SUT Journal of Mathematics
Vol. 52, No. 1 (2016), 21–39

# A test for subvector of mean vector with two-step monotone missing data

**Tamae Kawasaki and Takashi Seo**

(Received November 19, 2015; Revised May 19, 2016)

**Abstract.** In this paper, we consider the one-sample problem of testing for the subvector of a mean vector with two-step monotone missing data. In the case that the data set consists of complete data with $p(= p_1 + p_2 + p_3)$ dimensions and incomplete data with $(p_1 + p_2)$ dimensions, we derive the likelihood ratio criterion for testing the $(p_2 + p_3)$ mean vector under the given mean vector of $p_1$ dimensions. Furthermore, we propose an approximation for the upper percentile of the likelihood ratio test (LRT) statistic. We investigate the accuracy and asymptotic behavior of this approximation using Monte Carlo simulation. An example is presented in order to illustrate the method.

*AMS* 2010 *Mathematics Subject Classification.* 62H20, 62H10.

*Key words and phrases.* Likelihood ratio test, maximum likelihood estimators, Monte Carlo simulation, Rao's $U$ statistic.

## §1. Introduction

When analyzing data, it is important to consider missing observations. The existence of missing data is a common problem that is present in almost all statistical data analyses. However, the majority of statistical methods require a comparatively strict assumption concerning the cause of missing data, and are prone to substantial bias. Methods for dealing with missing data by removing incomplete cases or imputing missing values are more vulnerable to the propagation of bias throughout. Statistical analysis involving monotone missing data has been discussed by many authors, because in this case the mathematical complexity is reduced. For example, Anderson (1957) demonstrated an approach for deriving the maximum likelihood estimators (MLEs) of the mean vector and covariance matrix using the likelihood equations for monotone missing data. Kanda and Fujikoshi (1998) described the properties of MLEs based on two-step and three-step monotone missing samples and a general $k$-step. Among the many papers that propose methods for testing mean vectors with monotone missing data, we mention those by Krishnamoorthy and Pannala (1999); Yu, Krishnamoorthy, and Pannala (2006); and Chang and Richards (2009). In particular, for testing the mean vector with two-step monotone missing data, Seko, Yamazaki, and Seo (2012), and Seko, Kawasaki, and Seo (2011) have provided a simple approach to deriving the approximate upper percentiles of the Hotelling's $T^2$ type statistic and LRT statistic for one-sample and two-sample problems. Moreover, various statistical methods have been developed to analyze data with non-monotone missing values by Srivastava (1985), Srivastava and Carter (1986), and Shutoh, Kusumi, Morinaga, Yamada, and Seo (2010), and others. In the case of general $k$-step monotone missing data, many difficult problems remain unsolved. For simplicity, we assume that $k = 2$. Let the data set $\{x_{i,j}\}$ be of the form

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,p_1} & x_{1,p_1+1} & \cdots & x_{1,p_1+p_2} & x_{1,p_1+p_2+1} & \cdots & x_{1,p} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{n_1,1} & \cdots & x_{n_1,p_1} & x_{n_1,p_1+1} & \cdots & x_{n_1,p_1+p_2} & x_{n_1,p_1+p_2+1} & \cdots & x_{n_1,p} \\ x_{n_1+1,1} & \cdots & x_{n_1+1,p_1} & x_{n_1+1,p_1+1} & \cdots & x_{n_1+1,p_1+p_2} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{n,1} & \cdots & x_{n,p_1} & x_{n,p_1+1} & \cdots & x_{n,p_1+p_2} & * & \cdots & * \end{pmatrix},$$

where $n_2 = n - n_1$ and $n_1 > p$. Here, "$*$" indicates missing data. That is, we have complete data for $n_1$ mutually independent observations with $p$ dimensions, and incomplete data for $n_2$ mutually independent observations with $(p_1 + p_2)$ dimensions. Such a data set is described as two-step monotone missing data.

In this paper, based on two-step monotone missing data, we consider the one-sample problem of testing for the subvector of a mean vector. We derive the MLEs of the mean vector and the covariance matrix and the MLE of the covariance matrix under the null hypothesis. Using these MLEs, we propose the likelihood ratio test statistic and its approximate upper percentile.

In Section 2, we review the test for a subvector based on non-missing data when the first $p_1$ dimensions of the mean vector $\boldsymbol{\mu}$ is given. In Section 3, we derive the MLEs and the MLEs under the null hypothesis, with two-step monotone missing data. In Section 4, we propose the LRT statistic and its approximate upper percentiles. The accuracy of the approximate upper percentiles of the test statistic is investigated using Monte Carlo simulation in Section 5. In Section 6, we present a numerical example to illustrate our method using the approximate upper percentiles of the test statistic. Finally, Section 7 concludes this paper.

## §2. Test for a subvector

We review the case of non-missing data. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)'$ and $\Sigma$ are unknown. Let $\boldsymbol{\mu} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_{(23)}')'$, where $\boldsymbol{\mu}_1 = (\mu_1, \mu_2, \ldots, \mu_{p_1})'$ and $\boldsymbol{\mu}_{(23)} = (\mu_{p_1+1}, \mu_{p_1+2}, \ldots, \mu_p)'$, $p_1 < p < n$. Then, the sample mean vector and unbiased covariance matrix are defined as

$$\overline{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i = \begin{pmatrix} \overline{\boldsymbol{x}}_1 \\ \overline{\boldsymbol{x}}_{(23)} \end{pmatrix}, S = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})' = \begin{pmatrix} S_{11} & S_{1(23)} \\ S_{(23)1} & S_{(23)(23)} \end{pmatrix},$$

respectively, where $\overline{\boldsymbol{x}}_1$ is a $p_1$-vector and $S_{11}$ is a $p_1 \times p_1$ matrix. Consider the following hypothesis test problem for the case of two-step monotone missing data in the one-sample problem:

(1.1)

$$H_0 : \boldsymbol{\mu}_{(23)} = \boldsymbol{\mu}_{(23)0} \text{ given } \boldsymbol{\mu}_1 = \boldsymbol{\mu}_{10} \text{ vs. } H_1 : \boldsymbol{\mu}_{(23)} \neq \boldsymbol{\mu}_{(23)0} \text{ given } \boldsymbol{\mu}_1 = \boldsymbol{\mu}_{10},$$

where $\boldsymbol{\mu}_{(23)0}$ and $\boldsymbol{\mu}_{10}$ are known. A criterion that is equivalent to the likelihood ratio can be written as

$$U = \frac{T_p^2 - T_{p_1}^2}{n - 1 + T_{p_1}^2},$$

where $T_p^2 = n(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)'S^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)$ and $T_{p_1}^2 = n(\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_{10})'S_{11}^{-1}(\overline{\boldsymbol{x}}_1 - \boldsymbol{\mu}_{10})$. We note that $U = \lambda^{-2/n} - 1$, where $\lambda$ is the likelihood ratio criterion. Under $H_0$, it follows that $(n-p)U/(p-p_1)$ is distributed as an $F$ distribution with $p - p_1$ and $n - p$ degrees of freedom. This result follows from the one in Siotani, Hayakawa, and Fujikoshi (1985, p. 215). The criterion is called Rao's $U$ statistic (See, Rao (1949) and Giri (1964)). The situation in which a subvector of $\boldsymbol{\mu}$ can be known is not rare. In some situations, partial information concerning the population means may be available to the experimenter. Furthermore, this hypothesis in the two-sample problem is equivalent to a test for additional information. That is, a problem that is closely related to the testing of the mean vectors $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$ is determining whether $\boldsymbol{x}_{(23)} = (\boldsymbol{x}_2', \boldsymbol{x}_3')'$ has additional information in the presence of $\boldsymbol{x}_1$, where $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{x}_{(23)}')'$ arises from one of two groups $\Pi^{(1)} : N_p(\boldsymbol{\mu}^{(1)}, \Sigma)$ and $\Pi^{(2)} : N_p(\boldsymbol{\mu}^{(2)}, \Sigma)$. Eaton and Kariya (1975) derived tests for the independence of two normally distributed subvectors in the case that an additional random sample is available. Provost (1990) obtained explicit expressions in the case that the MLEs of all of the parameters of the multinormal random vector are given, and the likelihood ratio statistic for testing the independence between subvectors has been obtained. In the next section, we derive the MLEs and MLEs under $H_0$, with two-step monotone missing data, to obtain the LRT statistic.

## §3. MLEs with two-step monotone missing data

In this section, we obtain the MLEs using the decomposition of the density into conditional densities, which is called the conditional method (Kanda and Fujikoshi, 1998). Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n_1}$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$, and let $\boldsymbol{x}_{n_1+1}, \boldsymbol{x}_{n_1+2}, \ldots, \boldsymbol{x}_n$ be distributed as $N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}, \Sigma_{(12)(12)})$, where each $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,p})'$, $j = 1, 2, \ldots, n_1$ is $p \times 1$, each $\boldsymbol{x}_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,p_1+p_2})'$, $j = n_1+1, n_1+2, \ldots, n$ is $(p_1+p_2) \times 1$, and

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{10} \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{(12)} \\ \boldsymbol{\mu}_3 \end{pmatrix}, \quad \Sigma = \left(\begin{array}{cc|c} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \hline \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{array}\right) = \begin{pmatrix} \Sigma_{(12)(12)} & \Sigma_{(12)3} \\ \Sigma_{3(12)} & \Sigma_{33} \end{pmatrix}.$$

We partition $\boldsymbol{x}_j$ into a $p_1 \times 1$ random vector, a $p_2 \times 1$ random vector, and a $p_3 \times 1$ random vector as $\boldsymbol{x}_j = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j}, \boldsymbol{x}'_{3j})' = (\boldsymbol{x}'_{(12)j}, \boldsymbol{x}'_{3j})'$, where $\boldsymbol{x}_{ij}, i = 1, 2, 3, j = 1, 2, \ldots, n_1$ is $p_i \times 1$, and $p = p_1 + p_2 + p_3$. In addition, $\boldsymbol{x}_{(12)j}$ is partitioned into a $p_1 \times 1$ random vector and a $p_2 \times 1$ random vector as $\boldsymbol{x}_{(12)j} = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j})'$, where $\boldsymbol{x}_{ij}, i = 1, 2, j = n_1+1, n_1+2, \ldots, n$ is $p_i \times 1$. Then, the joint density function of the observed data set $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n_1}, \boldsymbol{x}_{(12)n_1+1}, \boldsymbol{x}_{(12)n_1+2}, \ldots, \boldsymbol{x}_{(12)n}$ can be written as

$$\prod_{j=1}^{n_1} f(\boldsymbol{x}_j; \boldsymbol{\mu}, \Sigma) \times \prod_{j=n_1+1}^{n} f(\boldsymbol{x}_{(12)j}; \boldsymbol{\mu}_{(12)}, \Sigma_{(12)(12)}),$$

where $f(\boldsymbol{x}_j; \boldsymbol{\mu}, \Sigma)$ and $f(\boldsymbol{x}_{(12)j}; \boldsymbol{\mu}_{(12)}, \Sigma_{(12)(12)})$ are the density functions of $N_p(\boldsymbol{\mu}, \Sigma)$ and $N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}, \Sigma_{(12)(12)})$, respectively. That is, the likelihood function is given by

$$L(\boldsymbol{\mu}, \Sigma) = \prod_{j=1}^{n_1} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}_j - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu})\right\}$$
$$\times \prod_{j=n_1+1}^{n} \frac{1}{(2\pi)^{(p_1+p_2)/2}|\Sigma_{(12)(12)}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)})'\Sigma_{(12)(12)}^{-1}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)})\right\}.$$

The sample mean vectors are defined as

$$\overline{\boldsymbol{x}}_{1T} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{x}_{1j}, \ \overline{\boldsymbol{x}}_{2T} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{x}_{2j},$$
$$\overline{\boldsymbol{x}}_F = (\overline{\boldsymbol{x}}'_{(12)F}, \overline{\boldsymbol{x}}'_{3F})' = \left(\frac{1}{n_1}\sum_{j=1}^{n_1} \boldsymbol{x}'_{(12)j}, \frac{1}{n_1}\sum_{j=1}^{n_1} \boldsymbol{x}'_{3j}\right)'.$$

In our situation, we first multiply the observation vectors $\boldsymbol{x}_j$ by the transformation matrix

$$\Gamma_1 = \left(\begin{array}{cc|c} I_{p_1} & O & \\ O & I_{p_2} & O \\ \hline \multicolumn{2}{c|}{-\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}} & I_{p_3} \end{array}\right)$$

on the left side, so that the transformed observation vectors are

$$\begin{aligned} &\boldsymbol{x}_{(12)j} \sim N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}, \Sigma_{(12)(12)}), \quad j = 1, 2, \ldots, n, \\ &\boldsymbol{x}_{3j} - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\boldsymbol{x}_{(12)j} \\ &\qquad \sim N_{p_3}(\boldsymbol{\mu}_3 - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\boldsymbol{\mu}_{(12)}, \Sigma_{33\cdot(12)}), \quad j = 1, 2, \ldots, n_1, \end{aligned}$$

where $\Sigma_{33\cdot(12)} = \Sigma_{33} - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}$. Next, we multiply the above observation vectors by the transformation matrix

$$\left(\begin{array}{cc|c} I_{p_1} & O & \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{p_2} & O \\ \hline \multicolumn{2}{c|}{O} & I_{p_3} \end{array}\right)$$

on the left side, so that the transformed observation vectors are

$$\begin{aligned} &\boldsymbol{x}_{1j} \sim N_{p_1}(\boldsymbol{\eta}_1, \Psi_{11}), \quad j = 1, 2, \ldots, n, \\ &\boldsymbol{x}_{2j} - \Psi_{21}\boldsymbol{x}_{1j} \sim N_{p_2}(\boldsymbol{\eta}_2, \Psi_{22}), \quad j = 1, 2, \ldots, n, \\ &\boldsymbol{x}_{3j} - \Psi_{3(12)}\boldsymbol{x}_{(12)j} \sim N_{p_3}(\boldsymbol{\eta}_3, \Psi_{33}), \quad j = 1, 2, \ldots, n_1, \end{aligned}$$

where

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{10} \\ \boldsymbol{\mu}_2 - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_{10} \\ \boldsymbol{\mu}_3 - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\boldsymbol{\mu}_{(12)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{10} \\ \boldsymbol{\mu}_2 - \Psi_{21}\boldsymbol{\mu}_{10} \\ \boldsymbol{\mu}_3 - \Psi_{3(12)}\boldsymbol{\mu}_{(12)} \end{pmatrix},$$

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

$$\Psi_{(12)(12)} = \begin{pmatrix} \Sigma_{11} & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22\cdot 1} \end{pmatrix}, \Psi_{3(12)} = \Psi'_{(12)3} = \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}, \Psi_{33} = \Sigma_{33\cdot(12)},$$

and $\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{21}$. It should be noted that $\boldsymbol{x}_{1j}$, $\boldsymbol{x}_{2j} - \Psi_{21}\boldsymbol{x}_{1j}$, and $\boldsymbol{x}_{3j} - \Psi_{3(12)}\boldsymbol{x}_{(12)j}$ are independent. Because $(\boldsymbol{\eta}, \Psi)$ has a one-to-one correspondence with $(\boldsymbol{\mu}, \Sigma)$, it is sufficient to derive the MLEs of $(\boldsymbol{\eta}, \Psi)$ instead of $(\boldsymbol{\mu}, \Sigma)$. Using the above transformation matrices, we will derive the MLEs of $\boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \Psi_{(12)(12)}, \Psi_{3(12)}$, and $\Psi_{33}$.

**Theorem 1.** *Suppose that the data set has a two-step monotone missing pattern. Then, the maximum likelihood estimators of* $\boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \Psi_{11}, \Psi_{21}, \Psi_{22}, \Psi_{3(12)}$, *and* $\Psi_{33}$ *are given by*

$$
\begin{aligned}
\widehat{\boldsymbol{\eta}}_2 &= \overline{\boldsymbol{x}}_{2T} - \widehat{\Psi}_{21}\overline{\boldsymbol{x}}_{1T},\ \widehat{\boldsymbol{\eta}}_3 = \overline{\boldsymbol{x}}_{3F} - \widehat{\Psi}_{3(12)}\overline{\boldsymbol{x}}_{(12)F},\\
\widehat{\Psi}_{11} &= \frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{x}_{1j} - \boldsymbol{\mu}_{10})(\boldsymbol{x}_{1j} - \boldsymbol{\mu}_{10})',\\
\widehat{\Psi}_{21} &= \left\{\sum_{j=1}^{n}(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{2T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})'\right\}\left\{\sum_{j=1}^{n}(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})'\right\}^{-1},\\
\widehat{\Psi}_{22} &= \frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{x}_{2j} - \widehat{\Psi}_{21}\boldsymbol{x}_{1j} - \widehat{\boldsymbol{\eta}}_2)(\boldsymbol{x}_{2j} - \widehat{\Psi}_{21}\boldsymbol{x}_{1j} - \widehat{\boldsymbol{\eta}}_2)',\\
\widehat{\Psi}_{3(12)} &= \left\{\sum_{j=1}^{n_1}(\boldsymbol{x}_{3j} - \overline{\boldsymbol{x}}_{3F})(\boldsymbol{x}_{(12)j} - \overline{\boldsymbol{x}}_{(12)F})'\right\}\\
&\quad \times\left\{\sum_{j=1}^{n_1}(\boldsymbol{x}_{(12)j} - \overline{\boldsymbol{x}}_{(12)F})(\boldsymbol{x}_{(12)j} - \overline{\boldsymbol{x}}_{(12)F})'\right\}^{-1},\\
\widehat{\Psi}_{33} &= \frac{1}{n_1}\sum_{j=1}^{n_1}(\boldsymbol{x}_{3j} - \widehat{\Psi}_{3(12)}\boldsymbol{x}_{(12)j} - \widehat{\boldsymbol{\eta}}_3)(\boldsymbol{x}_{3j} - \widehat{\Psi}_{3(12)}\boldsymbol{x}_{(12)j} - \widehat{\boldsymbol{\eta}}_3)',
\end{aligned}
$$

*respectively.*

*Proof.* The likelihood function for the parameters $\boldsymbol{\eta}$ and $\Psi$ can be written as

$$
\begin{aligned}
L(\boldsymbol{\eta}, \Psi) = &\prod_{j=1}^{n}\frac{1}{(2\pi)^{p_1/2}|\Psi_{11}|^{1/2}}\exp\left\{-\frac{1}{2}(\boldsymbol{x}_{1j} - \boldsymbol{\eta}_1)'\Psi_{11}^{-1}(\boldsymbol{x}_{1j} - \boldsymbol{\eta}_1)\right\}\\
&\times\prod_{j=1}^{n}\frac{1}{(2\pi)^{p_2/2}|\Psi_{22}|^{1/2}}\\
&\times\exp\left\{-\frac{1}{2}(\boldsymbol{x}_{2j} - \Psi_{21}\boldsymbol{x}_{1j} - \boldsymbol{\eta}_2)'\Psi_{22}^{-1}(\boldsymbol{x}_{2j} - \Psi_{21}\boldsymbol{x}_{1j} - \boldsymbol{\eta}_2)\right\}\\
&\times\prod_{j=1}^{n_1}\frac{1}{(2\pi)^{p_3/2}|\Psi_{33}|^{1/2}}\\
&\times\exp\left\{-\frac{1}{2}(\boldsymbol{x}_{3j} - \Psi_{3(12)}\boldsymbol{x}_{(12)j} - \boldsymbol{\eta}_3)'\Psi_{33}^{-1}(\boldsymbol{x}_{3j} - \Psi_{3(12)}\boldsymbol{x}_{(12)j} - \boldsymbol{\eta}_3)\right\}.
\end{aligned}
$$

Then, the partial derivative of $\log L(\boldsymbol{\eta}, \Psi)$ with respect to $\Psi_{11}$ (see Seber (1984, p.530)) is

$$\frac{\partial \log L(\boldsymbol{\eta}, \Psi)}{\partial \Psi_{11}} = -\frac{n}{2}\Psi_{11}^{-1} + \frac{1}{2}\sum_{j=1}^{n} \Psi_{11}^{-1}(\boldsymbol{x}_{1j} - \boldsymbol{\eta}_1)(\boldsymbol{x}_{1j} - \boldsymbol{\eta}_1)'\Psi_{11}^{-1}.$$

Thus, by solving $\partial \log L(\boldsymbol{\eta}, \Psi)/\partial \Psi_{11} = 0$ we obtain

$$\widehat{\Psi}_{11} = \frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{x}_{1j} - \boldsymbol{\mu}_{10})(\boldsymbol{x}_{1j} - \boldsymbol{\mu}_{10})'.$$

Similarly, the partial derivative of $\log L(\boldsymbol{\eta}, \Psi)$ with respect to $\Psi_{21}$ is

$$\begin{aligned}\frac{\partial \log L(\boldsymbol{\eta}, \Psi)}{\partial \Psi_{21}} =& \sum_{j=1}^{n}\{\Psi_{22}^{-1}(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{2T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})' \\ & - \Psi_{22}^{-1}\Psi_{21}(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})'\}.\end{aligned}$$

Thus, by solving $\partial \log L(\boldsymbol{\eta}, \Psi)/\partial \Psi_{21} = 0$ we obtain

$$\widehat{\Psi}_{21} = \left\{\sum_{j=1}^{n}(\boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{2T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})'\right\}\left\{\sum_{j=1}^{n}(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{1T})'\right\}^{-1}.$$

In the same manner as for $\Psi_{11}$ and $\Psi_{21}$, we solve the equations resulting from setting the partial derivative of $\log L(\boldsymbol{\eta}, \Psi)$ with respect to each of $\boldsymbol{\eta}_2$, $\boldsymbol{\eta}_3$, $\Psi_{22}$, $\Psi_{3(12)}$, and $\Psi_{33}$ to zero, and obtain the MLEs. □

Then, the MLEs of $\boldsymbol{\mu}_{(23)}$ and $\Sigma$ are expressed as

$$\widehat{\boldsymbol{\mu}}_{(23)} = \begin{pmatrix} \overline{\boldsymbol{x}}_{2T} - \widehat{\Sigma}_{21}\widehat{\Sigma}_{11}^{-1}(\overline{\boldsymbol{x}}_{1T} - \boldsymbol{\mu}_{10}) \\ \overline{\boldsymbol{x}}_{3T} - \widehat{\Sigma}_{3(12)}\widehat{\Sigma}_{(12)(12)}^{-1}\begin{pmatrix}\overline{\boldsymbol{x}}_{1F} - \boldsymbol{\mu}_{10} \\ \overline{\boldsymbol{x}}_{2F} - \widehat{\boldsymbol{\mu}}_2\end{pmatrix}\end{pmatrix},$$

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

where

$$\widehat{\Sigma}_{(12)(12)} = \begin{pmatrix}\widehat{\Psi}_{11} & \widehat{\Psi}_{11}\widehat{\Psi}_{12} \\ \widehat{\Psi}_{21}\widehat{\Psi}_{11} & \widehat{\Psi}_{22} + \widehat{\Psi}_{21}\widehat{\Psi}_{11}\widehat{\Psi}_{12}\end{pmatrix},\ \widehat{\Sigma}_{3(12)} = \widehat{\Sigma}'_{(12)3} = \widehat{\Psi}_{3(12)}\widehat{\Sigma}_{(12)(12)},$$

$$\widehat{\Sigma}_{33} = \widehat{\Psi}_{33} + \widehat{\Psi}_{3(12)}\widehat{\Psi}_{(12)(12)}\widehat{\Psi}_{(12)3}.$$

Next, we represent the MLEs under $H_0$ in order to obtain the LRT statistic. The null hypothesis in (1.1) can be written as $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ $(= (\boldsymbol{\mu}_{10}', \boldsymbol{\mu}_{20}', \boldsymbol{\mu}_{30}')' = (\boldsymbol{\mu}_{(12)0}', \boldsymbol{\mu}_{30}')')$. Let $\boldsymbol{x}_j = (\boldsymbol{x}_{(12)j}', \boldsymbol{x}_{3j}')'$ be distributed as $N_p(\boldsymbol{\mu}_0, \Sigma), j = 1, 2, \ldots, n_1$, and $\boldsymbol{x}_{(12)j}$ be distributed as $N_{p_1+p_2}(\boldsymbol{\mu}_{(12)0}, \Sigma_{(12)(12)})$, for $j = n_1 + 1, n_1 + 2, \ldots, n$. Then, the likelihood function is given by

$$
\begin{aligned}
L(\boldsymbol{\mu}_0, \Sigma) = &\prod_{j=1}^{n_1} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}_j - \boldsymbol{\mu}_0)'\Sigma^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu}_0)\right\} \\
&\times \prod_{j=n_1+1}^{n} \frac{1}{(2\pi)^{(p_1+p_2)/2}|\Sigma_{(12)(12)}|^{1/2}} \\
&\times \exp\left\{-\frac{1}{2}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})'\Sigma_{(12)(12)}^{-1}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})\right\}.
\end{aligned}
$$

By multiplying the observation vectors by $\Gamma_1$ on the left side, we obtain

$$
\begin{aligned}
&\boldsymbol{x}_{(12)j} \sim N_{p_1+p_2}(\boldsymbol{\xi}_{(12)}, \Phi_{(12)(12)}),\ j = 1, 2, \ldots, n, \\
&\boldsymbol{x}_{3j} - \Phi_{3(12)}\boldsymbol{x}_{(12)j} \sim N_{p_3}(\boldsymbol{\xi}_3, \Phi_{33}),\ j = 1, 2, \ldots, n_1,
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\xi} &= \begin{pmatrix} \boldsymbol{\xi}_{(12)} \\ \boldsymbol{\xi}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{(12)0} \\ \boldsymbol{\mu}_{30} - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\boldsymbol{\mu}_{(12)0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{(12)0} \\ \boldsymbol{\mu}_{30} - \Phi_{3(12)}\boldsymbol{\mu}_{(12)0} \end{pmatrix}, \\
\Phi &= \begin{pmatrix} \Phi_{(12)(12)} & \Phi_{(12)3} \\ \Phi_{3(12)} & \Phi_{33} \end{pmatrix} = \begin{pmatrix} \Sigma_{(12)(12)} & \Sigma_{(12)(12)}^{-1}\Sigma_{(12)3} \\ \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1} & \Sigma_{33\cdot(12)} \end{pmatrix}.
\end{aligned}
$$

These have a one-to-one correspondence with $\boldsymbol{\mu}_0$ and $\Sigma$. For the parameters $\boldsymbol{\xi}$ and $\Phi$, the likelihood function can be written as

$$
\begin{aligned}
L(\boldsymbol{\xi}, \Phi) = &\prod_{j=1}^{n} \frac{1}{(2\pi)^{(p_1+p_2)/2}|\Phi_{(12)(12)}|^{1/2}} \\
&\times \exp\left\{-\frac{1}{2}(\boldsymbol{x}_{(12)j} - \boldsymbol{\xi}_{(12)})'\Phi_{(12)(12)}^{-1}(\boldsymbol{x}_{(12)j} - \boldsymbol{\xi}_{(12)})\right\} \\
&\times \prod_{j=1}^{n_1} \frac{1}{(2\pi)^{p_3/2}|\Phi_{33}|^{1/2}} \\
&\times \exp\left\{-\frac{1}{2}(\boldsymbol{x}_{3j} - \Phi_{3(12)}\boldsymbol{x}_{(12)j} - \boldsymbol{\xi}_3)'\Phi_{33}^{-1}(\boldsymbol{x}_{3j} - \Phi_{3(12)}\boldsymbol{x}_{(12)j} - \boldsymbol{\xi}_3)\right\}.
\end{aligned}
$$

Similarly, as Theorem 1, we have the following Corollary. Note that $\Phi_{3(12)}$ and $\Phi_{33}$ correspond to the $\Psi_{3(12)}$ and $\Psi_{33}$ of Theorem 1, respectively.

**Corollary 1.** *Suppose that the data have a two-step monotone missing pattern. The maximum likelihood estimators of* $\boldsymbol{\xi}_3, \Phi_{(12)(12)}, \Phi_{3(12)}$ *and* $\Phi_{33}$ *under* $H_0$ *are given by*

$$\widetilde{\boldsymbol{\xi}}_3 = \boldsymbol{\mu}_{30} - \widetilde{\Phi}_{3(12)}\boldsymbol{\mu}_{(12)0},\ \widetilde{\Phi}_{(12)(12)} = \frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})',$$

$$\widetilde{\Phi}_{3(12)} = \left\{\sum_{j=1}^{n_1}(\boldsymbol{x}_{3j} - \boldsymbol{\mu}_{30})(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})'\right\} \times \left\{\sum_{j=1}^{n_1}(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})(\boldsymbol{x}_{(12)j} - \boldsymbol{\mu}_{(12)0})'\right\}^{-1},$$

$$\widetilde{\Phi}_{33} = \frac{1}{n_1}\sum_{j=1}^{n_1}(\boldsymbol{x}_{3j} - \widetilde{\Phi}_{3(12)}\boldsymbol{x}_{(12)j} - \widetilde{\boldsymbol{\xi}}_3)(\boldsymbol{x}_{3j} - \widetilde{\Phi}_{3(12)}\boldsymbol{x}_{(12)j} - \widetilde{\boldsymbol{\xi}}_3)',$$

*respectively.*

## §4. Likelihood ratio test

In this section, we derive the LRT statistic for testing the subvector of a mean vector with two-step monotone missing data. In the hypothesis in (1.1), the parameter space $\Omega$ and the subspace $\omega$ when $H_0$ holds, respectively, are as follows:

$$\Omega = \{(\boldsymbol{\mu}, \Sigma) : -\infty < \mu_i < \infty, i = p_1+1, p_1+2, \ldots, p, \boldsymbol{\mu}_1 = \boldsymbol{\mu}_{10}, \Sigma > 0 \text{ and } \Sigma_{(23)(23)} > 0\},$$
$$\omega = \{(\boldsymbol{\mu}, \Sigma) : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \Sigma > 0 \text{ and } \Sigma_{(23)(23)} > 0\},$$

where $\Sigma > 0$ and $\Sigma_{(23)(23)} > 0$ indicate that $\Sigma$ and $\Sigma_{(23)(23)}$ are positive definite matrices. We note that $L(\widetilde{\boldsymbol{\eta}}, \widetilde{\Psi}) = L(\widetilde{\boldsymbol{\xi}}, \widetilde{\Phi})$, where $\widetilde{\boldsymbol{\eta}}$ and $\widetilde{\Psi}$ are the MLEs of $\boldsymbol{\eta}$ and $\Psi$ under $H_0$. We have that $|\Phi_{(12)(12)}| = |\Psi_{11}|\cdot|\Psi_{22}|$ by Siotani, Hayakawa, and Fujikoshi (1985, p.591). Therefore, using the MLEs in Section 2, the likelihood ratio criterion is given by

$$\lambda_{\mathrm{M}} = \frac{\max_{\omega} L(\boldsymbol{\mu}, \Sigma)}{\max_{\Omega} L(\boldsymbol{\mu}, \Sigma)} = \lambda_{M(12)}^{\frac{n}{2}} \cdot \lambda_{M3}^{\frac{n_1}{2}},$$

where $\lambda_{M(12)} = |\widehat{\Psi}_{11}|\cdot|\widehat{\Psi}_{22}|/|\widetilde{\Phi}_{(12)(12)}|$ and $\lambda_{M3} = |\widehat{\Psi}_{33}|/|\widetilde{\Phi}_{33}|$ are independent.

Next, we consider the null distribution of $-2\log\lambda_M$. The characteristic function of $-2\log\lambda_M$ can be written as

$$\mathrm{E}[e^{it(-2\log\lambda_M)}] = \mathrm{E}[\lambda_M^{-2it}] = \mathrm{E}[\lambda_{M(12)}^{-itn}\cdot\lambda_{M3}^{-itn_1}].$$

We set

$$\begin{aligned}
\boldsymbol{z}_{(23)F} &= (\boldsymbol{z}_{2F}', \boldsymbol{z}_{3F}')' = \sqrt{n_1}\begin{pmatrix}\overline{\boldsymbol{x}}_{2F}-\boldsymbol{\mu}_2\\ \overline{\boldsymbol{x}}_{3F}-\boldsymbol{\mu}_3\end{pmatrix},\ V_F = \sqrt{n_1-1}(S_F - I_p),\\
\boldsymbol{z}_{2L} &= \sqrt{n_2}(\overline{\boldsymbol{x}}_{2L}-\boldsymbol{\mu}_2),\ V_L = \sqrt{n_2-1}(S_L - I_{p_{(12)}}),
\end{aligned}$$

where $p_{(12)} = p_1 + p_2$, $\boldsymbol{x}_{(23)j} = (\boldsymbol{x}_{2j}', \boldsymbol{x}_{3j}')'$, and the following hold:

$$\begin{aligned}
\overline{\boldsymbol{x}}_{(23)F} &= \frac{1}{n_1}\sum_{j=1}^{n_1}\boldsymbol{x}_{(23)j},\ \overline{\boldsymbol{x}}_L = (\overline{\boldsymbol{x}}_{1L}', \overline{\boldsymbol{x}}_{2L}')' = \left(\frac{1}{n_2}\sum_{j=n_1+1}^{n}\boldsymbol{x}_{1j}',\ \frac{1}{n_2}\sum_{j=n_1+1}^{n}\boldsymbol{x}_{2j}'\right)',\\
S_F &= \frac{1}{n_1-1}\sum_{j=1}^{n_1}(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_F)(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_F)',\\
S_L &= \frac{1}{n_2-1}\sum_{j=n_1+1}^{n}(\boldsymbol{x}_{(12)j} - \overline{\boldsymbol{x}}_L)(\boldsymbol{x}_{(12)j} - \overline{\boldsymbol{x}}_L)'.
\end{aligned}$$

Then, we can obtain the expansions of $\mathrm{E}[\lambda_{M(12)}^{-itn}]$ and $\mathrm{E}[\lambda_{M3}^{-itn_1}]$ as follows:

$$\begin{aligned}
\mathrm{E}[\lambda_{M(12)}^{-itn}] &= \mathrm{E}[e^{-itn\log\lambda_{M(12)}}] = \mathrm{E}[e^{it(\boldsymbol{z}_{2T}'\boldsymbol{z}_{2T}+O_p(n^{-\frac{1}{2}}))}] = (1-2it)^{-\frac{p_2}{2}} + O(n^{-1}),\\
\mathrm{E}[\lambda_{M3}^{-itn_1}] &= \mathrm{E}[e^{-itn_1\log\lambda_{M3}}] = \mathrm{E}[e^{it(\boldsymbol{z}_{3F}'\boldsymbol{z}_{3F}+O_p(n_1^{-\frac{1}{2}}))}] = (1-2it)^{-\frac{p_3}{2}} + O(n_1^{-1}),
\end{aligned}$$

where $\boldsymbol{z}_{2T} = \sqrt{n}(\overline{\boldsymbol{x}}_{2T}-\boldsymbol{\mu}_2)$. Thus, we have that

$$\mathrm{E}[e^{it(-2\log\lambda_M)}] = (1-2it)^{-\frac{p_2+p_3}{2}} + O(n^{-1}).$$

From the above that under the null hypothesis with $\Sigma = I_p$, the LRT statistic $-2\log\lambda_M$ is asymptotically distributed as $\chi^2$ with $p_2+p_3$ degrees of freedom, when $n_1, n_2 \to \infty$ with $n_i/n \to \delta \in (0,1]$, $i = 1, 2$. Even for the case of general $\Sigma$ it should be possible to prove that this holds in a similar manner. However, this becomes very complicated, and is left as a problem for a future study.

However, the upper percentile of the $\chi^2$ distribution is not a good approximation to that of the LRT statistic when the sample size is not large. We will consider an approximate upper percentile of the LRT statistic, because the exact one is not easy to obtain. In this paper, we present a simple approximation using the $n_1 \times p$ and $n \times p$ complete data sets ( see, e.g., Seko et al.

(2012)). As in Section 1, we make use of a property that is present in the case of complete data. That is, the exact upper $100\alpha$ percentile of $\lambda$ is given by

$$q_n(\alpha) = \left\{1 + \frac{(p_2+p_3)F_{p_2+p_3,n-p}(\alpha)}{n-p}\right\}^{-\frac{n}{2}},$$

and $F_{a,b}(\alpha)$ is the upper $100\alpha$ percentile of the $F$ distribution with $a$ and $b$ degrees of freedom. Thus, we can formulate an approximate upper $100\alpha$ percentile of the LRT statistic $-2\log\lambda_{\mathrm{M}}$ as

$$q^*_{\mathrm{M}}(\alpha) = -2\log\left\{\frac{p_3}{p}\, q_{n_1}(\alpha) + \frac{p_1+p_2}{p} q_n(\alpha)\right\},$$

where

$$q_{n_1}(\alpha) = \left\{1 + \frac{(p_2+p_3)F_{p_2+p_3,n_1-p}(\alpha)}{n_1-p}\right\}^{-\frac{n_1}{2}}.$$

Therefore, we reject $H_0$ if $-2\log\lambda_{\mathrm{M}} > q^*_{\mathrm{M}}(\alpha)$. In the next section, the accuracy and asymptotic behavior of the approximation are investigated using Monte Carlo simulation.

## §5. Simulation studies

In this section, we compute the upper $100\alpha$ percentiles of the LRT statistic $q_{\mathrm{sim}}(\alpha)$ using Monte Carlo simulation for $\alpha = 0.05$ and $0.01$. We generate artificial two-step missing data from $N_p(\mathbf{0}, I_p)$ for various conditions of $p_1, p_2, p_3, n_1$, and $n_2$. We simulate the upper percentiles of the LRT statistic, $q^*_{\mathrm{M}}(\alpha)$, and the type I error rates under the simulated LRT statistic when the null hypothesis is rejected using $q^*_{\mathrm{M}}(\alpha)$ and $\chi^2_{p_2+p_3}$, where

$$P_{q^*} = \Pr\{-2\log\lambda_{\mathrm{M}} > q^*_{\mathrm{M}}(\alpha)\}, \quad P_c = \Pr\{-2\log\lambda_{\mathrm{M}} > \chi^2_{p_2+p_3}(\alpha)\},$$

and $\chi^2_f(\alpha)$ is the upper $100\alpha$ percentile of the $\chi^2$ distribution with $f$ degrees of freedom. In Tables 1-4, we present the simulation results for the following four cases:

$$\begin{aligned}
&\text{Case I}: (p_1, p_2, p_3) = (2,2,4), (2,3,3), (2,4,2),\\
&\qquad (n_1, n_2) = (n_1, 2n_1), (n_1, n_1), (n_1, n_1/2), n_1 = 20, 40, 80, 160;\\
&\text{Case II}: (p_1, p_2, p_3) = (2,2,4), (2,3,3), (2,4,2),\\
&\qquad (n_1, n_2), n_1 = 20, 40, 80, 160, n_2 = 10, 20, 40;\\
&\text{Case III}: (p_1, p_2, p_3) = (2,2,2), (4,2,2), (8,2,2),\\
&\qquad (n_1, n_2) = (n_1, 2n_1), (n_1, n_1), (n_1, n_1/2), n_1 = 20, 40, 80, 160;\\
&\text{Case IV}: (p_1, p_2, p_3) = (2,2,4), (4,3,3), (6,2,2),\\
&\qquad (n_1, n_2), n_1 = 20, 40, 80, 160, n_2 = 10, 20, 40.
\end{aligned}$$

We note that the cases for $p = 8$ and $p_1 = 2$ are given in Tables 1 and 2. That is, the values of $p$ and $p_1$ are fixed. Furthermore, Tables 3 and 4 present the case where $p_2 = p_3$, and $p_2$ and $p_3$ are fixed.

From Tables 1 and 2, we can see that the proposed approximation $q^*_{\mathrm{M}}(\alpha)$ provides a good result in the case that the sample sizes $n_1$ and $n_2$ are large or the sample size $n_1$ is large and $n_2$ is fixed. Our results also indicate that the type I error rate is close to $\alpha$ when the sample size $n_1$ is large. From Tables 3 and 4, we can see that the approximation $q^*_{\mathrm{M}}(\alpha)$ is good in the case that $p_2 = p_3 = 2$ and the sample size $n_1$ is large. It can be seen from Tables 3 and 4 that the value of $q^*_{\mathrm{M}}(\alpha)$ is close to that of the LRT when $p_1$ is small. However, we note that the proposed approximation performs better than the $\chi^2$ approximation for all cases.

In addition, we used Monte Carlo simulation for some selected parameters to estimate the powers of the LRT based on two-step monotone missing data and the LRT based on partially complete data of $n_1 \times p$. In the case that the type I error is close to $\alpha$, each part of the data is set to the same degree. We expected the results for the powers of the LRT based on $q^*_{\mathrm{M}}(\alpha)$ to be larger than the corresponding powers of the LRT based on $q_{n_1}(\alpha)$. Because the type I error is not stable, the power of the LRT based on $\chi^2_{p_2+p_3}(\alpha)$ is not comparing. We note that the upper $100\alpha$ percentile of the $\chi^2_{p_2+p_3}$ is smaller than $q^*_M$, the power becomes large. This should be investigated in further detail using Monte Carlo simulation. Furthermore, we plan to discuss the power in a theoretical context in future work. In particular, we will consider the non-null distribution under local alternatives.

TABLE 1 : $p_1$ and $p$ are fixed, and $\alpha = 0.05, 0.01$

| | | $\alpha$=0.05 | | | | $\alpha$=0.01 | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $q_{\rm sim}(\alpha)$ | $q^*_{\rm M}(\alpha)$ | $P_{q^*}$ | $P_c$ | $q_{\rm sim}(\alpha)$ | $q^*_{\rm M}(\alpha)$ | $P_{q^*}$ | $P_c$ |
| $(p_1, p_2, p_3) = (2, 2, 4)$ | | | | | | | | | |
| 20 | 40 | 17.69 | 15.17 | .093 | .171 | 23.81 | 19.99 | .028 | .062 |
| 40 | 80 | 14.54 | 13.90 | .061 | .091 | 19.50 | 18.50 | .014 | .024 |
| 80 | 160 | 13.48 | 13.24 | .054 | .067 | 18.06 | 17.66 | .011 | .016 |
| 160 | 320 | 13.01 | 12.91 | .052 | .058 | 17.40 | 17.24 | .011 | .012 |
| 20 | 20 | 17.83 | 15.92 | .080 | .176 | 24.00 | 21.05 | .022 | .064 |
| 40 | 40 | 14.61 | 14.15 | .058 | .093 | 19.52 | 18.85 | .013 | .025 |
| 80 | 80 | 13.52 | 13.33 | .053 | .068 | 18.05 | 17.80 | .011 | .016 |
| 160 | 160 | 13.02 | 12.95 | .051 | .058 | 17.43 | 17.29 | .011 | .013 |
| 20 | 10 | 17.93 | 16.71 | .068 | .180 | 23.98 | 22.21 | .016 | .066 |
| 40 | 20 | 14.69 | 14.39 | .055 | .095 | 19.63 | 19.20 | .012 | .025 |
| 80 | 40 | 13.54 | 13.43 | .052 | .069 | 18.10 | 17.93 | .011 | .016 |
| 160 | 80 | 13.06 | 13.00 | .051 | .059 | 17.43 | 17.35 | .010 | .013 |
| $(p_1, p_2, p_3) = (2, 3, 3)$ | | | | | | | | | |
| 20 | 40 | 17.09 | 14.81 | .090 | .154 | 23.07 | 19.58 | .026 | .054 |
| 40 | 80 | 14.32 | 13.72 | .060 | .086 | 19.16 | 18.27 | .014 | .022 |
| 80 | 160 | 13.35 | 13.15 | .054 | .065 | 17.82 | 17.55 | .011 | .015 |
| 160 | 320 | 12.96 | 12.87 | .052 | .057 | 17.31 | 17.18 | .010 | .012 |
| 20 | 20 | 17.37 | 15.59 | .079 | .163 | 23.35 | 20.68 | .021 | .058 |
| 40 | 40 | 14.46 | 14.00 | .058 | .089 | 19.31 | 18.66 | .013 | .023 |
| 80 | 80 | 13.43 | 13.27 | .053 | .067 | 17.95 | 17.71 | .011 | .015 |
| 160 | 160 | 13.00 | 12.92 | .051 | .058 | 17.38 | 17.25 | .010 | .012 |
| 20 | 10 | 17.62 | 16.45 | .067 | .171 | 23.66 | 21.90 | .016 | .061 |
| 40 | 20 | 14.56 | 14.29 | .055 | .092 | 19.45 | 19.06 | .011 | .024 |
| 80 | 40 | 13.49 | 13.39 | .052 | .068 | 18.00 | 17.87 | .010 | .015 |
| 160 | 80 | 13.01 | 12.98 | .051 | .058 | 17.35 | 17.33 | .010 | .012 |
| $(p_1, p_2, p_3) = (2, 4, 2)$ | | | | | | | | | |
| 20 | 40 | 16.34 | 14.51 | .081 | .134 | 22.11 | 19.25 | .022 | .044 |
| 40 | 80 | 14.06 | 13.55 | .059 | .080 | 18.78 | 18.06 | .013 | .020 |
| 80 | 160 | 13.24 | 13.07 | .053 | .063 | 17.67 | 17.44 | .011 | .014 |
| 160 | 320 | 12.89 | 12.83 | .051 | .055 | 17.26 | 17.13 | .011 | .012 |
| 20 | 20 | 16.74 | 15.32 | .073 | .146 | 22.55 | 20.36 | .019 | .049 |
| 40 | 40 | 14.23 | 13.87 | .056 | .084 | 19.02 | 18.49 | .012 | .021 |
| 80 | 80 | 13.34 | 13.21 | .052 | .064 | 17.85 | 17.63 | .011 | .015 |
| 160 | 160 | 12.93 | 12.89 | .051 | .056 | 17.31 | 17.21 | .010 | .012 |
| 20 | 10 | 17.24 | 16.22 | .065 | .159 | 23.21 | 21.62 | .015 | .056 |
| 40 | 20 | 14.44 | 14.19 | .054 | .089 | 19.27 | 18.94 | .011 | .023 |
| 80 | 40 | 13.43 | 13.34 | .051 | .067 | 17.92 | 17.81 | .010 | .015 |
| 160 | 80 | 12.98 | 12.96 | .050 | .057 | 17.33 | 17.30 | .010 | .012 |

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

TABLE 2 : $p_1$, $p$ and $n_2$ are fixed, and $\alpha = 0.05, 0.01$

| | | $\alpha$=0.05 | | | | $\alpha$=0.01 | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $q_{\text{sim}}(\alpha)$ | $q^*_{\text{M}}(\alpha)$ | $P_{q^*}$ | $P_c$ | $q_{\text{sim}}(\alpha)$ | $q^*_{\text{M}}(\alpha)$ | $P_{q^*}$ | $P_c$ |
| $(p_1, p_2, p_3) = (2, 2, 4)$ | | | | | | | | | |
| 20 | 10 | 17.93 | 16.71 | .068 | .180 | 23.98 | 22.21 | .016 | .066 |
| 40 | 10 | 14.75 | 14.58 | .053 | .097 | 19.72 | 19.47 | .011 | .026 |
| 80 | 10 | 13.58 | 13.56 | .050 | .070 | 18.13 | 18.10 | .010 | .016 |
| 160 | 10 | 13.08 | 13.07 | .050 | .059 | 17.50 | 17.45 | .010 | .013 |
| 20 | 20 | 17.83 | 15.92 | .080 | .176 | 24.00 | 21.05 | .022 | .064 |
| 40 | 20 | 14.69 | 14.39 | .055 | .095 | 19.63 | 19.20 | .012 | .025 |
| 80 | 20 | 13.56 | 13.51 | .051 | .069 | 18.07 | 18.04 | .010 | .016 |
| 160 | 20 | 13.08 | 13.06 | .050 | .059 | 17.43 | 17.43 | .010 | .013 |
| 20 | 40 | 17.69 | 15.17 | .093 | .171 | 23.81 | 19.99 | .028 | .062 |
| 40 | 40 | 14.61 | 14.15 | .058 | .093 | 19.52 | 18.85 | .013 | .025 |
| 80 | 40 | 13.54 | 13.43 | .052 | .069 | 18.10 | 17.93 | .011 | .016 |
| 160 | 40 | 13.06 | 13.03 | .051 | .059 | 17.42 | 17.40 | .010 | .013 |
| $(p_1, p_2, p_3) = (2, 3, 3)$ | | | | | | | | | |
| 20 | 10 | 17.62 | 16.45 | .067 | .171 | 23.66 | 21.90 | .016 | .061 |
| 40 | 10 | 14.68 | 14.51 | .053 | .094 | 19.60 | 19.38 | .011 | .025 |
| 80 | 10 | 13.56 | 13.54 | .050 | .069 | 18.11 | 18.08 | .010 | .016 |
| 160 | 10 | 13.07 | 13.07 | .050 | .059 | 17.47 | 17.44 | .010 | .013 |
| 20 | 20 | 17.37 | 15.59 | .079 | .163 | 23.35 | 20.68 | .021 | .058 |
| 40 | 20 | 14.56 | 14.29 | .055 | .092 | 19.45 | 19.06 | .011 | .024 |
| 80 | 20 | 13.54 | 13.48 | .051 | .069 | 18.09 | 18.00 | .010 | .016 |
| 160 | 20 | 13.06 | 13.05 | .050 | .059 | 17.43 | 17.42 | .010 | .013 |
| 20 | 40 | 17.09 | 14.81 | .090 | .154 | 23.07 | 19.58 | .026 | .054 |
| 40 | 40 | 14.46 | 14.00 | .058 | .089 | 19.31 | 18.66 | .013 | .023 |
| 80 | 40 | 13.49 | 13.39 | .052 | .068 | 18.00 | 17.87 | .010 | .015 |
| 160 | 40 | 13.05 | 13.02 | .050 | .059 | 17.40 | 17.38 | .010 | .013 |
| $(p_1, p_2, p_3) = (2, 4, 2)$ | | | | | | | | | |
| 20 | 10 | 17.24 | 16.22 | .065 | .159 | 23.21 | 21.62 | .015 | .056 |
| 40 | 10 | 14.58 | 14.45 | .052 | .092 | 19.47 | 19.30 | .011 | .024 |
| 80 | 10 | 13.57 | 13.53 | .051 | .070 | 18.12 | 18.06 | .010 | .016 |
| 160 | 10 | 13.09 | 13.06 | .050 | .060 | 17.51 | 17.44 | .010 | .013 |
| 20 | 20 | 16.74 | 15.32 | .073 | .146 | 22.55 | 20.36 | .019 | .049 |
| 40 | 20 | 14.44 | 14.19 | .054 | .089 | 19.27 | 18.94 | .011 | .023 |
| 80 | 20 | 13.50 | 13.45 | .051 | .068 | 18.04 | 17.96 | .010 | .016 |
| 160 | 20 | 13.08 | 13.04 | .051 | .059 | 17.49 | 17.41 | .010 | .013 |
| 20 | 40 | 16.34 | 14.51 | .081 | .134 | 22.11 | 19.25 | .022 | .044 |
| 40 | 40 | 14.23 | 13.87 | .056 | .084 | 19.02 | 18.49 | .012 | .021 |
| 80 | 40 | 13.43 | 13.34 | .051 | .067 | 17.92 | 17.81 | .010 | .015 |
| 160 | 40 | 13.02 | 13.01 | .050 | .058 | 17.41 | 17.37 | .010 | .013 |

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

TABLE 3 : $p_2 = p_3$, and $\alpha = 0.05, 0.01$

| | | $\alpha$=0.05 | | | | $\alpha$=0.01 | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $q_{\text{sim}}(\alpha)$ | $q^*_{\text{M}}(\alpha)$ | $P_{q^*}$ | $P_c$ | $q_{\text{sim}}(\alpha)$ | $q^*_{\text{M}}(\alpha)$ | $P_{q^*}$ | $P_c$ |
| $(p_1, p_2, p_3) = (2, 2, 2)$ | | | | | | | | | |
| 20 | 40 | 11.93 | 10.88 | .070 | .109 | 16.78 | 15.12 | .018 | .032 |
| 40 | 80 | 10.51 | 10.17 | .057 | .073 | 14.72 | 14.21 | .012 | .018 |
| 80 | 160 | 9.96 | 9.82 | .053 | .060 | 13.96 | 13.74 | .011 | .013 |
| 160 | 320 | 9.70 | 9.65 | .051 | .055 | 13.58 | 13.51 | .010 | .011 |
| 20 | 20 | 12.10 | 11.31 | .064 | .113 | 16.98 | 15.75 | .015 | .034 |
| 40 | 40 | 10.57 | 10.34 | .055 | .075 | 14.78 | 14.45 | .011 | .018 |
| 80 | 80 | 10.01 | 9.90 | .052 | .061 | 14.02 | 13.85 | .011 | .013 |
| 160 | 160 | 9.75 | 9.69 | .051 | .056 | 13.61 | 13.56 | .010 | .012 |
| 20 | 10 | 12.29 | 11.76 | .059 | .119 | 17.30 | 16.42 | .013 | .037 |
| 40 | 20 | 10.66 | 10.51 | .053 | .077 | 14.97 | 14.70 | .011 | .019 |
| 80 | 40 | 10.05 | 9.97 | .052 | .062 | 14.06 | 13.95 | .010 | .014 |
| 160 | 80 | 9.75 | 9.72 | .050 | .056 | 13.60 | 13.61 | .010 | .011 |
| $(p_1, p_2, p_3) = (4, 2, 2)$ | | | | | | | | | |
| 20 | 40 | 13.35 | 11.23 | .091 | .146 | 18.87 | 15.57 | .026 | .051 |
| 40 | 80 | 10.98 | 10.35 | .062 | .084 | 15.39 | 14.45 | .014 | .022 |
| 80 | 160 | 10.13 | 9.91 | .054 | .064 | 14.19 | 13.86 | .011 | .014 |
| 160 | 320 | 9.81 | 9.70 | .052 | .057 | 13.70 | 13.57 | .011 | .012 |
| 20 | 20 | 13.59 | 11.96 | .079 | .155 | 19.19 | 16.62 | .021 | .055 |
| 40 | 40 | 11.06 | 10.63 | .058 | .086 | 15.49 | 14.85 | .013 | .022 |
| 80 | 80 | 10.20 | 10.03 | .053 | .066 | 14.29 | 14.03 | .011 | .015 |
| 160 | 160 | 9.82 | 9.75 | .051 | .057 | 13.75 | 13.65 | .010 | .012 |
| 20 | 10 | 13.94 | 12.78 | .069 | .165 | 19.60 | 17.82 | .017 | .060 |
| 40 | 20 | 11.22 | 10.91 | .056 | .091 | 15.72 | 15.26 | .012 | .024 |
| 80 | 40 | 10.27 | 10.15 | .052 | .067 | 14.40 | 14.21 | .011 | .015 |
| 160 | 80 | 9.86 | 9.81 | .051 | .058 | 13.78 | 13.73 | .010 | .012 |
| $(p_1, p_2, p_3) = (8, 2, 2)$ | | | | | | | | | |
| 20 | 40 | 18.24 | 11.99 | .170 | .271 | 26.27 | 16.63 | .068 | .132 |
| 40 | 80 | 12.01 | 10.71 | .076 | .111 | 16.91 | 14.92 | .019 | .033 |
| 80 | 160 | 10.55 | 10.09 | .059 | .074 | 14.79 | 14.10 | .013 | .018 |
| 160 | 320 | 9.99 | 9.78 | .054 | .061 | 13.99 | 13.69 | .011 | .013 |
| 20 | 20 | 18.67 | 13.47 | .137 | .289 | 26.83 | 18.71 | .050 | .142 |
| 40 | 40 | 12.28 | 11.23 | .069 | .118 | 17.28 | 15.67 | .017 | .036 |
| 80 | 80 | 10.64 | 10.30 | .057 | .076 | 14.91 | 14.41 | .012 | .018 |
| 160 | 160 | 10.03 | 9.88 | .053 | .062 | 14.03 | 13.83 | .011 | .014 |
| 20 | 10 | 19.33 | 15.38 | .106 | .311 | 27.60 | 21.42 | .033 | .158 |
| 40 | 20 | 12.47 | 11.81 | .062 | .125 | 17.48 | 16.50 | .014 | .039 |
| 80 | 40 | 10.77 | 10.53 | .055 | .079 | 15.06 | 14.73 | .011 | .020 |
| 160 | 80 | 10.07 | 9.98 | .052 | .063 | 14.14 | 13.97 | .011 | .014 |

Note : $\chi^2_4(0.05) = 9.49$, $\chi^2_4(0.01) = 13.28$

TABLE 4 : $n_2$ is fixed, $p_2 = p_3$, and $\alpha = 0.05, 0.01$

| | | $\alpha$=0.05 | | | | $\alpha$=0.01 | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n_1$ | $n_2$ | $q_{\rm sim}(\alpha)$ | $q^*_{\rm M}(\alpha)$ | $P_{q^*}$ | $P_c$ | $q_{\rm sim}(\alpha)$ | $q^*_{\rm M}(\alpha)$ | $P_{q^*}$ | $P_c$ |
| $(p_1, p_2, p_3) = (2, 2, 2)$ | | | | | | | | | |
| 20 | 10 | 12.29 | 11.76 | .059 | .119 | 17.30 | 16.42 | .013 | .037 |
| 40 | 10 | 10.73 | 10.65 | .052 | .078 | 15.04 | 14.90 | .011 | .020 |
| 80 | 10 | 10.11 | 10.07 | .051 | .063 | 14.16 | 14.10 | .010 | .014 |
| 160 | 10 | 9.79 | 9.78 | .050 | .056 | 13.69 | 13.69 | .010 | .012 |
| 20 | 20 | 12.10 | 11.31 | .064 | .113 | 16.98 | 15.75 | .015 | .034 |
| 40 | 20 | 10.66 | 10.51 | .053 | .077 | 14.97 | 14.70 | .011 | .019 |
| 80 | 20 | 10.06 | 10.03 | .050 | .062 | 14.10 | 14.04 | .010 | .014 |
| 160 | 20 | 9.77 | 9.77 | .050 | .056 | 13.68 | 13.67 | .010 | .012 |
| 20 | 40 | 11.93 | 10.88 | .070 | .109 | 16.78 | 15.12 | .018 | .032 |
| 40 | 40 | 10.57 | 10.34 | .055 | .075 | 14.78 | 14.45 | .011 | .018 |
| 80 | 40 | 10.05 | 9.97 | .052 | .062 | 14.06 | 13.95 | .010 | .014 |
| 160 | 40 | 9.79 | 9.75 | .051 | .056 | 13.70 | 13.65 | .010 | .012 |
| $(p_1, p_2, p_3) = (4, 2, 2)$ | | | | | | | | | |
| 20 | 10 | 13.94 | 12.78 | .069 | .165 | 19.60 | 17.82 | .017 | .060 |
| 40 | 10 | 11.34 | 11.15 | .054 | .094 | 15.86 | 15.60 | .011 | .025 |
| 80 | 10 | 10.34 | 10.32 | .050 | .069 | 14.49 | 14.44 | .010 | .016 |
| 160 | 10 | 9.90 | 9.90 | .050 | .059 | 13.83 | 13.86 | .010 | .013 |
| 20 | 20 | 13.59 | 11.96 | .079 | .155 | 19.19 | 16.62 | .021 | .055 |
| 40 | 20 | 11.22 | 10.91 | .056 | .091 | 15.72 | 15.26 | .012 | .024 |
| 80 | 20 | 10.32 | 10.25 | .051 | .068 | 14.44 | 14.34 | .010 | .016 |
| 160 | 20 | 9.89 | 9.88 | .050 | .059 | 13.84 | 13.83 | .010 | .013 |
| 20 | 40 | 13.35 | 11.23 | .091 | .146 | 18.87 | 15.57 | .026 | .051 |
| 40 | 40 | 11.06 | 10.63 | .058 | .086 | 15.49 | 14.85 | .013 | .022 |
| 80 | 40 | 10.27 | 10.15 | .052 | .067 | 14.40 | 14.21 | .011 | .015 |
| 160 | 40 | 9.88 | 9.85 | .051 | .059 | 13.82 | 13.79 | .010 | .012 |
| $(p_1, p_2, p_3) = (8, 2, 2)$ | | | | | | | | | |
| 20 | 10 | 19.33 | 15.38 | .106 | .311 | 27.60 | 21.42 | .033 | .158 |
| 40 | 10 | 12.75 | 12.30 | .057 | .132 | 17.88 | 17.21 | .012 | .043 |
| 80 | 10 | 10.92 | 10.84 | .051 | .083 | 15.30 | 15.17 | .011 | .021 |
| 160 | 10 | 10.16 | 10.15 | .050 | .065 | 14.23 | 14.21 | .010 | .015 |
| 20 | 20 | 18.67 | 13.47 | .137 | .289 | 26.83 | 18.71 | .050 | .142 |
| 40 | 20 | 12.47 | 11.81 | .062 | .125 | 17.48 | 16.50 | .014 | .039 |
| 80 | 20 | 10.84 | 10.72 | .052 | .081 | 15.19 | 14.99 | .011 | .020 |
| 160 | 20 | 10.15 | 10.12 | .051 | .064 | 14.18 | 14.16 | .010 | .015 |
| 20 | 40 | 18.24 | 11.99 | .170 | .271 | 26.27 | 16.63 | .068 | .132 |
| 40 | 40 | 12.28 | 11.23 | .069 | .118 | 17.28 | 15.67 | .017 | .036 |
| 80 | 40 | 10.77 | 10.53 | .055 | .079 | 15.06 | 14.73 | .011 | .020 |
| 160 | 40 | 10.13 | 10.06 | .051 | .064 | 14.21 | 14.08 | .010 | .014 |

Note : $\chi^2_4(0.05) = 9.49$, $\chi^2_4(0.01) = 13.28$

## §6. Numerical example

Now, we illustrate the results of this study using an example given in Wei and Lachin (1984). The sample data set consists of serum cholesterol values that were measured under the treatment group at five different time points: the baseline and at months 6, 12, 20, and 24. The original data set contains 36 complete observations, and we create two-step monotone missing data by randomly selecting 30 observations and deleting the values for 10 observations for each of the months 20 and 24. Thus, we have $n = 30$, $n_1 = 20$, $n_2 = 10$, $p = 5$, $p_1 = 1$, and $p_2 = p_3 = 2$. We are interested in the change from the baseline at each post-baseline time point. It is known that the mean for all baseline value was 220. We consider the hypothesis $H : (\mu_2, \mu_3, \mu_4, \mu_5)' = (220, 220, 220, 220)'$, given $\mu_1 = 220$. Then, we compute $-2 \log \lambda_{\mathrm{M}} = 10.92$. Because we have $q_{\mathrm{sim}}(0.05) = 11.63$ from the simulation study, we do not reject the null hypothesis at the 0.05 significance level. Moreover, when we use $q^*_{\mathrm{M}}(0.05) = 11.30$, the null hypothesis is not rejected. When we use $\chi^2_{4(0.05)} = 9.49$, the null hypothesis is rejected. However, $q_{\mathrm{sim}}(0.01) = 16.32$ from the simulation study, and the null hypothesis is rejected at the 0.01 significance level. When we use $q^*_{\mathrm{M}}(0.01) = 15.79$ or $\chi^2_{4(0.01)} = 13.28$, the null hypothesis is also rejected.

## §7. Concluding remarks

In this paper, we have considered the one-sample problem of testing for the subvector of a mean vector with two-step monotone missing data. First, we provided an introduction to two-step monotone missing data. Then, we reviewed the test for the subvector of a mean vector with non-missing data. In the case that the data set consists of complete data with $p$ dimensions and incomplete data with $(p_1 + p_2)$ dimensions, we derived the likelihood ratio criterion for testing the $(p_2 + p_3)$ mean vector under the given mean vector of $p_1$ dimensions, which is given by (1.1). This test procedure only treats the $(p_2 + p_3)$-components as if observations are present. Next, we derived the MLEs, and provided the LRT statistic and the approximate upper $100\alpha$ percentiles of the LRT, $q^*_M(\alpha)$, for a subvector. The approximate values can easily be calculated, and the simulation results suggest that the type I error rates are close to $\alpha$ when the sample size $n_1$ is large. In all cases, it appears that the approximate upper $100\alpha$ percentiles $q^*_M(\alpha)$ are preferable to $\chi^2_{p_2+p_3}$.

**Acknowledgments**

The authors would like to thank the referee for helpful comments and suggestions. Second author's research was in part supported by Grant-in-Aid for Scientific Research (C) (26330050).

## References

[1] Anderson, T. W. (1957). Maximum likelihood estimates for a multivariate normal distribution where some observations are missing. *J. Amer. Statist. Assoc.*, **52**, 200–203.

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

[3] Eaton, M. L. and Kariya, T. (1975). Tests on means with additional information. Technical Report #243, *University of Minnesota*.

[4] Giri, N. C. (1964). On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Stat.*, **35**, 181–189, 1388.

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

[6] Krishnamoorthy, K. and Pannala, K. M. (1999). Confidence estimation of a normal mean vector with incomplete data. *Canad. J. Statist.*, **27**, 395–407.

[7] Provost, S. B. (1990). Estimators for the parameters of a multivariate normal random vector with incomplete data on two subvectors and test of independence. *Comput. Statist. Data Anal.*, **9**, 37–46.

[8] Rao, C. R. (1949). On some problems arising out of discrimination with multiple characters. *Sankhyā*, **9**, 343–364.

[9] Seber, G. A. F. (1984). *Multivariate Observations*, John Wiley & Sons, Inc.

[10] Seko, N., Kawasaki, T. and Seo, T. (2011). Testing equality of two mean vectors with two-step monotone missing data. *Amer. J. Math. Management Sci.*, **31**, 117–135.

[11] Seko, N., Yamazaki, A. and Seo, T. (2012). Tests for mean vector with two-step monotone missing data. *SUT J. Math.*, **48**, 13–38.

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

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

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

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

[16] Wei, L. J. and Lachin, J. M. (1984). Two-sample asymptotically distribution-free tests for incomplete multivariate observations. *J. Amer. Statist. Assoc.*, **79**, 653–661.

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

Tamae Kawasaki
Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
*E-mail*: tm.kawasaki@rs.tus.ac.jp

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