Tests for mean vector and simultaneous confidence intervals with two-step monotone missing data

Noriko Seko[1], Akiko Yamazaki[2] and Takashi Seo[1]

[1]Department of Mathematical Information Science
Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan
[2]Clinical Data Analiysis Div.,
The Institute of Japanese Union of Scientists & Engineers
5-10-9, Sendagaya, Shibuya-ku, Tokyo, 151-0051, Japan

Abstract

In this paper, we consider the tests for mean vector and simultaneous confidence intervals in one sample problem when the data has two-step monotone pattern missing observations. MLEs of mean vector and covariance matrix with two-step monotone missing data have been introduced by Anderson and Olkin (1985), and the distribution of MLEs has been discussed by Kanda and Fujikoshi (1998). Using these MLEs and the distribution, we give Hotelling's T^2 type statistic and likelihood ratio test statistic for mean vector. The accuracy of asymptotic distributions of these test statistics is investigated by Monte Carlo simulation for some selected parameters. Simultaneous confidence interval are also obtained.

Key Words and Phrases: Hotelling's T^2 type statistic; Likelihood ratio test statistic; Maximum likelihood estimator; Simultaneous confidence interval; Two-step monotone missing data

1 Introduction

In statistical data analyses, we often face to the data with missing observations. For a general missing pattern, many statistical methods have been developed by Srivastava(1985), Srivastava and Carter (1986) and Shutoh et al. (2009). When the missing pattern is monotone, Seo and Srivastava (2000) discussed the test of equality of means and simultaneous confidence intervals in one sample problem, and Koizumi and Seo (2009a, 2009b) considered testing equality of means and simultaneous confidence intervals in k samples problem for k-step monotone missing data. For a two-step monotone missing data, Anderson and Olkin (1985) obtained the MLEs of mean and covariance vector for one sample problem, and Kanda and Fujikoshi (1998) discussed the distribution of these MLEs.

A two-step monotone missing data is a data set that missing occurs in all observations after one specific point with some samples,

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p_1} & x_{1p_1+1} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p_1} & x_{2p_1+1} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{N_11} & x_{N_12} & \dots & x_{N_1p_1} & x_{N_1p_1+1} & \dots & x_{N_1p} \\ x_{N_1+11} & x_{N_1+12} & \dots & x_{N_1+1p_1} & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Np_1} & * & \dots & * \end{pmatrix}$$

where $N = N_1 + N_2$ and $p = p_1 + p_2$. The data can be written in a vector expression as below;

$$\begin{pmatrix} \boldsymbol{x}'_{11} & \boldsymbol{x}'_{21} \\ \boldsymbol{x}'_{12} & \boldsymbol{x}'_{22} \\ \vdots & \vdots \\ \boldsymbol{x}'_{1N_1} & \boldsymbol{x}'_{2N_1} \\ \boldsymbol{x}'_{1N_1+1} & * \\ \vdots & \vdots \\ \boldsymbol{x}'_{1N} & * \end{pmatrix}$$

Let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_1}$ be distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and let $\boldsymbol{x}_{1N_1+1}, \dots, \boldsymbol{x}_{1N}$ be distributed as $N_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$. We partition $\boldsymbol{x}_j$ as

$$\boldsymbol{x}_j = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j})',$$

where $\boldsymbol{x}_{ij} : p_i \times 1,\ i = 1, 2,\ j = 1, \dots, N_1$. Then the marginal density function of the observed data set $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_1}, \boldsymbol{x}_{1N_1+1}, \dots, \boldsymbol{x}_{1N}$ can be written by

$$\prod_{j=1}^{N_1} f(\boldsymbol{x}_j; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \times \prod_{j=N_1+1}^{N} f(\boldsymbol{x}_{1j}; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}), \tag{1}$$

where $f(\boldsymbol{x}_j; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $f(\boldsymbol{x}_{1j}; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ are the density functions of $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $N_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, respectively, and

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

We introduce some notation for the sample mean vectors and covariance matrices. Let $\overline{\boldsymbol{x}}^{(1)}$ denote the sample mean vector of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_1}$, and $\overline{\boldsymbol{x}}^{(1)} = (\overline{\boldsymbol{x}}_1^{(1)'}, \overline{\boldsymbol{x}}_2^{(1)'})',\ \overline{\boldsymbol{x}}_i^{(1)} : p_i \times 1$. Let $\overline{\boldsymbol{x}}^{(2)}$ denote the sample mean vector of $\boldsymbol{x}_{1N_1+1}, \dots, \boldsymbol{x}_{1N}$. That is

$$\overline{\boldsymbol{x}}_1^{(1)} = \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{1j},$$

$$\overline{\boldsymbol{x}}_2^{(1)} = \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{2j},$$

$$\overline{\boldsymbol{x}}^{(2)} = \frac{1}{N_2} \sum_{j=N_1+1}^{N} \boldsymbol{x}_{1j}.$$

The sample covariance matrices based on the N_1 and N_2 observations are expressed as

$$\boldsymbol{S}^{(1)} = \frac{1}{N_1 - 1} \sum_{j=1}^{N_1} \left(\boldsymbol{x}_j - \overline{\boldsymbol{x}}^{(1)} \right) \left(\boldsymbol{x}_j - \overline{\boldsymbol{x}}^{(1)} \right)' = \begin{pmatrix} \boldsymbol{S}_{11}^{(1)} & \boldsymbol{S}_{12}^{(1)} \\ \boldsymbol{S}_{21}^{(1)} & \boldsymbol{S}_{22}^{(1)} \end{pmatrix},$$

$$\boldsymbol{S}^{(2)} = \frac{1}{N_2 - 1} \sum_{j=N_1+1}^{N} \left(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}^{(2)} \right) \left(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}^{(2)} \right)',$$

respectively. In this paper, we give T^2 type statistic and likelihood ratio test statistic for hypothesis $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, where $\boldsymbol{\mu}_0$ is known. We also study the asymptotic distributions of these statistics in a situation when

$$\rho_i = \frac{n_i}{n} \to \text{positive} \quad \text{constants}$$

as N_i's tend to infinity$(i = 1, 2)$, where $n_i = N_i - 1$ and $n = n_1 + n_2$. In addition, we examined for the case that $\rho_1 = 1$ as N_1 is large and N_2 is fixed.

The paper is organized as follows. Section 2 will introduce MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in general and MLE of $\boldsymbol{\Sigma}$ under $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 (= \boldsymbol{0})$. Section 3 will discuss T^2 type statistic and likelihood ratio test statistic for hypothesis H_0 and section 4 will give simultaneous confidence intervals for $\boldsymbol{\mu}$. Simulation results will be provided in section 5 to evaluate the accuracy of asymptotic null distributions of T^2 type statistic and likelihood ratio test statistic. An numerical example will be provided for simultaneous confidence intervals in section 6.

2 Maximum likelihood estimators

2.1 MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

Let the MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denote by $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$, which are partitioned in the same way as $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. We assume observation vectors are distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $N_1 > p$ which is necessary and sufficient condition for existence and uniqueness of the MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Anderson and Olkin (1985) derived the MLEs $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ given by

$$\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{N} \left(N_1 \overline{\boldsymbol{x}}_1^{(1)} + N_2 \overline{\boldsymbol{x}}^{(2)} \right) \\ \overline{\boldsymbol{x}}_2^{(1)} - \widehat{\boldsymbol{\Sigma}}_{21} \widehat{\boldsymbol{\Sigma}}_{11}^{-1} \left(\overline{\boldsymbol{x}}_1^{(1)} - \widehat{\boldsymbol{\mu}}_1 \right) \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_{11} & \widehat{\boldsymbol{\Sigma}}_{12} \\ \widehat{\boldsymbol{\Sigma}}_{21} & \widehat{\boldsymbol{\Sigma}}_{22} \end{pmatrix},$$

where

$$\widehat{\boldsymbol{\Sigma}}_{11} = \frac{1}{N}\left(\boldsymbol{W}_{11}^{(1)} + \boldsymbol{W}^{(2)}\right),$$
$$\widehat{\boldsymbol{\Sigma}}_{12} = \widehat{\boldsymbol{\Sigma}}_{11}\left(\boldsymbol{W}_{11}^{(1)}\right)^{-1}\boldsymbol{W}_{12}^{(1)},$$
$$\widehat{\boldsymbol{\Sigma}}_{22} = \frac{1}{N_1}\boldsymbol{W}_{22\cdot 1}^{(1)} + \widehat{\boldsymbol{\Sigma}}_{21}\widehat{\boldsymbol{\Sigma}}_{11}^{-1}\widehat{\boldsymbol{\Sigma}}_{12},$$

and

$$\boldsymbol{W}^{(1)} = (N_1 - 1)\boldsymbol{S}^{(1)} = \begin{pmatrix} \boldsymbol{W}_{11}^{(1)} & \boldsymbol{W}_{12}^{(1)} \\ \boldsymbol{W}_{21}^{(1)} & \boldsymbol{W}_{22}^{(1)} \end{pmatrix},$$
$$\boldsymbol{W}^{(2)} = (N_2 - 1)\boldsymbol{S}^{(2)} + \frac{N_1 N_2}{N}\left(\overline{\boldsymbol{x}}_1^{(1)} - \overline{\boldsymbol{x}}^{(2)}\right)\left(\overline{\boldsymbol{x}}_1^{(1)} - \overline{\boldsymbol{x}}^{(2)}\right)',$$
$$\boldsymbol{W}_{22\cdot 1}^{(1)} = \boldsymbol{W}_{22}^{(1)} - \boldsymbol{W}_{21}^{(1)}\left(\boldsymbol{W}_{11}^{(1)}\right)^{-1}\boldsymbol{W}_{12}^{(1)}.$$

These MLEs are derived using the usual transformed parameters defined by

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1 \end{pmatrix},$$

$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{\Psi}_{12} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{\Sigma}_{22\cdot 1} \end{pmatrix},$$

which are one to one correspondence to $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. Multiplying the observation vector $\boldsymbol{x}_j$ by the transformation matrix;

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_{p_1} & \boldsymbol{O} \\ -\boldsymbol{\Psi}_{21} & \boldsymbol{I}_{p_2} \end{pmatrix}$$

on the left side, the mean vector and the covariance matrix of transformed observation vector are

$$\boldsymbol{A}\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 - \boldsymbol{\Psi}_{21}\boldsymbol{\mu}_1 \end{pmatrix} = \boldsymbol{\eta}, \qquad \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}' = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Psi}_{22} \end{pmatrix},$$

respectively. The MLEs of $(\boldsymbol{\eta}, \boldsymbol{\Psi})$ are expressed as follows:

$$\widehat{\boldsymbol{\eta}}_1 = \widehat{\boldsymbol{\mu}}_1, \quad \widehat{\boldsymbol{\eta}}_2 = \overline{\boldsymbol{x}}_2^{(1)} - \widehat{\boldsymbol{\Psi}}_{21}\overline{\boldsymbol{x}}_1^{(1)},$$
$$\widehat{\boldsymbol{\Psi}}_{11} = \widehat{\boldsymbol{\Sigma}}_{11}, \quad \widehat{\boldsymbol{\Psi}}_{12} = \left(\boldsymbol{W}_{11}^{(1)}\right)^{-1}\boldsymbol{W}_{12}^{(1)}, \quad \widehat{\boldsymbol{\Psi}}_{22} = \tfrac{1}{N_1}\boldsymbol{W}_{22\cdot 1}^{(1)}.$$

Kanda and Fujikoshi (1998) derived the next result.

Theorem 2.1. (Kanda and Fujikoshi (1998))

The mean vector and the covariance matrix of $\widehat{\boldsymbol{\mu}}$ are given by

$$\begin{aligned}\mathrm{E}[\widehat{\boldsymbol{\mu}}] &= \boldsymbol{\mu},\\ \mathrm{Cov}[\widehat{\boldsymbol{\mu}}] &= \begin{pmatrix} \frac{1}{N}\boldsymbol{\Sigma}_{11} & \frac{1}{N}\boldsymbol{\Sigma}_{12} \\ \frac{1}{N}\boldsymbol{\Sigma}_{21} & \mathrm{Cov}[\widehat{\boldsymbol{\mu}}_2] \end{pmatrix},\end{aligned}$$

respectively, where

$$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}_2] = \frac{1}{N_1}\left(\boldsymbol{\Sigma}_{22} - \frac{N_2}{N}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\right) + \frac{N_2 p_1}{N N_1 (N_1 - p_1 - 2)}\boldsymbol{\Sigma}_{22\cdot 1} \qquad (N_1 > p_1 + 2).$$

2.2 MLE of $\boldsymbol{\Sigma}$ under $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 (= \mathbf{0})$

In this section, we consider the MLE of $\boldsymbol{\Sigma}$ under $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0(= \mathbf{0})$. Let $\boldsymbol{x}_j = (\boldsymbol{x}_{1j}, \boldsymbol{x}_{2j})$ be distributed as $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ $(j = 1, \ldots, N_1)$ and let $\boldsymbol{x}_{1j}$ be distributed as $N_{p1}(\mathbf{0}, \boldsymbol{\Sigma}_{11})$ $(j = N_1 + 1, \ldots, N)$. The likelihood function is

$$\begin{aligned} L(\mathbf{0}, \boldsymbol{\Sigma}) &= \prod_{j=1}^{N_1} \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{x}_j'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_j\right) \\ &\quad \times \prod_{j=N_1+1}^{N} \frac{1}{(2\pi)^{p_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{x}_{1j}'\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_{1j}\right). \end{aligned}$$

Multiplying the observation vector by $\boldsymbol{A}$ on the left side, we have

$$\boldsymbol{A}\boldsymbol{x}_j = \begin{pmatrix} \boldsymbol{x}_{1j} \\ \boldsymbol{x}_{2j} - \boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j} \end{pmatrix} \sim N_p\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Psi}_{22} \end{pmatrix}\right), \ j = 1, \ldots, N_1.$$

We note that $\boldsymbol{\Sigma}$ is one to one correspondence to $\boldsymbol{\Psi}$. For parameter $\boldsymbol{\Psi}$, the likelihood function can be written as

$$\begin{aligned} L(\mathbf{0}, \boldsymbol{\Psi}) &= \prod_{j=1}^{N} \frac{1}{(2\pi)^{p_1/2}|\boldsymbol{\Psi}_{11}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{x}_{1j}'\boldsymbol{\Psi}_{11}^{-1}\boldsymbol{x}_{1j}\right) \\ &\quad \times \prod_{j=1}^{N_1} \frac{1}{(2\pi)^{p_2/2}|\boldsymbol{\Psi}_{22}|^{1/2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{x}_{2j} - \boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)'\boldsymbol{\Psi}_{22}^{-1}\left(\boldsymbol{x}_{2j} - \boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)\right\}. \end{aligned}$$

Thus, the log likelihood function is

$$\begin{aligned} \log L(\mathbf{0}, \boldsymbol{\Psi}) &= -\left(\frac{p_1 N}{2} + \frac{p_2 N_1}{2}\right)\log(2\pi) - \frac{N}{2}\log|\boldsymbol{\Psi}_{11}| - \frac{N_1}{2}\log|\boldsymbol{\Psi}_{22}| \\ &\quad + \sum_{j=1}^{N}\left(-\frac{1}{2}\boldsymbol{x}_{1j}'\boldsymbol{\Psi}_{11}^{-1}\boldsymbol{x}_{1j}\right) + \sum_{j=1}^{N_1}\left\{-\frac{1}{2}\left(\boldsymbol{x}_{2j} - \boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)'\boldsymbol{\Psi}_{22}^{-1}\left(\boldsymbol{x}_{2j} - \boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)\right\}. \end{aligned}$$

The partial derivative of $\log L(\mathbf{0},\boldsymbol{\Psi})$ with respect to $\boldsymbol{\Psi}_{11}$ is given by

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

Solving the equation, the MLE of $\boldsymbol{\Psi}_{11}$ is obtained as follows;

$$\widetilde{\boldsymbol{\Psi}}_{11}=\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'.$$

Similarly, the partial derivatives of $\log L(\mathbf{0},\boldsymbol{\Psi})$ with respect to $\boldsymbol{\Psi}_{21}$ and $\boldsymbol{\Psi}_{22}$ are

$$\frac{\partial \log L\left(\boldsymbol{\eta},\boldsymbol{\Psi}\right)}{\partial \boldsymbol{\Psi}_{21}} \;=\; \sum_{j=1}^{N_1}\left(\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{x}_{2j}\boldsymbol{x}_{1j}'-\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'\right),$$

and

$$\frac{\partial \log L(\boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}_{22}} \;=\; -\frac{N_1}{2}\boldsymbol{\Psi}_{22}^{-1}+\sum_{j=1}^{N_1}\frac{1}{2}\boldsymbol{\Psi}_{22}^{-1}\left(\boldsymbol{x}_{2j}-\boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)\left(\boldsymbol{x}_{2j}-\boldsymbol{\Psi}_{21}\boldsymbol{x}_{1j}\right)'\boldsymbol{\Psi}_{22}^{-1},$$

respectively. Solving these equations, the MLEs of $\boldsymbol{\Psi}_{21}$ and $\boldsymbol{\Psi}_{22}$ are

$$\widetilde{\boldsymbol{\Psi}}_{21}=\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{1j}'\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'\right)^{-1},$$

and

$$\begin{aligned}\widetilde{\boldsymbol{\Psi}}_{22} &= \frac{1}{N_1}\sum_{j=1}^{N_1}\left(\boldsymbol{x}_{2j}-\widetilde{\boldsymbol{\Psi}}_{21}\boldsymbol{x}_{1j}\right)\left(\boldsymbol{x}_{2j}-\widetilde{\boldsymbol{\Psi}}_{21}\boldsymbol{x}_{1j}\right)'\\ &= \frac{1}{N_1}\left\{\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{2j}'-\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{1j}'\right)\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'\right)^{-1}\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{2j}'\right)\right\}.\end{aligned}$$

The MLE of $\boldsymbol{\Psi}$ is expressed as follows:

$$\widetilde{\boldsymbol{\Psi}}=\begin{pmatrix}\widetilde{\boldsymbol{\Psi}}_{11} & \widetilde{\boldsymbol{\Psi}}_{12}\\ \widetilde{\boldsymbol{\Psi}}_{21} & \widetilde{\boldsymbol{\Psi}}_{22}\end{pmatrix}=\begin{pmatrix}\widetilde{\boldsymbol{\Sigma}}_{11} & \widetilde{\boldsymbol{\Sigma}}_{11}^{-1}\widetilde{\boldsymbol{\Sigma}}_{12}\\ \widetilde{\boldsymbol{\Sigma}}_{21}\widetilde{\boldsymbol{\Sigma}}_{11}^{-1} & \widetilde{\boldsymbol{\Sigma}}_{22\cdot 1}\end{pmatrix}.$$

Since $\boldsymbol{\Psi}$ is one to one correspondence to $\boldsymbol{\Sigma}$, the MLE of $\boldsymbol{\Sigma}$ under H_0 is given by

$$\widetilde{\boldsymbol{\Sigma}}=\begin{pmatrix}\widetilde{\boldsymbol{\Sigma}}_{11} & \widetilde{\boldsymbol{\Sigma}}_{12}\\ \widetilde{\boldsymbol{\Sigma}}_{21} & \widetilde{\boldsymbol{\Sigma}}_{22}\end{pmatrix},$$

where

$$\widetilde{\boldsymbol{\Sigma}}_{11}=\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}',$$

$$\widetilde{\boldsymbol{\Sigma}}_{21}=\frac{1}{N}\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{1j}'\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'\right)^{-1}\sum_{j=1}^{N}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}',$$

$$\widetilde{\boldsymbol{\Sigma}}_{22}=\frac{1}{N_1}\left\{\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{2j}'-\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{2j}\boldsymbol{x}_{1j}'\right)\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{1j}'\right)^{-1}\left(\sum_{j=1}^{N_1}\boldsymbol{x}_{1j}\boldsymbol{x}_{2j}'\right)\right\}+\widetilde{\boldsymbol{\Sigma}}_{21}\widetilde{\boldsymbol{\Sigma}}_{11}^{-1}\widetilde{\boldsymbol{\Sigma}}_{12}.$$

3 Test statistics for mean vector

In this section, we provide a test statistic for testing the following hypothesis,

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

where $\boldsymbol{\mu}_0$ is known.

3.1 T^2 type statistic

In case of complete data, the statistic for this hypothesis is known as Hotelling's T^2 statistic. For a two-step monotone missing data, we can construct a test statistic based on Hotelling's T^2 statistic structure,

$$T^2 = (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)' \widehat{\boldsymbol{\Gamma}}^{-1} (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0), \tag{2}$$

where $\widehat{\boldsymbol{\Gamma}}$ is the estimator of $\boldsymbol{\Gamma}$ which is the covariance matrix of $\widehat{\boldsymbol{\mu}}$. We call this statistic the T^2 type statistic. Under H_0, since this T^2 type statistic is asymptotically distributed as χ^2 with degree of freedom p when N_1 and N_2 are large, H_0 is rejected when $T^2 > \chi^2_{p,\alpha}$. However, it seems that the upper percentiles of chi-square distribution is not good approximation for T^2 type statistic when sample size is not large. In case of complete data, Hotelling T^2 statistic has Fdistribution as follows;

$$T^2 \quad \sim \frac{(N-1)p}{N-p} F_{p,N-p}$$

Using this property, the upper percentile of T^2 type statistic of two-step monotone missing data should lie between the two upper percentiles of Hotelling T^2 statistic for non-missing data, that is, the data with N observations and the data with N_1 observations. As an approximation, we propose F^* for the upper percentile of T^2 type statistic

$$F^*_\alpha = cT^2_{p,N_1-p,\alpha} + (1-c)T^2_{p,N-p,\alpha}$$

where

$$c = \frac{N_2 p_2}{Np}, \quad T^2_{p,N_1-p,\alpha} = \frac{(N_1-1)p}{N_1-p} F_{\alpha;p,N_1-p}, \quad T^2_{p,N-p,\alpha} = \frac{(N-1)p}{N-p} F_{\alpha;p,N-p}$$

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

3.2 Likelihood ratio test statistic

Using the MLEs derived in Section 2, the likelihood ratio test statistic for the hypothesis can be obtained. Without loss of generality, we can put $\boldsymbol{\mu}_0 = \mathbf{0}$. The likelihood ratio test statistic, $-2\log\lambda$, is asymptotically distributed chi-square distribution with p degrees of freedom, where

$$\begin{aligned}\lambda &= \frac{L(\boldsymbol{\mu}_0, \widetilde{\boldsymbol{\Sigma}})}{L(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})} = \frac{L(\mathbf{0}, \widetilde{\boldsymbol{\Psi}})}{L(\widehat{\boldsymbol{\eta}}, \widehat{\boldsymbol{\Psi}})} \\ &= \frac{|\widehat{\boldsymbol{\Psi}}_{11}|^{N/2}}{|\widetilde{\boldsymbol{\Psi}}_{11}|^{N/2}} \times \frac{|\widehat{\boldsymbol{\Psi}}_{22}|^{N_1/2}}{|\widetilde{\boldsymbol{\Psi}}_{22}|^{N_1/2}}.\end{aligned}$$

4 Simultaneous confidence intervals

Using T^2 type statistic derived in Section 3.1, the simultaneous confidence intervals can be obtained. Suppose that we have a sample of N observations with two-step monotone missing pattern observations with mean vector $\boldsymbol{\mu}$, for any vector $\boldsymbol{a}' = [a_1, \ldots, a_p]$,

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

and from the distribution of the T^2 type statistic it follows that the probability statement

$$P[\text{all} \;\; T^2(\boldsymbol{a}) \leq t^2_{p,\alpha}] = 1 - \alpha$$

holds for all $\boldsymbol{a}$, where $t^2_{p,\alpha}$ denotes the upper 100α percentile of the T^2 type statistic distribution. Then the simultaneous confidence intervals can be obtained as follows;

$$\boldsymbol{a}'\widehat{\boldsymbol{\mu}} - \sqrt{\boldsymbol{a}'\widehat{\Gamma}\boldsymbol{a}t^2_{p,\alpha}} \leq \boldsymbol{a}'\boldsymbol{\mu} \leq \boldsymbol{a}'\widehat{\boldsymbol{\mu}} + \sqrt{\boldsymbol{a}'\widehat{\Gamma}\boldsymbol{a}t^2_{p,\alpha}}.$$

Since the asymptotic distribution of T^2 is χ^2, asymptotic simultaneous confidence intervals can be given using the upper 100α percentile of the chi-squared distribution, $\chi^2_{p,\alpha}$, instead of $t^2_{p,\alpha}$. As stated in Section 3.1, however, when sample size is not large, F^* is better approximation of the upper 100α percentile of the T^2 type statistic distribution. The asymptotic simultaneous confidence intervals can be improved as follows;

$$\boldsymbol{a}'\widehat{\boldsymbol{\mu}} - \sqrt{\boldsymbol{a}'\widehat{\Gamma}\boldsymbol{a}F^*_{\alpha}} \leq \boldsymbol{a}'\boldsymbol{\mu} \leq \boldsymbol{a}'\widehat{\boldsymbol{\mu}} + \sqrt{\boldsymbol{a}'\widehat{\Gamma}\boldsymbol{a}F^*_{\alpha}}.$$

5 Simulation studies

We compute the upper 100α percentiles of T^2 type statistic and likelihood ratio test statistic by Monte Carlo simulation based on replications of 10^5 times and compare those values to χ^2. We generate an artificial complete data set with N samples from $N_p(\mathbf{0}, \boldsymbol{I}_p)$, and delete p_2 consecutive data of N_2 samples to obtain a two-step monotone missing data. The upper percentiles of T^2 type statistic and F^* values are shown in Table 1. The T^2 type statistic is closer to the one of chi-square distribution with p-degrees of freedom as the sample size N_1 and N_2 are large. Meanwhile, F^* is much closer to the upper percentiles of T^2 type statistic even when the sample size is not large. Table 2 shows the same results when N_2 is fixed. Here, we need to note that the obtained upper percentiles of T^2 type statistic are slightly overestimated in a simulation when N_2 is very small relative to N_1. Table 3 and Table 4 present the results for comparing the type I error rates under T^2 type statistic when the null hypothesis H_0 is rejected using F^* and χ_p^2. The rejection regions are bigger than true values when the sample size is small, however, it can be seen that F^* gives smaller rejection regions comparing to χ_p^2. The upper percentiles of likelihood ratio test statistic and type I rates using χ_p^2 under likelihood ratio test statistic show the same tendency in Table 5 and Table 6.

Table 1: Upper percentiles of T^2 type statistic and F^* value

								$\alpha = 0.05$		$\alpha = 0.01$	
p	p_1	p_2	ρ_1	ρ_2	N	N_1	N_2	T^2	F^*	T^2	F^*
4	2	2	1/2	1/2	20	10	10	23.81	17.51	47.95	30.72
$\chi^2_{4,0.05} = 9.49$					40	20	20	13.47	12.13	20.87	18.31
$\chi^2_{4,0.01} = 13.28$					100	50	50	10.73	10.37	15.44	14.90
					200	100	100	10.06	9.91	14.30	14.04
					300	150	150	9.86	9.76	13.90	13.77
					400	200	200	9.78	9.69	13.75	13.65
			2/3	1/3	30	20	10	13.94	12.58	44.87	30.61
					60	40	20	11.27	10.81	16.47	15.71
					120	80	40	10.30	10.10	14.71	14.40
					240	160	80	9.90	9.79	13.96	13.81
					480	320	160	9.67	9.63	13.59	13.54
			1/3	2/3	30	10	20	22.16	17.22	21.75	19.17
					60	20	40	12.99	11.89	20.07	17.88
					120	40	80	10.90	10.51	15.83	15.15
					240	80	160	10.13	9.96	14.41	14.14
					480	160	320	9.80	9.72	13.79	13.69
8	4	4	1/2	1/2	20	10	10	510.79	201.40	2633.73	937.11
$\chi^2_{8,0.05} = 15.51$					40	20	20	31.42	25.43	49.03	37.11
$\chi^2_{8,0.01} = 20.09$					100	50	50	19.19	18.23	26.00	24.43
					200	100	100	17.15	16.75	22.60	22.03
					300	150	150	16.53	16.31	21.64	21.34
					400	200	200	16.26	16.10	21.26	21.01
			2/3	1/3	30	20	10	33.29	27.07	52.30	39.84
					60	40	20	21.07	19.70	29.13	26.82
					120	80	40	17.86	17.35	23.76	22.99
					240	160	80	16.60	16.38	21.72	21.45
					480	320	160	16.03	15.93	20.88	20.75
			1/3	2/3	30	10	20	460.49	249.30	52.13	39.84
					60	20	40	29.68	24.87	46.58	36.39
					120	40	80	19.93	18.75	27.18	25.32
					240	80	160	17.32	16.93	22.88	22.32
					480	160	320	16.33	16.18	21.34	21.13
20	10	10	1/2	1/2	100	50	50	54.91	47.39	71.55	60.08
$\chi^2_{20,0.05} = 31.41$					200	100	100	39.48	37.56	48.57	45.95
$\chi^2_{20,0.05} = 37.57$					300	150	150	36.25	35.23	44.07	42.74
					400	200	200	34.88	34.18	42.23	41.30
					500	250	250	34.11	33.58	41.23	40.49
					600	300	300	33.66	33.20	40.56	39.97
			2/3	1/3	240	160	80	36.48	35.54	44.35	43.15
					480	320	160	33.74	33.35	40.66	40.17
					960	640	320	32.52	32.35	39.02	38.83
					1920	1280	640	31.99	31.87	38.27	38.19
			1/3	2/3	240	80	160	41.07	38.79	50.94	47.72
					480	160	320	35.35	34.59	42.87	41.87
					960	320	640	33.24	32.90	39.97	39.57
					1920	640	1280	32.32	32.13	38.77	38.54

Table 2: Upper percentiles of T^2 type statistic and F^* value when N_2 is fixed

p p_1 p_2	N	N_1	N_2	$\alpha = 0.05$ T^2	$\alpha = 0.05$ F^*	$\alpha = 0.01$ T^2	$\alpha = 0.01$ F^*
4 2 2	20	10	10	23.81	17.51	47.95	30.72
$\chi^2_{4,0.05} = 9.49$	30	20	10	13.94	12.58	21.75	19.17
$\chi^2_{4,0.01} = 13.28$	60	50	10	11.04	10.71	16.09	15.53
	110	100	10	10.26	10.11	14.62	14.41
	60	10	50	20.95	17.57	42.69	31.90
	70	20	50	12.90	11.85	19.87	17.82
	100	50	50	10.73	10.37	15.44	14.90
	150	100	50	10.14	9.97	14.44	14.16
	110	10	100	20.48	17.88	41.43	32.81
	120	20	100	12.54	11.82	19.30	17.79
	150	50	100	10.57	10.28	15.21	14.73
	200	100	100	10.06	9.91	14.30	14.04
8 4 4	20	10	10	510.79	201.40	2648.20	937.11
$\chi^2_{8,0.05} = 15.51$	30	20	10	33.29	27.07	52.13	39.84
$\chi^2_{8,0.01} = 20.09$	60	50	10	20.14	19.34	27.42	26.23
	110	100	10	17.61	17.37	23.38	23.02
	60	10	50	419.47	301.80	2174.61	1505.83
	70	20	50	29.29	24.86	45.76	36.45
	100	50	50	19.19	18.23	25.89	24.43
	150	100	50	17.34	16.95	22.89	22.35
	110	10	100	401.03	326.45	2094.58	1638.67
	120	20	100	28.25	25.06	43.94	37.01
	150	50	100	18.76	17.96	25.26	24.00
	200	100	100	17.15	16.75	22.62	22.03
20 10 10	100	50	50	54.91	47.39	71.55	60.08
$\chi^2_{20,0.05} = 31.41$	150	100	50	40.43	38.58	49.88	47.36
$\chi^2_{20,0.01} = 37.57$	200	150	50	37.19	36.25	45.31	44.14
	150	50	100	52.38	46.28	68.26	58.66
	200	100	100	39.48	37.56	48.57	45.95
	250	150	100	36.62	35.57	44.53	43.21

Table 3: Type I error rate using F^* and χ^2 values under T^2 type statistic

								$\alpha = 0.05$		$\alpha = 0.01$	
p	p_1	p_2	ρ_1	ρ_2	N	N_1	N_2	F^*	χ^2	F^*	χ^2
4	2	2	1/2	1/2	20	10	10	0.094	0.264	0.029	0.156
					40	20	20	0.068	0.131	0.017	0.052
					100	50	50	0.057	0.076	0.012	0.021
					200	100	100	0.053	0.062	0.011	0.015
					300	150	150	0.052	0.058	0.011	0.013
					400	200	200	0.052	0.056	0.011	0.012
			2/3	1/3	30	20	10	0.068	0.140	0.017	0.058
					60	40	20	0.058	0.088	0.013	0.027
					120	80	40	0.054	0.067	0.011	0.017
					240	160	80	0.052	0.058	0.011	0.013
					480	320	160	0.051	0.054	0.010	0.011
			1/3	2/3	30	10	20	0.085	0.243	0.025	0.139
					60	20	40	0.066	0.121	0.016	0.047
					120	40	80	0.057	0.080	0.012	0.023
					240	80	160	0.053	0.064	0.011	0.015
					480	160	320	0.052	0.057	0.010	0.012
8	4	4	1/2	1/2	20	10	10	0.120	0.773	0.028	0.690
					40	20	20	0.094	0.334	0.029	0.176
					100	50	50	0.063	0.118	0.014	0.040
					200	100	100	0.056	0.079	0.012	0.021
					300	150	150	0.053	0.068	0.011	0.017
					400	200	200	0.053	0.063	0.011	0.015
			2/3	1/3	30	20	10	0.094	0.334	0.027	0.199
					60	40	20	0.066	0.154	0.016	0.061
					120	80	40	0.057	0.093	0.012	0.027
					240	160	80	0.053	0.069	0.011	0.017
					480	320	160	0.052	0.059	0.010	0.013
			1/3	2/3	30	10	20	0.089	0.742	0.019	0.653
					60	20	40	0.086	0.280	0.025	0.156
					120	40	80	0.065	0.015	0.015	0.048
					240	80	160	0.056	0.083	0.012	0.023
					480	160	320	0.052	0.064	0.011	0.015
20	10	10	1/2	1/2	100	50	50	0.104	0.424	0.030	0.257
					200	100	100	0.069	0.178	0.016	0.068
					300	150	150	0.061	0.122	0.013	0.039
					400	200	200	0.058	0.099	0.012	0.028
					500	250	250	0.056	0.088	0.012	0.023
					600	300	300	0.055	0.081	0.012	0.020
			2/3	1/3	240	160	80	0.060	0.126	0.013	0.041
					480	320	160	0.054	0.082	0.011	0.021
					960	640	320	0.052	0.064	0.011	0.015
					1920	1280	640	0.051	0.057	0.010	0.012
			1/3	2/3	240	80	160	0.071	0.206	0.017	0.086
					480	160	320	0.058	0.107	0.013	0.032
					960	320	640	0.054	0.074	0.011	0.018
					1920	640	1280	0.052	0.062	0.011	0.014

Table 4: Type I error rate using F^* and χ^2 values under T^2 type statistic when N_2 is fixed

						$\alpha = 0.05$		$\alpha = 0.01$	
p	p_1	p_2	N	N_1	N_2	F^*	χ^2	F^*	χ^2
4	2	2	20	10	10	0.094	0.264	0.029	0.156
			30	20	10	0.068	0.140	0.017	0.058
			60	50	10	0.055	0.082	0.012	0.024
			110	100	10	0.052	0.066	0.011	0.016
			60	10	50	0.072	0.223	0.020	0.125
			70	20	50	0.064	0.119	0.015	0.045
			100	50	50	0.057	0.076	0.012	0.021
			150	100	50	0.053	0.064	0.011	0.016
			110	10	100	0.066	0.214	0.018	0.118
			120	20	100	0.060	0.112	0.014	0.041
			150	50	100	0.055	0.073	0.012	0.020
			200	100	100	0.053	0.062	0.011	0.015
8	4	4	20	10	10	0.120	0.773	0.028	0.690
			30	20	10	0.094	0.334	0.027	0.199
			60	50	10	0.059	0.136	0.013	0.050
			110	100	10	0.054	0.088	0.011	0.025
			60	10	50	0.068	0.710	0.014	0.619
			70	20	50	0.083	0.274	0.023	0.151
			100	50	50	0.063	0.118	0.014	0.040
			150	100	50	0.056	0.083	0.012	0.023
			110	10	100	0.061	0.697	0.013	0.605
			120	20	100	0.073	0.254	0.019	0.137
			150	50	100	0.061	0.110	0.014	0.036
			200	100	100	0.056	0.079	0.012	0.021
20	10	10	100	50	50	0.104	0.424	0.030	0.257
			150	100	50	0.068	0.196	0.016	0.079
			200	150	50	0.059	0.138	0.013	0.047
			150	50	100	0.094	0.383	0.026	0.221
			200	100	100	0.069	0.178	0.016	0.068
			250	150	100	0.061	0.127	0.013	0.042

Table 5: Upper percentiles of LRT statistic and type I error rate using χ^2 value

								$\alpha = 0.05$		$\alpha = 0.01$	
p	p_1	p_2	ρ_1	ρ_2	N	N_1	N_2	LRT	Type I	LRT	Type I
4	2	2	1/2	1/2	20	10	10	13.33	0.146	18.90	0.051
$\chi^2_{4,0.05} = 9.49$					40	20	20	10.96	0.084	15.36	0.022
$\chi^2_{4,0.01} = 13.28$					100	50	50	10.00	0.061	13.99	0.014
					200	100	100	9.73	0.055	13.64	0.012
					300	150	150	9.65	0.053	13.48	0.011
					400	200	200	9.62	0.053	13.45	0.011
			2/3	1/3	30	20	10	11.03	0.086	18.65	0.022
					60	40	20	10.17	0.065	14.22	0.015
					120	80	40	9.81	0.057	13.73	0.012
					240	160	80	9.66	0.054	13.50	0.011
					480	320	160	9.56	0.051	13.37	0.010
			1/3	2/3	30	10	20	13.14	0.141	18.65	0.048
					60	20	40	10.86	0.081	15.24	0.021
					120	40	80	10.10	0.063	14.17	0.014
					240	80	160	9.79	0.056	13.70	0.012
					480	160	320	9.64	0.053	13.47	0.011
8	4	4	1/2	1/2	20	10	10	42.05	0.570	58.39	0.396
$\chi^2_{8,0.05} = 15.51$					40	20	20	20.59	0.162	26.88	0.057
$\chi^2_{8,0.01} = 20.09$					100	50	50	17.02	0.078	22.11	0.019
					200	100	100	16.24	0.063	21.01	0.014
					300	150	150	15.95	0.058	20.67	0.012
					400	200	200	15.84	0.056	20.56	0.012
			2/3	1/3	30	20	10	20.79	0.168	27.14	0.059
					60	40	20	17.58	0.090	22.86	0.024
					120	80	40	16.46	0.067	21.33	0.015
					240	160	80	15.97	0.058	20.64	0.012
					480	320	160	15.72	0.054	20.37	0.011
			1/3	2/3	30	10	20	41.77	0.558	27.11	0.385
					60	20	40	20.37	0.156	26.68	0.054
					120	40	80	17.42	0.087	22.57	0.022
					240	80	160	16.33	0.065	21.17	0.014
					480	160	320	15.89	0.057	20.59	0.012
20	10	10	1/2	1/2	100	50	50	40.24	0.217	48.28	0.081
$\chi^2_{20,0.05} = 31.41$					200	100	100	34.97	0.104	41.86	0.028
$\chi^2_{20,0.01} = 37.57$					300	150	150	33.63	0.081	40.23	0.020
					400	200	200	33.03	0.072	39.52	0.017
					500	250	250	32.69	0.067	39.14	0.015
					600	300	300	32.48	0.064	38.87	0.014
			2/3	1/3	240	160	80	33.60	0.081	40.15	0.019
					480	320	160	32.45	0.064	38.78	0.014
					960	640	320	31.90	0.056	38.14	0.012
					1920	1280	640	31.68	0.053	37.84	0.011
			1/3	2/3	240	80	160	35.87	0.121	42.92	0.035
					480	160	320	33.37	0.077	39.91	0.018
					960	320	640	32.36	0.062	38.70	0.014
					1920	640	1280	31.90	0.056	38.15	0.012

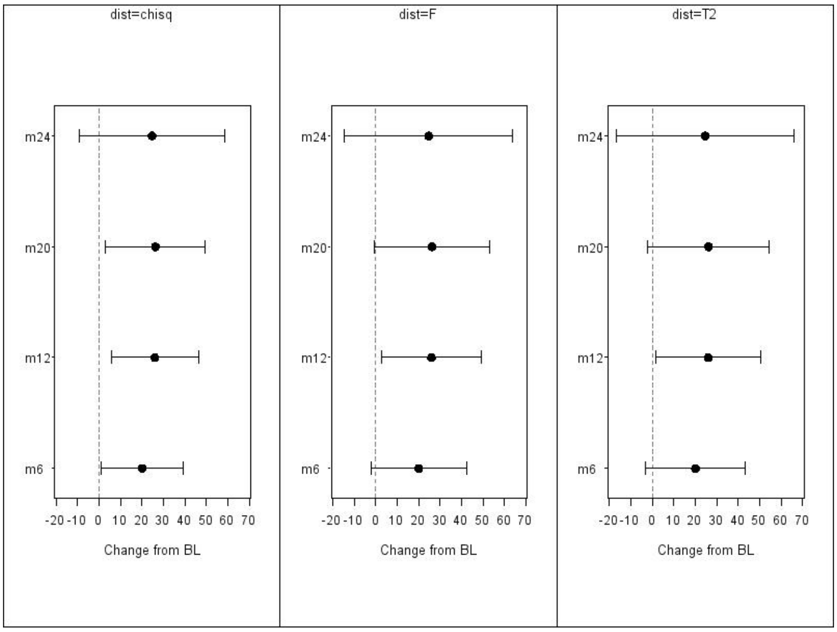
Table 6: Upper percentiles of LRT statistic and type I error rate using χ^2 value when N_2 is fixed

						$\alpha = 0.05$		$\alpha = 0.01$	
p	p_1	p_2	N	N_1	N_2	LRT	Type I	LRT	Type I
4	2	2	20	10	10	13.33	0.146	18.90	0.051
$\chi^2_{4,0.05} = 9.49$			30	20	10	11.03	0.086	15.46	0.022
$\chi^2_{4,0.01} = 13.28$			60	50	10	10.08	0.063	14.14	0.014
			110	100	10	9.80	0.056	13.70	0.012
			60	10	50	13.03	0.137	18.52	0.047
			70	20	50	10.87	0.081	15.23	0.020
			100	50	50	10.00	0.061	13.99	0.014
			150	100	50	9.76	0.055	13.68	0.012
			110	10	100	12.97	0.134	18.41	0.045
			120	20	100	10.78	0.079	15.14	0.020
			150	50	100	9.97	0.060	13.96	0.013
			200	100	100	9.73	0.055	13.64	0.012
8	4	4	20	10	10	42.05	0.570	58.33	0.396
$\chi^2_{8,0.05} = 15.51$			30	20	10	20.79	0.168	27.11	0.059
$\chi^2_{8,0.01} = 20.09$			60	50	10	17.20	0.082	22.28	0.020
			110	100	10	16.30	0.064	21.14	0.014
			60	10	50	41.50	0.547	57.64	0.377
			70	20	50	20.34	0.156	26.60	0.053
			100	50	50	17.02	0.078	22.05	0.019
			150	100	50	16.26	0.063	21.06	0.014
			110	10	100	41.34	0.544	57.60	0.374
			120	20	100	20.22	0.151	26.44	0.051
			150	50	100	16.95	0.077	21.99	0.019
			200	100	100	16.24	0.063	21.05	0.014
20	10	10	100	50	50	40.24	0.217	48.28	0.081
$\chi^2_{20,0.05} = 31.41$			150	100	50	35.11	0.107	42.00	0.029
$\chi^2_{20,0.01} = 37.57$			200	150	50	33.82	0.084	40.40	0.021
			150	50	100	39.92	0.207	47.99	0.077
			200	100	100	34.97	0.104	41.86	0.028
			250	150	100	33.69	0.082	40.24	0.020

6 Numerical example

We will illustrate how F^* improve approximation of simultaneous confidence intervals using a numerical example. The sample data is a serum cholesterol values under treatment at 5 different time points, baseline, month 6, 12, 20 and 24, by Wei and Lachin(1984). The original data has 36 complete observations. We randomly chose 30 observations from this data and calculate the difference from baseline at each post-baseline time point. We moreover chose 10 observations randomly and delete the data at month 20 and 24. Hence, we got a two-step monotone missing data with 20 complete observations at 4 time points and 10 incomplete observations without the last 2 time points, i.e. $N_1 = 20, N_2 = 10$ and $p_1 = p_2 = 2$. On this data, we obtained $T^2 = 19.62$ for the hypothesis $H_0 : \boldsymbol{\mu} = 0$. Since $t^2_{4,0.05} = 13.94$ from the simulation study in section 5, $F^*_{0.05} = 12.58$ and $\chi^2_{4,0.05} = 9.46$, the hypothesis is rejected by both percentiles at significant level of 0.05. 95 % simultaneous confidence intervals for the change from baseline at each time point are shown in Figure 1. Considering the confidence intervals using T^2 are true values, the results shows that using χ^2 distribution may lead to incorrect conclusions. In contrast, the confidence intervals using F^* tend to have the same results as T^2.

Figure 1: Mean and 95 % simultaneous confidence interval for change from baseline



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] 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.

[3] Koizumi, K. and Seo, T. (2009a). Testing equality of two mean vectors and simultaneous confidence intervals in repeated measures with missing data. *Journal of the Japanese Society of Computational Statistics*, **22**.

[4] Koizumi, K. and Seo, T. (2009b). Simultaneous confidence intervals among k mean vectors in repeated measures with missing data. *American Journal of Mathematical and Management Sciences*.

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

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

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

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

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