2_{\rm YS\cdot L}(\alpha))=1-\Pr\{\widetilde{T}^2>t^2_{\rm YS\cdot L}(\alpha)\},$$

$$\widetilde{\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))=1-\Pr\{\widetilde{T}^2>t^2_{{\rm YS}\cdot F}(\alpha)\},$$

$$\widetilde{\rm CP}(\chi^2_{p,\alpha})=1-\Pr\{\widetilde{T}^2>\chi^2_{p,\alpha}\}.$$

It may be noted from Tables 1 and 2 that the simulated values, $t^2_{\rm simu}(\alpha)$, are closer to the upper percentiles of the χ^2 distribution when the sample size n_1 becomes large. However, the upper percentiles of the χ^2 distribution, $\chi^2_{p,\alpha}$, are not good approximations to those of the $\widetilde{T}^2$ statistic for small sample sizes. At the same time, the proposed approximate upper percentiles $t^2_{\rm YS\cdot L}$ and $t^2_{{\rm YS}\cdot F}$ are good even for small sample sizes, in particular, $t^2_{\rm YS\cdot L}$ is considerably good for all cases.

In Tables 3 and 4, we provide the simulated upper 100α percentiles of the T^2 and $\widetilde{T}^2$ statistics, $t^2_{\rm simu}(=t^2_{\rm simu}(\alpha))$ and $\tilde{t}^2_{\rm simu}(=\tilde{t}^2_{\rm simu}(\alpha))$, and the two kinds of approximation in Corollary 8, $t^2_{\rm YS\cdot L}(=t^2_{\rm YS\cdot L}(\alpha))$ and $t^2_{{\rm YS}\cdot F}(=t^2_{{\rm YS}\cdot F}(\alpha))$, for $(p_1,p_2)=(4,2),(8,4)$; $\alpha=0.05,0.01$; and $(n_1,n_2)=(m_1,m_2)$, where $m_1=10,20,30,50,100$ and $m_2=10,20,50,100$, and where the sets of (n_1,n_2) are combinations of m_1 and m_2. Further, these tables list the simulated coverage probabilities for the approximate values of $t^2_{\rm YS\cdot L}$ and $t^2_{{\rm YS}\cdot F}$ as well as $\chi^2_{p,\alpha}$, which are given by the same coverage probabilities as those in Tables 1 and 2. It may be noted from Tables 3 and 4 that the approximate values $t^2_{\rm YS\cdot L}$ and $t^2_{{\rm YS}\cdot F}$ are closer to the simulated values $\tilde{t}^2_{\rm simu}$ or $t^2_{\rm simu}$ when the sample size becomes large.

In addition, we also provide the simulated coverage probabilities given by

$${\rm CP}(t^2_{\rm YS\cdot L}(\alpha))=1-\Pr\{T^2>t^2_{\rm YS\cdot L}(\alpha)\},$$

$${\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))=1-\Pr\{T^2>t^2_{{\rm YS}\cdot F}(\alpha)\},$$

$${\rm CP}(\chi^2_{p,\alpha})=1-\Pr\{T^2>\chi^2_{p,\alpha}\}.$$

In Tables 1 and 2, we do not provide the simulated values, $t^2_{\rm simu}(\alpha)$ and the simulated coverage probabilities, $\mathrm{CP}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\mathrm{CP}(t^2_{\mathrm{YS}\cdot F}(\alpha))$ and $\mathrm{CP}(\chi^2_{p,\alpha})$ since the T^2 statistic is very complicated and is not derived in the case of three-step monotone missing data.

In particular, it can be seen from the coverage probabilities in Tables 3 and 4 that the proposed approximate upper percentiles, $t^2_{\mathrm{YS\cdot L}}$, are considerably good for all cases, even $n_1 < n_2$. Therefore, it can be concluded that our approximation procedures are very accurate for most of the cases, though $t^2_{\mathrm{YS}\cdot F}$ is not a good approximation for a small n_1 when $n_1 < n_2$.

4.2. Two-sample problem

Tables 5 and 6 give the simulated upper 100α percentiles of the T^2 and $\widetilde{T}^2$ statistics, $t^2_{\mathrm{simu}}(= t^2_{\mathrm{simu}}(\alpha))$ and $\widetilde{t}^2_{\mathrm{simu}}(= \widetilde{t}^2_{\mathrm{simu}}(\alpha))$, for the case where two data sets have the same pattern of two-step monotone missing data. Computations are conducted for $(p_1, p_2) = (4,2), (8,4); \alpha = 0.05, 0.01$; and $(n_1^{(\ell)}, n_2^{(\ell)}) = (m_1, m_2)$, $\ell = 1, 2$, where $m_1 = 10, 20, 30, 50, 100$ and $m_2 = 10, 20, 50, 100$, and where the sets of $(n_1^{(\ell)}, n_2^{(\ell)})$ are combinations of m_1 and m_2. We note that this setting for two data sets is the same with respect to the sample size. Tables 5 and 6 also list the approximations to the upper percentiles of T^2 (or $\widetilde{T}^2$), $t^2_{\mathrm{YS\cdot L}}(= t^2_{\mathrm{YS\cdot L}}(\alpha))$ and $t^2_{\mathrm{YS}\cdot F}(= t^2_{\mathrm{YS}\cdot F}(\alpha))$, in Theorem 10, and the simulated coverage probabilities for the approximate values of $t^2_{\mathrm{YS\cdot L}}$, $t^2_{\mathrm{YS}\cdot F}$, and $\chi^2_{p,\alpha}$. It may be noted from Tables 5 and 6 that the simulated values and the approximate values are closer to the upper percentiles of the χ^2 distribution when $n_1^{(\ell)}$ becomes large. It is seen that $t^2_{\mathrm{YS}\cdot F}$ and $t^2_{\mathrm{YS\cdot L}}$ are considerably good approximate values of t^2_{simu} and $\widetilde{t}^2_{\mathrm{simu}}$, respectively, even for small samples.

In Table 7, we provide the simulated values, $t^2_{\mathrm{simu}}(= t^2_{\mathrm{simu}}(\alpha))$ and $\widetilde{t}^2_{\mathrm{simu}}(= \widetilde{t}^2_{\mathrm{simu}}(\alpha))$, and the approximate values, $t^2_{\mathrm{YS\cdot L}}(= t^2_{\mathrm{YS\cdot L}}(\alpha))$ and $t^2_{\mathrm{YS}\cdot F}(= t^2_{\mathrm{YS}\cdot F}(\alpha))$, for the case where the sample sizes are unequal. That is, computations are conducted for $(p_1, p_2) = (4,2), (8,4); \alpha = 0.05, 0.01$; and $\{(n_1^{(1)}, n_2^{(1)}), (n_1^{(2)}, n_2^{(2)})\} = \{(2m, m), (m, m)\}, \{(2m, 2m), (m, m)\}, \{(2m, m), (2m, m)\}$, $\{(2m, 2m), (2m, m)\}$, and $\{(2m, 2m), (2m, 2m)\}$, where $m = 15, 20$. In this case, it may be seen from the simulation results that $t^2_{\mathrm{YS}\cdot F}$ and $t^2_{\mathrm{YS\cdot L}}$ are considerably good approximate upper percentiles of the T^2 and $\widetilde{T}^2$ statistics, respectively. We note that these results for the case of unequal sample sizes are similar to the results for the case of equal sample sizes in Tables 5 and 6.

In conclusion, we have developed the usual Hotelling's T^2 statistic and its simplified Hotelling's T^2-type statistic for two-step monotone missing data under both one-sample and two-sample problems, though we have developed only the simplified Hotelling's T^2-type statistic for one-sample problem under three-step monotone missing data. Further, we have proposed the approximate upper percentiles of these statistics and our approximation procedures are considerably more accurate than the χ^2 approximation, even for small samples.

Table 1: Simulated and approximate values and coverage probabilities for one-sample problem with three-step monotone sample

Sample Size			Upper Percentile			Coverage Probability		
n_1	n_2	n_3	$\tilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L}$	$t^2_{{\rm YS}\cdot F}$	$\widetilde{\rm CP}_{\rm YS\cdot L}$	$\widetilde{\rm CP}_{{\rm YS}\cdot F}$	$\widetilde{\rm CP}_{\chi^2}$
				$(p_1, p_2, p_3) = (8, 4, 2)$				
				$\alpha = 0.05$				
20	10	10	35.34	32.15	27.77	0.933	0.898	0.635
50	10	10	20.19	19.99	19.81	0.948	0.946	0.862
100	10	10	17.65	17.59	17.57	0.949	0.949	0.912
200	10	10	16.52	16.52	16.52	0.950	0.950	0.932
400	10	10	16.01	16.01	16.01	0.950	0.950	0.941
800	10	10	15.73	15.76	15.76	0.950	0.950	0.946
20	10	20	35.24	32.20	26.14	0.934	0.881	0.637
50	10	20	20.16	19.90	19.59	0.947	0.943	0.862
100	10	20	17.62	17.56	17.52	0.949	0.949	0.912
200	10	20	16.54	16.51	16.50	0.950	0.949	0.932
400	10	20	16.00	16.00	16.00	0.950	0.950	0.942
800	10	20	15.76	15.76	15.76	0.950	0.950	0.946
20	10	30	35.12	32.23	24.89	0.935	0.866	0.638
50	10	30	20.13	19.83	19.39	0.947	0.941	0.863
100	10	30	17.62	17.53	17.47	0.949	0.948	0.912
200	10	30	16.53	16.50	16.49	0.950	0.949	0.932
400	10	30	16.01	16.00	16.00	0.950	0.950	0.941
800	10	30	15.76	15.75	15.75	0.950	0.950	0.946
20	20	10	34.63	30.76	24.89	0.928	0.870	0.645
50	20	10	20.02	19.71	19.39	0.946	0.942	0.865
100	20	10	17.59	17.51	17.47	0.949	0.948	0.913
200	20	10	16.49	16.50	16.49	0.950	0.950	0.933
400	20	10	16.02	16.00	16.00	0.950	0.950	0.941
800	20	10	15.76	15.75	15.75	0.950	0.950	0.946
20	20	20	34.52	31.08	23.91	0.931	0.858	0.647
50	20	20	20.03	19.67	19.21	0.946	0.940	0.865
100	20	20	17.53	17.48	17.42	0.949	0.948	0.913
200	20	20	16.50	16.49	16.48	0.950	0.950	0.933
400	20	20	16.01	16.00	16.00	0.950	0.950	0.941
800	20	20	15.74	15.75	15.75	0.950	0.950	0.946
20	20	30	34.50	31.30	23.11	0.933	0.846	0.647
50	20	30	19.94	19.63	19.04	0.946	0.939	0.867
100	20	30	17.55	17.46	17.38	0.949	0.948	0.913
200	20	30	16.47	16.48	16.47	0.950	0.950	0.933
400	20	30	16.00	16.00	15.99	0.950	0.950	0.942
800	20	30	15.77	15.75	15.75	0.950	0.950	0.946
20	30	10	34.19	29.94	23.11	0.925	0.849	0.651
50	30	10	19.90	19.50	19.04	0.945	0.939	0.868
100	30	10	17.53	17.44	17.38	0.949	0.948	0.913
200	30	10	16.48	16.48	16.47	0.950	0.950	0.933
400	30	10	15.99	16.00	15.99	0.950	0.950	0.942
800	30	10	15.76	15.75	15.75	0.950	0.950	0.946
20	30	20	34.18	30.35	22.44	0.928	0.838	0.652
50	30	20	19.85	19.49	18.89	0.946	0.938	0.868
100	30	20	17.53	17.42	17.33	0.949	0.947	0.914
200	30	20	16.50	16.47	16.46	0.949	0.949	0.933
400	30	20	16.00	15.99	15.99	0.950	0.950	0.942
800	30	20	15.75	15.75	15.75	0.950	0.950	0.946
20	30	30	34.07	30.64	21.89	0.931	0.829	0.654
50	30	30	19.79	19.47	18.75	0.946	0.937	0.869
100	30	30	17.50	17.41	17.29	0.949	0.947	0.914
200	30	30	16.52	16.46	16.45	0.949	0.949	0.932
400	30	30	16.00	15.99	15.99	0.950	0.950	0.941
800	30	30	15.73	15.75	15.75	0.950	0.950	0.946

Note. $\chi^2_{8,0.05} = 15.51$, $\widetilde{\rm CP}_{\rm YS\cdot L} = \widetilde{\rm CP}(t^2_{\rm YS\cdot L}(\alpha))$, $\widetilde{\rm CP}_{{\rm YS}\cdot F} = \widetilde{\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))$, $\widetilde{\rm CP}_{\chi^2} = \widetilde{\rm CP}(\chi^2_{p,\alpha})$.

Table 2: Simulated and approximate values and coverage probabilities for one-sample problem with three-step monotone sample

Sample Size			Upper Percentile			Coverage Probability		
n_1	n_2	n_3	$\widetilde{t}^2_{\text{simu}}$	$t^2_{\text{YS}\cdot\text{L}}$	$t^2_{\text{YS}\cdot F}$	$\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}}$	$\widetilde{\text{CP}}_{\text{YS}\cdot F}$	$\widetilde{\text{CP}}_{\chi^2}$
				$(p_1, p_2, p_3) = (8, 4, 2)$				
				$\alpha = 0.01$				
20	10	10	55.66	49.22	40.58	0.984	0.968	0.775
50	10	10	27.47	27.19	26.89	0.989	0.989	0.949
100	10	10	23.39	23.33	23.30	0.990	0.990	0.974
200	10	10	21.63	21.65	21.65	0.990	0.990	0.984
400	10	10	20.84	20.86	20.86	0.990	0.990	0.987
800	10	10	20.43	20.47	20.47	0.990	0.990	0.989
20	10	20	55.58	49.42	37.67	0.984	0.960	0.777
50	10	20	27.47	27.05	26.53	0.989	0.988	0.949
100	10	20	23.34	23.27	23.21	0.990	0.990	0.975
200	10	20	21.70	21.63	21.63	0.990	0.990	0.983
400	10	20	20.83	20.85	20.85	0.990	0.990	0.987
800	10	20	20.48	20.47	20.47	0.990	0.990	0.988
20	10	30	55.20	49.54	35.48	0.985	0.952	0.778
50	10	30	27.30	26.94	26.21	0.989	0.987	0.949
100	10	30	23.36	23.23	23.13	0.990	0.989	0.975
200	10	30	21.65	21.62	21.61	0.990	0.990	0.983
400	10	30	20.84	20.85	20.85	0.990	0.990	0.987
800	10	30	20.44	20.47	20.47	0.990	0.990	0.989
20	20	10	54.57	46.77	35.48	0.982	0.954	0.784
50	20	10	27.18	26.74	26.21	0.989	0.987	0.951
100	20	10	23.29	23.20	23.13	0.990	0.990	0.975
200	20	10	21.58	21.62	21.61	0.990	0.990	0.984
400	20	10	20.82	20.85	20.85	0.990	0.990	0.987
800	20	10	20.45	20.47	20.47	0.990	0.990	0.989
20	20	20	54.56	47.44	33.76	0.983	0.946	0.786
50	20	20	27.16	26.68	25.92	0.989	0.987	0.951
100	20	20	23.21	23.16	23.06	0.990	0.990	0.976
200	20	20	21.66	21.60	21.59	0.990	0.990	0.984
400	20	20	20.88	20.84	20.84	0.990	0.990	0.987
800	20	20	20.51	20.47	20.47	0.990	0.990	0.988
20	20	30	54.80	47.89	32.39	0.984	0.939	0.786
50	20	30	27.12	26.62	25.65	0.989	0.986	0.952
100	20	30	23.20	23.13	22.99	0.990	0.989	0.975
200	20	30	21.54	21.59	21.57	0.990	0.990	0.984
400	20	30	20.81	20.84	20.84	0.990	0.990	0.987
800	20	30	20.51	20.47	20.47	0.990	0.990	0.988
20	30	10	54.09	45.35	32.39	0.981	0.941	0.790
50	30	10	26.93	26.41	25.65	0.989	0.987	0.952
100	30	10	23.20	23.09	22.99	0.990	0.989	0.976
200	30	10	21.55	21.58	21.57	0.990	0.990	0.984
400	30	10	20.81	20.84	20.84	0.990	0.990	0.987
800	30	10	20.42	20.47	20.47	0.990	0.990	0.989
20	30	20	54.11	46.16	31.27	0.982	0.934	0.790
50	30	20	26.89	26.39	25.40	0.989	0.986	0.953
100	30	20	23.16	23.06	22.92	0.990	0.989	0.976
200	30	20	21.64	21.57	21.55	0.990	0.990	0.984
400	30	20	20.87	20.84	20.83	0.990	0.990	0.987
800	30	20	20.40	20.47	20.46	0.990	0.990	0.989
20	30	30	53.86	46.72	30.33	0.983	0.929	0.792
50	30	30	26.78	26.37	25.18	0.989	0.986	0.953
100	30	30	23.10	23.04	22.86	0.990	0.989	0.976
200	30	30	21.59	21.56	21.53	0.990	0.990	0.984
400	30	30	20.83	20.83	20.83	0.990	0.990	0.987
800	30	30	20.45	20.46	20.46	0.990	0.990	0.989

Note. $\chi^2_{8,0.01} = 20.09$, $\widetilde{\text{CP}}_{\text{YS}\cdot\text{L}} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot\text{L}}(\alpha))$, $\widetilde{\text{CP}}_{\text{YS}\cdot F} = \widetilde{\text{CP}}(t^2_{\text{YS}\cdot F}(\alpha))$, $\widetilde{\text{CP}}_{\chi^2} = \widetilde{\text{CP}}(\chi^2_{p,\alpha})$.

Table 3: Simulated and approximate values and coverage probabilities for one-sample problem with two-step monotone sample

Sample Size		Upper Percentile				Coverage Probability					
n_1	n_2	$t^2_{\rm simu}$	$\widetilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L}$	$t^2_{{\rm YS}\cdot F}$	${\rm CP}_{\rm YS\cdot L}$	${\rm CP}_{{\rm YS}\cdot F}$	${\rm CP}_{\chi^2}$	$\widetilde{\rm CP}_{\rm YS\cdot L}$	$\widetilde{\rm CP}_{{\rm YS}\cdot F}$	$\widetilde{\rm CP}_{\chi^2}$
						$(p_1, p_2) = (4, 2)$					
						$\alpha = 0.05$					
10	10	23.88	27.04	22.63	18.31	0.944	0.914	0.736	0.928	0.893	0.699
20	10	13.91	14.26	13.85	13.52	0.949	0.946	0.860	0.945	0.941	0.853
30	10	12.20	12.33	12.18	12.10	0.950	0.949	0.893	0.948	0.947	0.890
50	10	11.03	11.07	11.03	11.01	0.950	0.950	0.918	0.949	0.949	0.917
100	10	10.25	10.26	10.24	10.24	0.950	0.950	0.934	0.950	0.950	0.934
10	20	22.33	26.39	21.44	15.03	0.945	0.889	0.757	0.924	0.855	0.708
20	20	13.41	13.93	13.37	12.66	0.950	0.940	0.869	0.943	0.933	0.858
30	20	11.96	12.17	11.92	11.70	0.949	0.946	0.898	0.946	0.943	0.894
50	20	10.92	10.99	10.92	10.87	0.950	0.949	0.920	0.949	0.948	0.918
100	20	10.21	10.23	10.21	10.20	0.950	0.950	0.935	0.950	0.949	0.934
10	50	20.98	25.87	20.55	12.10	0.948	0.852	0.777	0.921	0.802	0.717
20	50	12.84	13.59	12.84	11.41	0.950	0.928	0.882	0.940	0.917	0.866
30	50	11.55	11.87	11.58	11.01	0.950	0.941	0.906	0.945	0.935	0.899
50	50	10.72	10.85	10.73	10.57	0.950	0.947	0.924	0.948	0.945	0.921
100	50	10.12	10.16	10.14	10.11	0.950	0.950	0.937	0.950	0.949	0.936
10	100	20.41	25.64	20.21	10.87	0.949	0.830	0.786	0.920	0.771	0.722
20	100	12.55	13.41	12.59	10.65	0.950	0.918	0.888	0.939	0.903	0.869
30	100	11.34	11.74	11.37	10.49	0.950	0.935	0.911	0.944	0.927	0.902
50	100	10.57	10.74	10.60	10.28	0.950	0.945	0.928	0.948	0.942	0.924
100	100	10.06	10.11	10.07	10.00	0.950	0.949	0.938	0.949	0.948	0.937
						$\alpha = 0.01$					
10	10	47.63	54.84	42.43	30.92	0.987	0.972	0.844	0.982	0.962	0.814
20	10	21.73	22.33	21.49	20.80	0.990	0.988	0.942	0.988	0.987	0.938
30	10	18.25	18.47	18.20	18.03	0.990	0.989	0.963	0.989	0.989	0.961
50	10	16.03	16.09	16.04	16.01	0.990	0.990	0.976	0.990	0.990	0.975
100	10	14.62	14.63	14.60	14.60	0.990	0.990	0.984	0.990	0.990	0.983
10	20	44.89	54.09	40.05	23.86	0.987	0.957	0.860	0.980	0.938	0.822
20	20	20.84	21.76	20.58	19.11	0.989	0.986	0.948	0.987	0.983	0.942
30	20	17.76	18.11	17.72	17.29	0.990	0.989	0.966	0.989	0.988	0.963
50	20	15.80	15.91	15.84	15.75	0.990	0.990	0.977	0.990	0.989	0.976
100	20	14.55	14.58	14.55	14.53	0.990	0.990	0.984	0.990	0.990	0.984
10	50	42.61	53.67	38.37	18.03	0.987	0.932	0.875	0.978	0.899	0.829
20	50	19.80	21.10	19.61	16.75	0.990	0.980	0.955	0.986	0.976	0.946
30	50	17.14	17.67	17.08	16.01	0.990	0.986	0.970	0.988	0.984	0.967
50	50	15.47	15.66	15.50	15.19	0.990	0.989	0.979	0.989	0.988	0.978
100	50	14.41	14.46	14.43	14.37	0.990	0.990	0.985	0.990	0.990	0.984
10	100	41.64	53.41	37.76	15.75	0.987	0.915	0.881	0.978	0.874	0.832
20	100	19.37	20.88	19.14	15.35	0.990	0.975	0.958	0.986	0.968	0.949
30	100	16.67	17.32	16.72	15.06	0.990	0.984	0.972	0.988	0.981	0.968
50	100	15.23	15.49	15.26	14.67	0.990	0.988	0.981	0.989	0.987	0.979
100	100	14.28	14.37	14.31	14.18	0.990	0.990	0.985	0.990	0.989	0.985

Note. $\chi^2_{4,0.05} = 9.49$, $\chi^2_{4,0.01} = 13.28$, ${\rm CP}_{\rm YS\cdot L} = {\rm CP}(t^2_{\rm YS\cdot L}(\alpha))$, ${\rm CP}_{{\rm YS}\cdot F} = {\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))$, ${\rm CP}_{\chi^2} = {\rm CP}(\chi^2_{p,\alpha})$, $\widetilde{\rm CP}_{\rm YS\cdot L} = \widetilde{\rm CP}(t^2_{\rm YS\cdot L}(\alpha))$, $\widetilde{\rm CP}_{{\rm YS}\cdot F} = \widetilde{\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))$, $\widetilde{\rm CP}_{\chi^2} = \widetilde{\rm CP}(\chi^2_{p,\alpha})$.

Table 4: Simulated and approximate values and coverage probabilities for one-sample problem with two-step monotone sample

Sample Size		Upper Percentile				Coverage Probability					
n_1	n_2	t^2_{simu}	$\widetilde{t}^2_{\mathrm{simu}}$	$t^2_{\mathrm{YS\cdot L}}$	$t^2_{\mathrm{YS}\cdot F}$	$\mathrm{CP}_{\mathrm{YS\cdot L}}$	$\mathrm{CP}_{\mathrm{YS}\cdot F}$	CP_{χ^2}	$\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L}}$	$\widetilde{\mathrm{CP}}_{\mathrm{YS}\cdot F}$	$\widetilde{\mathrm{CP}}_{\chi^2}$
					$(p_1, p_2) = (8, 4)$						
					$\alpha = 0.05$						
10	10	509.02	760.63	406.41	63.87	0.938	0.678	0.227	0.910	0.579	0.150
20	10	33.30	35.62	32.06	29.98	0.943	0.930	0.666	0.931	0.915	0.632
30	10	24.55	25.20	24.29	23.91	0.948	0.945	0.784	0.942	0.939	0.771
50	10	20.16	20.33	20.12	20.06	0.950	0.949	0.863	0.948	0.947	0.860
100	10	17.63	17.66	17.63	17.63	0.950	0.950	0.912	0.950	0.950	0.911
10	20	463.70	768.86	400.49	37.98	0.943	0.562	0.258	0.908	0.422	0.155
20	20	31.43	34.81	30.21	26.14	0.943	0.912	0.693	0.923	0.885	0.643
30	20	23.69	24.71	23.40	22.44	0.948	0.939	0.800	0.938	0.928	0.779
50	20	19.80	20.10	19.77	19.59	0.950	0.948	0.871	0.946	0.944	0.864
100	20	17.56	17.62	17.54	17.52	0.950	0.949	0.914	0.949	0.949	0.912
10	50	418.50	763.45	397.22	23.91	0.947	0.443	0.289	0.908	0.281	0.159
20	50	29.25	33.88	28.44	21.41	0.945	0.873	0.727	0.915	0.822	0.657
30	50	22.50	24.07	22.27	20.06	0.948	0.923	0.823	0.933	0.903	0.792
50	50	19.20	19.70	19.20	18.63	0.950	0.943	0.882	0.944	0.936	0.871
100	50	17.34	17.47	17.34	17.25	0.950	0.949	0.917	0.948	0.947	0.915
10	100	400.95	761.26	396.18	19.59	0.949	0.386	0.303	0.908	0.224	0.162
20	100	28.17	33.42	27.66	18.89	0.947	0.839	0.745	0.912	0.772	0.665
30	100	21.71	23.58	21.66	18.40	0.950	0.906	0.837	0.931	0.878	0.799
50	100	18.77	19.45	18.79	17.74	0.950	0.936	0.890	0.942	0.926	0.876
100	100	17.13	17.33	17.15	16.94	0.950	0.947	0.921	0.948	0.944	0.917
					$\alpha = 0.01$						
10	10	2673.93	4009.12	2017.48	117.26	0.987	0.806	0.311	0.980	0.733	0.217
20	10	52.35	56.39	48.83	44.60	0.987	0.982	0.800	0.983	0.976	0.772
30	10	35.02	36.03	34.47	33.76	0.989	0.988	0.897	0.987	0.986	0.889
50	10	27.44	27.68	27.40	27.30	0.990	0.990	0.949	0.989	0.989	0.947
100	10	23.43	23.48	23.39	23.38	0.990	0.990	0.975	0.990	0.990	0.974
10	20	2364.88	3937.27	2006.32	59.99	0.988	0.689	0.346	0.980	0.563	0.223
20	20	49.03	54.89	45.63	37.67	0.987	0.973	0.823	0.980	0.961	0.783
30	20	33.58	35.16	32.97	31.27	0.989	0.986	0.908	0.986	0.982	0.895
50	20	26.93	27.35	26.84	26.53	0.990	0.989	0.953	0.989	0.988	0.950
100	20	23.26	23.35	23.25	23.21	0.990	0.990	0.975	0.990	0.990	0.975
10	50	2150.78	3938.12	2000.75	33.76	0.989	0.558	0.381	0.980	0.390	0.228
20	50	45.65	53.67	42.70	29.54	0.987	0.952	0.849	0.977	0.924	0.794
30	50	31.83	34.29	31.14	27.30	0.989	0.979	0.923	0.984	0.971	0.904
50	50	25.93	26.66	25.92	24.98	0.990	0.987	0.960	0.988	0.985	0.954
100	50	22.92	23.10	22.94	22.79	0.990	0.990	0.977	0.990	0.989	0.976
10	100	2076.76	3960.91	1999.09	26.53	0.990	0.492	0.395	0.980	0.315	0.231
20	100	44.02	53.10	41.46	25.40	0.987	0.931	0.863	0.975	0.889	0.801
30	100	30.64	33.61	30.17	24.61	0.989	0.971	0.932	0.983	0.958	0.909
50	100	25.19	26.16	25.28	23.57	0.990	0.985	0.964	0.988	0.981	0.957
100	100	22.56	22.84	22.64	22.30	0.990	0.989	0.979	0.989	0.988	0.977

Note. $\chi^2_{8,0.05} = 15.51$, $\chi^2_{8,0.01} = 20.09$, $\mathrm{CP}_{\mathrm{YS\cdot L}} = \mathrm{CP}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\mathrm{CP}_{\mathrm{YS}\cdot F} = \mathrm{CP}(t^2_{\mathrm{YS}\cdot F}(\alpha))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_{p,\alpha})$, $\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L}} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\widetilde{\mathrm{CP}}_{\mathrm{YS}\cdot F} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS}\cdot F}(\alpha))$, $\widetilde{\mathrm{CP}}_{\chi^2} = \widetilde{\mathrm{CP}}(\chi^2_{p,\alpha})$.

Table 5: Simulated and approximate values and coverage probabilities for two-sample problem with two-step monotone sample

Sample Size		Upper Percentile				Coverage Probability					
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$t^2_{\rm simu}$	$\tilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L}$	$t^2_{{\rm YS}\cdot F}$	${\rm CP_{YS\cdot L}}$	${\rm CP}_{{\rm YS}\cdot F}$	${\rm CP}_{\chi^2}$	$\widetilde{\rm CP}_{\rm YS\cdot L}$	$\widetilde{\rm CP}_{{\rm YS}\cdot F}$	$\widetilde{\rm CP}_{\chi^2}$
						$(p_1, p_2) = (4, 2)$					
						$\alpha = 0.05$					
10	10	13.34	14.69	14.19	13.24	0.959	0.949	0.872	0.944	0.932	0.843
20	10	11.41	11.67	11.58	11.46	0.953	0.951	0.909	0.949	0.947	0.903
30	10	10.82	10.93	10.86	10.83	0.951	0.950	0.922	0.949	0.948	0.919
50	10	10.30	10.34	10.31	10.30	0.950	0.950	0.933	0.949	0.949	0.932
100	10	9.89	9.90	9.90	9.90	0.950	0.950	0.942	0.950	0.950	0.941
10	20	12.56	14.27	13.69	12.08	0.962	0.943	0.888	0.943	0.918	0.851
20	20	11.15	11.54	11.36	11.08	0.953	0.949	0.915	0.947	0.943	0.906
30	20	10.64	10.81	10.74	10.64	0.952	0.950	0.926	0.949	0.947	0.922
50	20	10.20	10.27	10.25	10.23	0.951	0.951	0.935	0.950	0.949	0.933
100	20	9.88	9.89	9.88	9.88	0.950	0.950	0.942	0.950	0.950	0.942
10	50	11.88	13.92	13.26	10.83	0.966	0.933	0.902	0.942	0.898	0.859
20	50	10.75	11.29	11.10	10.50	0.955	0.945	0.924	0.947	0.936	0.911
30	50	10.37	10.64	10.56	10.30	0.953	0.949	0.932	0.949	0.944	0.926
50	50	10.08	10.20	10.15	10.07	0.951	0.950	0.937	0.949	0.948	0.935
100	50	9.82	9.85	9.85	9.83	0.951	0.950	0.943	0.950	0.950	0.942
10	100	11.55	13.72	13.09	10.23	0.967	0.927	0.909	0.942	0.886	0.863
20	100	10.58	11.19	10.96	10.12	0.956	0.941	0.927	0.946	0.929	0.914
30	100	10.24	10.56	10.45	10.03	0.954	0.946	0.934	0.948	0.940	0.927
50	100	9.98	10.13	10.08	9.92	0.952	0.949	0.940	0.949	0.946	0.936
100	100	9.79	9.84	9.81	9.77	0.950	0.950	0.944	0.949	0.949	0.943
						$\alpha = 0.01$					
10	10	20.79	23.18	21.98	20.05	0.992	0.988	0.949	0.988	0.982	0.933
20	10	16.73	17.14	16.96	16.74	0.991	0.990	0.972	0.989	0.989	0.969
30	10	15.56	15.73	15.67	15.60	0.990	0.990	0.978	0.990	0.990	0.977
50	10	14.65	14.71	14.69	14.68	0.990	0.990	0.983	0.990	0.990	0.983
100	10	13.96	13.97	13.98	13.98	0.990	0.990	0.987	0.990	0.990	0.987
10	20	19.39	22.45	21.08	17.87	0.993	0.986	0.958	0.987	0.976	0.938
20	20	16.23	16.81	16.57	16.06	0.991	0.989	0.975	0.989	0.987	0.971
30	20	15.28	15.54	15.45	15.28	0.991	0.990	0.980	0.990	0.989	0.979
50	20	14.51	14.61	14.60	14.55	0.990	0.990	0.984	0.990	0.990	0.984
100	20	13.98	14.01	13.95	13.95	0.990	0.990	0.987	0.990	0.990	0.987
10	50	18.21	21.94	20.33	15.60	0.994	0.981	0.966	0.987	0.965	0.942
20	50	15.57	16.41	16.11	15.03	0.992	0.988	0.979	0.989	0.985	0.973
30	50	14.87	15.28	15.14	14.68	0.991	0.989	0.982	0.989	0.988	0.980
50	50	14.33	14.49	14.43	14.28	0.990	0.990	0.985	0.990	0.989	0.984
100	50	13.83	13.88	13.89	13.86	0.990	0.990	0.987	0.990	0.990	0.987
10	100	17.68	21.62	20.02	14.55	0.994	0.978	0.969	0.987	0.958	0.945
20	100	15.23	16.20	15.88	14.36	0.992	0.987	0.981	0.989	0.982	0.975
30	100	14.58	15.08	14.95	14.21	0.991	0.989	0.984	0.990	0.986	0.981
50	100	14.11	14.33	14.30	14.02	0.991	0.990	0.986	0.990	0.989	0.985
100	100	13.79	13.87	13.83	13.76	0.990	0.990	0.988	0.990	0.990	0.987

Note. $\chi^2_{4,0.05} = 9.49$, $\chi^2_{4,0.01} = 13.28$, ${\rm CP_{YS\cdot L}} = {\rm CP}(t^2_{\rm YS\cdot L}(\alpha))$, ${\rm CP}_{{\rm YS}\cdot F} = {\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))$, ${\rm CP}_{\chi^2} = {\rm CP}(\chi^2_{p,\alpha})$, $\widetilde{\rm CP}_{\rm YS\cdot L} = \widetilde{\rm CP}(t^2_{\rm YS\cdot L}(\alpha))$, $\widetilde{\rm CP}_{{\rm YS}\cdot F} = \widetilde{\rm CP}(t^2_{{\rm YS}\cdot F}(\alpha))$, $\widetilde{\rm CP}_{\chi^2} = \widetilde{\rm CP}(\chi^2_{p,\alpha})$.

Table 6: Simulated and approximate values and coverage probabilities for two-sample problem with two-step monotone sample

Sample Size		Upper Percentile				Coverage Probability					
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$t^2_{\rm simu}$	$\widetilde{t}^2_{\rm simu}$	$t^2_{\rm YS\cdot L}$	$t^2_{{\rm YS}\cdot F}$	$\rm CP_{YS\cdot L}$	$\mathrm{CP}_{\mathrm{YS}\cdot F}$	CP_{χ^2}	$\widetilde{\rm CP}_{\rm YS\cdot L}$	$\widetilde{\mathrm{CP}}_{\mathrm{YS}\cdot F}$	$\widetilde{\mathrm{CP}}_{\chi^2}$
						$(p_1, p_2) = (8, 4)$					
						$\alpha = 0.05$					
10	10	27.29	36.75	33.08	27.66	0.975	0.952	0.763	0.932	0.888	0.625
20	10	21.02	22.21	21.64	21.21	0.956	0.952	0.848	0.944	0.939	0.823
30	10	19.13	19.58	19.34	19.23	0.952	0.951	0.883	0.947	0.946	0.874
50	10	17.65	17.79	17.72	17.70	0.951	0.951	0.912	0.949	0.949	0.909
100	10	16.59	16.62	16.59	16.59	0.950	0.950	0.931	0.950	0.949	0.931
10	20	24.12	35.17	31.44	23.28	0.981	0.944	0.816	0.930	0.845	0.648
20	20	19.98	21.68	20.98	20.01	0.960	0.950	0.868	0.943	0.930	0.832
30	20	18.63	19.33	18.98	18.67	0.954	0.951	0.893	0.946	0.942	0.878
50	20	17.40	17.64	17.57	17.50	0.952	0.951	0.916	0.949	0.948	0.911
100	20	16.50	16.56	16.55	16.54	0.951	0.950	0.933	0.950	0.950	0.931
10	50	21.51	33.86	30.19	19.23	0.986	0.927	0.859	0.930	0.782	0.670
20	50	18.87	21.16	20.23	18.27	0.964	0.942	0.890	0.939	0.910	0.843
30	50	17.86	18.90	18.48	17.70	0.958	0.948	0.908	0.945	0.933	0.886
50	50	17.04	17.46	17.30	17.06	0.953	0.950	0.923	0.948	0.944	0.914
100	50	16.36	16.48	16.45	16.41	0.951	0.951	0.935	0.950	0.949	0.933
10	100	20.44	33.22	29.71	17.50	0.988	0.915	0.876	0.930	0.744	0.678
20	100	18.25	20.84	19.86	17.18	0.967	0.935	0.902	0.939	0.893	0.849
30	100	17.43	18.69	18.19	16.96	0.959	0.943	0.916	0.943	0.923	0.890
50	100	16.77	17.31	17.10	16.65	0.955	0.948	0.928	0.947	0.940	0.917
100	100	16.24	16.42	16.35	16.25	0.952	0.950	0.937	0.949	0.947	0.934
						$\alpha = 0.01$					
10	10	42.28	58.83	50.76	40.06	0.995	0.988	0.874	0.983	0.962	0.765
20	10	28.94	30.71	29.81	29.08	0.992	0.990	0.939	0.988	0.986	0.925
30	10	25.80	26.44	26.06	25.87	0.991	0.990	0.960	0.989	0.989	0.956
50	10	23.40	23.59	23.50	23.46	0.990	0.990	0.975	0.990	0.990	0.973
100	10	21.70	21.74	21.74	21.74	0.990	0.990	0.983	0.990	0.990	0.983
10	20	36.90	56.49	48.07	32.50	0.997	0.983	0.910	0.982	0.937	0.784
20	20	27.42	29.93	28.75	27.13	0.992	0.989	0.951	0.988	0.983	0.931
30	20	25.01	26.00	25.49	24.99	0.991	0.990	0.965	0.989	0.987	0.958
50	20	23.02	23.34	23.25	23.14	0.991	0.990	0.977	0.990	0.989	0.975
100	20	21.63	21.71	21.67	21.66	0.990	0.990	0.983	0.990	0.990	0.983
10	50	32.83	54.82	46.08	25.87	0.998	0.975	0.937	0.981	0.893	0.802
20	50	25.64	29.11	27.58	24.35	0.994	0.986	0.963	0.986	0.974	0.938
30	50	23.81	25.33	24.71	23.46	0.992	0.989	0.973	0.988	0.984	0.963
50	50	22.51	23.07	22.84	22.47	0.991	0.990	0.979	0.989	0.988	0.976
100	50	21.40	21.55	21.52	21.46	0.990	0.990	0.985	0.990	0.990	0.984
10	100	30.97	53.73	45.33	23.14	0.998	0.968	0.947	0.982	0.863	0.809
20	100	24.72	28.62	27.01	22.66	0.994	0.983	0.969	0.986	0.966	0.942
30	100	23.17	24.98	24.26	22.30	0.993	0.987	0.976	0.988	0.980	0.965
50	100	22.02	22.78	22.53	21.83	0.991	0.989	0.982	0.989	0.987	0.978
100	100	21.19	21.43	21.38	21.22	0.991	0.990	0.986	0.990	0.989	0.984

Note. $\chi^2_{8,0.05} = 15.51$, $\chi^2_{8,0.01} = 20.09$, $\mathrm{CP}_{\mathrm{YS\cdot L}} = \mathrm{CP}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\mathrm{CP}_{\mathrm{YS}\cdot F} = \mathrm{CP}(t^2_{\mathrm{YS}\cdot F}(\alpha))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_{p,\alpha})$, $\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L}} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\widetilde{\mathrm{CP}}_{\mathrm{YS}\cdot F} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS}\cdot F}(\alpha))$, $\widetilde{\mathrm{CP}}_{\chi^2} = \widetilde{\mathrm{CP}}(\chi^2_{p,\alpha})$.

Table 7: Simulated and approximate values and coverage probabilities for two-sample problem with two-step monotone sample

Sample Size				Upper Percentile				Coverage Probability					
$n_1^{(1)}$	$n_2^{(1)}$	$n_1^{(2)}$	$n_2^{(2)}$	t^2_{simu}	$\widetilde{t}^2_{\mathrm{simu}}$	$t^2_{\mathrm{YS\cdot L}}$	$t^2_{\mathrm{YS\cdot}F}$	$\mathrm{CP}_{\mathrm{YS\cdot L}}$	$\mathrm{CP}_{\mathrm{YS\cdot}F}$	CP_{χ^2}	$\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L}}$	$\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot}F}$	$\widetilde{\mathrm{CP}}_{\chi^2}$
							$(p_1, p_2) = (4, 2)$						
							$\alpha = 0.05$						
30	15	15	15	10.90	11.31	11.23	11.08	0.955	0.953	0.920	0.949	0.946	0.911
30	30	15	15	10.79	11.22	11.11	10.88	0.955	0.951	0.923	0.948	0.944	0.913
30	15	30	15	10.70	10.85	10.79	10.73	0.951	0.950	0.924	0.949	0.948	0.921
30	30	30	15	10.63	10.81	10.71	10.60	0.951	0.950	0.926	0.948	0.946	0.922
30	30	30	30	10.52	10.73	10.66	10.50	0.952	0.950	0.928	0.949	0.946	0.923
40	20	20	20	10.54	10.81	10.74	10.64	0.953	0.952	0.928	0.949	0.947	0.922
40	40	20	20	10.45	10.74	10.66	10.50	0.954	0.951	0.930	0.949	0.946	0.923
40	20	40	20	10.37	10.47	10.43	10.39	0.951	0.950	0.931	0.949	0.949	0.929
40	40	40	20	10.31	10.44	10.38	10.30	0.951	0.950	0.932	0.949	0.947	0.930
40	40	40	40	10.24	10.39	10.34	10.23	0.952	0.950	0.934	0.949	0.947	0.931
							$\alpha = 0.01$						
30	15	15	15	15.84	16.46	16.33	16.06	0.991	0.991	0.977	0.990	0.989	0.973
30	30	15	15	15.65	16.31	16.13	15.70	0.991	0.990	0.978	0.989	0.988	0.974
30	15	30	15	15.39	15.60	15.54	15.43	0.991	0.990	0.979	0.990	0.989	0.978
30	30	30	15	15.27	15.55	15.41	15.21	0.990	0.990	0.980	0.990	0.989	0.978
30	30	30	30	15.08	15.39	15.31	15.03	0.991	0.990	0.981	0.990	0.989	0.979
40	20	20	20	15.16	15.57	15.45	15.28	0.991	0.990	0.981	0.990	0.989	0.978
40	40	20	20	14.99	15.42	15.31	15.03	0.991	0.990	0.982	0.990	0.989	0.979
40	20	40	20	14.84	15.00	14.91	14.83	0.990	0.990	0.983	0.990	0.989	0.982
40	40	40	20	14.76	14.94	14.82	14.68	0.990	0.990	0.983	0.990	0.989	0.982
40	40	40	40	14.66	14.89	14.74	14.55	0.990	0.990	0.984	0.989	0.989	0.982
							$(p_1, p_2) = (8, 4)$						
							$\alpha = 0.05$						
30	15	15	15	19.12	20.91	20.51	20.01	0.964	0.959	0.885	0.945	0.940	0.847
30	30	15	15	18.82	20.70	20.18	19.40	0.964	0.957	0.891	0.944	0.934	0.851
30	15	30	15	18.81	19.38	19.14	18.93	0.954	0.951	0.889	0.947	0.944	0.877
30	30	30	15	18.56	19.28	18.91	18.56	0.954	0.950	0.895	0.945	0.941	0.879
30	30	30	30	18.26	19.12	18.75	18.27	0.956	0.950	0.900	0.945	0.938	0.882
40	20	20	20	18.15	19.22	18.98	18.67	0.960	0.956	0.902	0.947	0.943	0.880
40	40	20	20	17.94	19.07	18.75	18.27	0.960	0.954	0.906	0.946	0.939	0.883
40	20	40	20	17.87	18.26	18.08	17.95	0.953	0.951	0.907	0.948	0.946	0.899
40	40	40	20	17.66	18.13	17.93	17.70	0.954	0.951	0.911	0.947	0.944	0.901
40	40	40	40	17.47	18.03	17.82	17.50	0.955	0.950	0.915	0.947	0.942	0.903
							$\alpha = 0.01$						
30	15	15	15	26.04	28.61	27.96	27.13	0.994	0.992	0.960	0.989	0.986	0.941
30	30	15	15	25.50	28.25	27.43	26.14	0.994	0.991	0.963	0.988	0.984	0.943
30	15	30	15	25.27	26.11	25.74	25.40	0.991	0.990	0.964	0.989	0.988	0.958
30	30	30	15	24.86	25.90	25.38	24.81	0.991	0.990	0.966	0.989	0.987	0.959
30	30	30	30	24.48	25.72	25.13	24.35	0.991	0.990	0.969	0.989	0.986	0.960
40	20	20	20	24.36	25.87	25.49	24.99	0.992	0.991	0.970	0.989	0.988	0.959
40	40	20	20	23.97	25.59	25.13	24.35	0.993	0.991	0.972	0.989	0.986	0.961
40	20	40	20	23.75	24.27	24.07	23.85	0.991	0.990	0.973	0.989	0.989	0.969
40	40	40	20	23.43	24.09	23.82	23.46	0.991	0.990	0.974	0.989	0.988	0.970
40	40	40	40	23.19	23.96	23.65	23.14	0.991	0.990	0.976	0.989	0.988	0.971

Note. $\chi^2_{4,0.05} = 9.49$, $\chi^2_{4,0.01} = 13.28$, $\chi^2_{8,0.05} = 15.51$, $\chi^2_{8,0.01} = 20.09$, $\mathrm{CP}_{\mathrm{YS\cdot L}} = \mathrm{CP}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\mathrm{CP}_{\mathrm{YS\cdot}F} = \mathrm{CP}(t^2_{\mathrm{YS\cdot}F}(\alpha))$, $\mathrm{CP}_{\chi^2} = \mathrm{CP}(\chi^2_{p,\alpha})$, $\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot L}} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS\cdot L}}(\alpha))$, $\widetilde{\mathrm{CP}}_{\mathrm{YS\cdot}F} = \widetilde{\mathrm{CP}}(t^2_{\mathrm{YS\cdot}F}(\alpha))$, $\widetilde{\mathrm{CP}}_{\chi^2} = \widetilde{\mathrm{CP}}(\chi^2_{p,\alpha})$.

Acknowledgments The authors would like to thank two referees for their helpful comments and suggestions. The second author's research was partly supported by a Grant-in-Aid for Scientific Research (C) (26330050).

References

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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

[9] 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.

[10] 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.

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

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

[13] Yagi, A. and Seo, T. (2014a). 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.

[14] Yagi, A. and Seo, T. (2014b). Tests for two mean vectors and simultaneous confidence intervals for multiple comparisons with three-step monotone missing data. *Technical Report No.14–09, Statistical Research Group, Hiroshima University, Hiroshima, Japan.*

Ayaka Yagi

Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
1413620@ed.tus.ac.jp

Takashi Seo

Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
seo@rs.tus.ac.jp