Tests for mean vectors
with two-step monotone missing data
or unequal covariance matrices
（２-ステップ単調欠測データまたは
異なる共分散行列をもつ平均ベクトルの検定）

Tamae Kawasaki
（川崎　玉恵）

Department of Mathematical Science
for Information Sciences
Graduate School of Science, Tokyo University of Science

Acknowledgements

I would like to express his sincere gratitude to Professor Takashi Seo for his enormous support and valuable comments throughout the course of this study. I am also greatly indebted to Professors Hiroki Hashiguchi, Keiko Kotani, Etsuo Miyaoka and Hiroshi Yabe for their careful readings. Further, I am grateful to Doctor Takahiro Nishiyama, Doctor Nobumichi Shutoh, and Doctor Hidetoshi Murakami, whose comments and suggestions have helped me very much. I also thank all members of Seo Laboratory at Tokyo University of Science for their helpful comments and fruitful discussions. I would also like to thank all people in the Department of Mathematical Science for Information Science who helped me toward the successful completion of the dissertation.

Finally, I must also thank my parents, younger sister and all of my friends for their encouragement and dedicated supports.

Contents

Chapter 1

Introduction

To consider the tests for mean vectors is a fundamental problem in multivariate statistical analysis. In particular, we consider the following two topics.

The first research topic is focused on the one-sample problem of testing for the mean vector with two-step monotone missing data. In almost all statistical analyses, missing data is a constantly occurring problem. Many statistical methods have been developed to analyze data with missing values. In the case of general k-step monotone missing data, many difficult problems remain unsolved. For simplicity, we assume that $k = 2$. We derive the asymptotic expansion of Hotelling's T^2 type statistic for the case where the sample size is large. The asymptotic first two moments are obtained using stochastic expansion. We also propose the Bartlett and modified Bartlett corrected statistics for two-step monotone missing data. Simulation studies are performed to demonstrate the performance of the proposed results by a Monte Carlo simulation.

In the second research topic, testing the equality of two mean vectors with unequal covariance matrices. In the case of equal covariance matrices, we can use Hotelling's T^2 statistic, which follows the F distribution under the null hypothesis. Meanwhile, in the case of unequal covariance matrices, the test statistic does not follow the F distribution, and it is also difficult to derive the exact distribution. Even in the case of univariate, mean comparison with unequal variances is intrinsically difficult, and is well known as the Behrens-Fisher problem. We provide two approximate solutions: F statistic by adjusting the degrees of freedom, the bias corrected statistic. Asymptotic expansions up to the term of order N^{-2} for the first and second moments of the test statistic are given, where N is the total sample size minus two. By simulations, we compare the two proposed results as well as some existing procedures.

The remainder of this paper is organized as follows. Chapter 2 presents the results derived by Kawasaki and Seo (2016), i.e., the first research topic. First of all, we review the missing mechanism and two-step monotone missing data. In the text that follows, we discuss the asymptotic expansion of maximum likelihood estimators (MLEs) and T^2 type statistics and propose the transformation statistics. Chapter 3 presents the results derived by Kawasaki and Seo (2015), i.e., the second research topic. First of all, we review the Behrens-Fisher problem. In the text the follows, we explain the approximate degrees of freedom and discuss the main results in more detail. The accuracy of our results is investigated by Monte Carlo simulation in each chapter. In the Appendix, we present certain formulas used to derive the main results.

Chapter 2

Bias correction with two-step monotone missing data

In this chapter, we consider the one-sample problem of testing for the multivariate mean vector with two-step monotone missing data. We derive the Bartlett correction and modified Bartlett correction statistics by using the stochastic expansion of the T^2 type statistic with two-step monotone missing data. In order to derive the asymptotic expectation and variance, the stochastic expansion of Hotelling's T^2 type statistic is calculated. Further, we propose Bartlett correction statistics for two-step monotone missing data. Using Monte Carlo simulations, we investigate the first and second moments, the variance, the mean square error (MSE), and the upper percentiles for the T^2 type statistic and the transformation statistics. Moreover, we show the type I error rate when the null hypothesis is rejected using χ_p^2 under the simulated transformation statistics.

2.1 Missing data

In almost all statistical analyses, missing data is a constantly occurring problem. Many statistical methods have been developed to analyze data with missing values. Anderson (1957) developed an approach to derive the maximum likelihood estimators (MLEs) of the mean and covariance matrix by solving the likelihood equation for monotone missing data. Jinadasa and Tracy (1992) proposed the MLEs in the case of a k-step monotone pattern using matrix derivatives. Kanda and Fujikoshi (1998) showed the properties of the MLEs based on k-step monotone missing samples for $k = 2, 3$ and a general k. Srivastava (1985) derived the MLEs for missing data from a growth curve and several other models and developed likelihood ratio tests. By using the idea of Srivastava (1985), Srivastava and Carter (1986) and Shutoh et al. (2010) obtained the MLEs for the mean vector and covariance matrix using the Newton-Raphson method with general missing data. However, these results by the Newton-Raphson method are numerically rather than analytically iterative solutions.

In the case of a two-step monotone missing data pattern, Seko et al. (2012) derived the T^2 type statistic of testing for the mean vector and its approximate upper percentile. Krishnamoorthy and Pannala (1999) provided an accurate and simple approach to construct a confidence region for a normal mean vector. We note that among the many other papers with a monotone missing data pattern, Bhargava (1962) derived likelihood ratio tests and their approximate null distributions. Seo and Srivastava (2000) derived a test of equality of means and simultaneous confidence intervals in a one sample problem under a covariance matrix with intra-class correlation. For the case of a two-sample problem, Seko et al. (2011) derived the T^2 type statistic of testing for two mean vectors and their approximate upper percentiles with a two-step monotone missing data pattern. Yu et al. (2006) obtained the T^2 type statistic and derived F-approximations for the distribution of the T^2 type statistic in the case of three-step monotone missing data. Wu et al. (2006) developed a pooled component test statistic to test the equality of two mean vectors.

Any analysis of the data is then dependent on assumptions about the missing data mechanism. However, because the causes of missing data are not our concern, we ignore the process that causes missing data by assuming that the data are missing at random (MAR). Rubin (1976) pointed out that in the case of MAR and the observed data that are observed at random (OAR) or, more simply, the missing data that are missing completely at random (MCAR), the missing data mechanism can be ignored for likelihood based inferences. The details of MAR, OAR, and MCAR follow from Little and Rubin (1987). In this paper, we assume that the data are MCAR.

First, we consider a one sample problem with a monotone missing data pattern. Let $\boldsymbol{y}$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$ and let $\boldsymbol{y}^{[i]}$ be the vector of the first $p_1 + p_2 + \cdots + p_{k-i+1}$ dimension of $\boldsymbol{y}$. Let $Y^{[1]} = (\boldsymbol{y}_1^{[1]}, \boldsymbol{y}_2^{[1]}, \ldots, \boldsymbol{y}_{n_1}^{[1]})'$ represent the fully observed variables with p $(= p_1 + p_2 + \cdots + p_k)$ dimensions, $Y^{[2]} = (\boldsymbol{y}_1^{[2]}, \boldsymbol{y}_2^{[2]}, \ldots, \boldsymbol{y}_{n_2}^{[2]})'$ the incompletely observed variable with the fewest missing values with $(p_1 + p_2 + \cdots + p_{k-1})$ dimensions, $Y^{[3]} = (\boldsymbol{y}_1^{[3]}, \boldsymbol{y}_2^{[3]}, \ldots, \boldsymbol{y}_{n_3}^{[3]})'$ the variable with the second fewest missing values and $(p_1 + p_2 + \cdots + p_{k-2})$ dimensions, … , and $Y^{[k]} = (\boldsymbol{y}_1^{[k]}, \boldsymbol{y}_2^{[k]}, \ldots, \boldsymbol{y}_{n_k}^{[k]})'$ the variable with the most missing values and p_1 dimensions, with $n_i > (p_1 + p_2 + \cdots + p_{k-i+1})$. If $\boldsymbol{y}_j^{[i]}$ denotes the j-th observation on $\boldsymbol{y}^{[i]}$, the sample of observations $\boldsymbol{y}_j^{[i]}$, $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i$, is called a monotone missing sample (Srivastava and Carter (1983)) or a monotone missing data pattern (Little and Rubin (1987)). Such a sample is called k-step monotone missing data.

Next, we consider two-step monotone missing data. Let $\boldsymbol{x}_1^{[1]}, \boldsymbol{x}_2^{[1]}, \ldots, \boldsymbol{x}_{n_1}^{[1]}$ be distributed as the multivariate normal $N_p(\boldsymbol{\mu}, \Sigma)$ and $\boldsymbol{x}_1^{[2]}, \boldsymbol{x}_2^{[2]}, \ldots, \boldsymbol{x}_{n_2}^{[2]}$ be distributed as the multivariate normal $N_{p_1}(\boldsymbol{\mu}_1, \Sigma_{11})$, where each $\boldsymbol{x}_j^{[1]} = (x_{j,1}^{[1]}, x_{j,2}^{[1]}, \ldots, x_{j,p}^{[1]})', j = 1, 2, \ldots, n_1$ is $p \times 1$ and each $\boldsymbol{x}_j^{[2]} = (x_{j,1}^{[2]}, x_{j,2}^{[2]}, \ldots, x_{j,p_1}^{[2]})', j = 1, 2, \ldots, n_2$ is $p_1 \times 1$, and

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

To be specific, two-step monotone missing data (the case of $k = 2$) are drawn from a multivariate normal population of the form

$$\begin{pmatrix} x^{[1]}_{1,1} & x^{[1]}_{1,2} & \cdots & x^{[1]}_{1,p_1} & x^{[1]}_{1,p_1+1} & \cdots & x^{[1]}_{1,p} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x^{[1]}_{n_1,1} & x^{[1]}_{n_1,2} & \cdots & x^{[1]}_{n_1,p_1} & x^{[1]}_{n_1,p_1+1} & \cdots & x^{[1]}_{n_1,p} \\ x^{[2]}_{1,1} & x^{[2]}_{1,2} & \cdots & x^{[2]}_{1,p_1} & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x^{[2]}_{n_2,1} & x^{[2]}_{n_2,2} & \cdots & x^{[2]}_{n_2,p_1} & * & \cdots & * \end{pmatrix},$$

where $n = n_1 + n_2$ and $p = p_1 + p_2$. "$*$" 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 dimensions. We partition $\boldsymbol{x}^{[1]}_j$ into a $p_1 \times 1$ random vector and a $p_2 \times 1$ random vector as $\boldsymbol{x}^{[1]}_j = (\boldsymbol{x}^{[1]\prime}_{1j}, \boldsymbol{x}^{[1]\prime}_{2j})'$, where $\boldsymbol{x}^{[1]}_{ij} : p_i \times 1, i = 1, 2, j = 1, 2, \ldots, n_1$.

We define the sample means as

$$\overline{\boldsymbol{x}}^{[1]}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} \boldsymbol{x}^{[1]}_{1j}, \quad \overline{\boldsymbol{x}}^{[1]}_2 = \frac{1}{n_1}\sum_{j=1}^{n_1} \boldsymbol{x}^{[1]}_{2j}, \quad \overline{\boldsymbol{x}}^{[2]} = \frac{1}{n_2}\sum_{j=1}^{n_2} \boldsymbol{x}^{[2]}_{j},$$

where $\overline{\boldsymbol{x}}^{[1]} = (\overline{\boldsymbol{x}}^{[1]\prime}_1, \overline{\boldsymbol{x}}^{[1]\prime}_2)'$, and the sample covariance matrices are given as

$$S^{[1]} = \frac{1}{N_1}\sum_{j=1}^{n_1} (\boldsymbol{x}^{[1]}_j - \overline{\boldsymbol{x}}^{[1]})(\boldsymbol{x}^{[1]}_j - \overline{\boldsymbol{x}}^{[1]})' = \begin{pmatrix} S^{[1]}_{11} & S^{[1]}_{12} \\ S^{[1]}_{21} & S^{[1]}_{22} \end{pmatrix},$$

$$S^{[2]} = \frac{1}{N_2}\sum_{j=1}^{n_2} (\boldsymbol{x}^{[2]}_j - \overline{\boldsymbol{x}}^{[2]})(\boldsymbol{x}^{[2]}_j - \overline{\boldsymbol{x}}^{[2]})',$$

where $N_i = n_i - 1$.

2.2 Hypothesis and test statistics

The problem of testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, where $\boldsymbol{\mu}_0$ is a specified vector, has been studied extensively for two-step monotone missing data. Seko et al. (2012) obtained the approximate upper percentiles of Hotelling's T^2 type statistic and the likelihood ratio test statistic for two-step monotone missing data in a one sample problem.

Let the MLEs of $\boldsymbol{\mu}$ and Σ be denoted by $\widehat{\boldsymbol{\mu}}$ and $\widehat{\Sigma}$, respectively, which are partitioned in the same manner as $\boldsymbol{\mu}$ and Σ. We assume that the observation vectors are distributed as $N_p(\boldsymbol{\mu}, \Sigma)$ and $n_1 > p$, which is a necessary and sufficient condition for the existence and uniqueness of the MLEs of $\boldsymbol{\mu}$ and Σ. Anderson and Olkin (1985) derived the MLEs of $\boldsymbol{\mu}$ and Σ (see Kanda and Fujikoshi (1998), Chang and Richards (2009), and Seko et al. (2012)) as

$$\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{n}(n_1\overline{\boldsymbol{x}}_1^{[1]} + n_2\overline{\boldsymbol{x}}^{[2]}) \\ \overline{\boldsymbol{x}}_2^{[1]} - \widehat{\Sigma}_{21}\widehat{\Sigma}_{11}^{-1}(\overline{\boldsymbol{x}}_1^{[1]} - \widehat{\boldsymbol{\mu}}_1) \end{pmatrix},$$

$$\widehat{\Sigma} = \begin{pmatrix} \frac{1}{n}\left(W_{11}^{[1]} + W^{[2]}\right) & \widehat{\Sigma}_{11}\left(W_{11}^{[1]}\right)^{-1} W_{12}^{[1]} \\ W_{21}^{[1]}\left(W_{11}^{[1]}\right)^{-1}\widehat{\Sigma}_{11} & \frac{1}{n_1}W_{22\cdot1}^{[1]} + \widehat{\Sigma}_{21}\widehat{\Sigma}_{11}^{-1}\widehat{\Sigma}_{12} \end{pmatrix},$$

where

$$W^{[1]} = N_1S^{[1]} = \begin{pmatrix} W_{11}^{[1]} & W_{12}^{[1]} \\ W_{21}^{[1]} & W_{22}^{[1]} \end{pmatrix}, \quad W_{22\cdot1}^{[1]} = W_{22}^{[1]} - W_{21}^{[1]}\left(W_{11}^{[1]}\right)^{-1} W_{12}^{[1]},$$

$$W^{[2]} = N_2S^{[2]} + \frac{n_1n_2}{n}(\overline{\boldsymbol{x}}_1^{[1]} - \overline{\boldsymbol{x}}^{[2]})(\overline{\boldsymbol{x}}_1^{[1]} - \overline{\boldsymbol{x}}^{[2]})'.$$

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

$$\mathrm{E}[\widehat{\boldsymbol{\mu}}] = \boldsymbol{\mu}, \quad \mathrm{Cov}[\widehat{\boldsymbol{\mu}}] = \begin{pmatrix} \frac{1}{n}\Sigma_{11} & \frac{1}{n}\Sigma_{12} \\ \frac{1}{n}\Sigma_{21} & \mathrm{Cov}[\widehat{\boldsymbol{\mu}}_2] \end{pmatrix},$$

respectively, where

$$\mathrm{Cov}[\widehat{\boldsymbol{\mu}}_2] = \frac{1}{n_1}\left(\Sigma_{22} - \frac{n_2}{n}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right) + \frac{n_2p_1}{nn_1(n_1-p_1-2)}\Sigma_{22\cdot1}, \ (n_1 > p_1 + 2).$$

For two-step monotone missing data, it is easy to construct a test statistic based on Hotelling's T^2 statistic structure:

$$T^2 = (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)'\left\{\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})\right\}^{-1}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0),$$

where

$$\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}] = \begin{pmatrix} \dfrac{1}{n}\widehat{\Sigma}_{11} & \dfrac{1}{n}\widehat{\Sigma}_{12} \\ \dfrac{1}{n}\widehat{\Sigma}_{21} & \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2] \end{pmatrix}.$$

Note that under the null hypothesis, the T^2 type statistic is asymptotically distributed as χ^2 with p degrees of freedom when $n_1, n \to \infty$ with $n_1/n \to \delta \in (0, 1]$ (see Chang and Richards (2009)). Kanda and Fujikoshi (1998) considered two- and three-step monotone missing data and obtained the asymptotic expansion of the distributions of $\widehat{\boldsymbol{\mu}}$ and $\widehat{\Sigma}$ when n_1 and n_2 are large, and the MLEs of $\boldsymbol{\mu}$ and the usual transformed matrix of Σ are given in explicit forms for a general k.

2.3 Expectation of T^2 type statistic

In this section, we consider the asymptotic expansion of the first and second order moments of T^2 in a situation when

$$\gamma_i = N_i/N \to \text{positive constants}$$

as n_i's tend to infinity, where $N_i = n_i - 1$ and $N = N_1 + N_2$. In our derivations, we consider the stochastic expansions of $\widehat{\boldsymbol{\mu}}$ and $\widehat{\Sigma}$ in terms of

$$\boldsymbol{z}^{[1]} = \begin{pmatrix} \boldsymbol{z}_1^{[1]} \\ \boldsymbol{z}_2^{[1]} \end{pmatrix} = \sqrt{N_1} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{[1]} \\ \overline{\boldsymbol{x}}_2^{[1]} \end{pmatrix}, \quad \boldsymbol{z}^{[2]} = \sqrt{N_2}\, \overline{\boldsymbol{x}}^{[2]},$$

$$V^{[1]} = \sqrt{N_1}(S^{[1]} - I_p) = \begin{pmatrix} V_{11}^{[1]} & V_{12}^{[1]} \\ V_{21}^{[1]} & V_{22}^{[1]} \end{pmatrix}, \quad V^{[2]} = \sqrt{N_2}(S^{[2]} - I_{p_1}).$$

We note that the stochastic expansions are derived under $\boldsymbol{\mu} = \boldsymbol{\mu}_0 = \mathbf{0}$ and $\Sigma = I_p$. T^2 can be written as

$$T^2 = (\sqrt{N}\widehat{\boldsymbol{\mu}})' \left\{ N\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}}) \right\}^{-1} (\sqrt{N}\widehat{\boldsymbol{\mu}}),$$

and $\sqrt{N}\widehat{\boldsymbol{\mu}}$ and $N\widehat{\mathrm{Cov}}\,(\widehat{\boldsymbol{\mu}})$ can be expanded as

$$\sqrt{N}\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \sqrt{\gamma_1}\boldsymbol{z}_1^{[1]} + \sqrt{\gamma_2}\boldsymbol{z}^{[2]} \\ \dfrac{1}{\sqrt{\gamma_1}}\boldsymbol{z}_2^{[1]} \end{pmatrix} + \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{0} \\ \dfrac{\sqrt{\gamma_2}}{\sqrt{\gamma_1}}V_{21}^{[1]}\boldsymbol{z}^{[2]} - \dfrac{\gamma_2}{\gamma_1}V_{21}^{[1]}\boldsymbol{z}_1^{[1]} \end{pmatrix}$$
$$+ \frac{1}{N} \begin{pmatrix} \left(\dfrac{1}{\sqrt{\gamma_1}} - 2\sqrt{\gamma_1}\right)\boldsymbol{z}_1^{[1]} + \left(\dfrac{1}{\sqrt{\gamma_2}} - 2\sqrt{\gamma_2}\right)\boldsymbol{z}^{[2]} \\ \dfrac{\gamma_2}{\gamma_1\sqrt{\gamma_1}}V_{21}^{[1]}V_{11}^{[1]}\boldsymbol{z}_1^{[1]} - \dfrac{\sqrt{\gamma_2}}{\gamma_1}V_{21}^{[1]}V_{11}^{[1]}\boldsymbol{z}^{[2]} \end{pmatrix} + \mathrm{O}_p(N^{-\frac{3}{2}}),$$

$$N\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})$$
$$= \begin{pmatrix} I_{p_1} & O_{12} \\ O_{21} & \dfrac{1}{\gamma_1}I_{p2} \end{pmatrix} + \frac{1}{\sqrt{N}} \begin{pmatrix} \sqrt{\gamma_1}V_{11}^{[1]} + \sqrt{\gamma_2}V^{[2]} & \dfrac{1}{\sqrt{\gamma_1}}V_{12}^{[1]} \\ \dfrac{1}{\sqrt{\gamma_1}}V_{21}^{[1]} & \dfrac{1}{\gamma_1\sqrt{\gamma_1}}V_{22}^{[1]} \end{pmatrix}$$
$$+\frac{1}{N}\begin{pmatrix} (\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})(\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})' - 4I_{p1} & \dfrac{\sqrt{\gamma_2}}{\sqrt{\gamma_1}}V^{[2]}V_{12}^{[1]} - \dfrac{\gamma_2}{\gamma_1}V_{11}^{[1]}V_{12}^{[1]} \\ \dfrac{\sqrt{\gamma_2}}{\sqrt{\gamma_1}}V_{21}^{[1]}V^{[2]} - \dfrac{\gamma_2}{\gamma_1}V_{21}^{[1]}V_{11}^{[1]} & \dfrac{p_1\gamma_2 - 2}{\gamma_1^2}I_{p_2} - \dfrac{\gamma_2}{\gamma_1^2}V_{21}^{[1]}V_{12}^{[1]} \end{pmatrix}$$
$$+ \mathrm{O}_p(N^{-\frac{3}{2}}),$$

respectively. Therefore, putting

$$\Delta = \begin{pmatrix} I_{p_1} & O_{p_{12}} \\ O_{p_{21}} & \gamma_1 I_{p_2} \end{pmatrix}, O_{p_{12}} = O_{p_{21}}' : p_1 \times p_2,$$

we have

$$T^2 = (\sqrt{N}\Delta^{\frac{1}{2}}\widehat{\boldsymbol{\mu}})' \left\{ N\Delta^{\frac{1}{2}}\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})\Delta^{\frac{1}{2}'} \right\}^{-1} (\sqrt{N}\Delta^{\frac{1}{2}}\widehat{\boldsymbol{\mu}})$$

and

$$\Delta^{\frac{1}{2}}(\sqrt{N}\widehat{\boldsymbol{\mu}}) = \boldsymbol{u}_0 + \frac{1}{\sqrt{N}}\boldsymbol{u}_1 + \frac{1}{N}\boldsymbol{u}_2 + \mathrm{O}_p(N^{-\frac{3}{2}}),$$
$$\Delta^{\frac{1}{2}}\ N\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})\Delta^{\frac{1}{2}'} = I_p + \frac{1}{\sqrt{N}}M_1 + \frac{1}{N}M_2 + \mathrm{O}_p(N^{-\frac{3}{2}}),$$

where

$$\boldsymbol{u}_0 = \begin{pmatrix} \sqrt{\gamma_1}\boldsymbol{z}_1^{[1]} + \sqrt{\gamma_2}\boldsymbol{z}^{[2]} \\ \boldsymbol{z}_2^{[1]} \end{pmatrix}, \quad \boldsymbol{u}_1 = \begin{pmatrix} \mathbf{0} \\ \sqrt{\gamma_2}V_{21}^{[1]}\boldsymbol{z}^{[2]} - \dfrac{\gamma_2}{\sqrt{\gamma_1}}V_{21}^{[1]}\boldsymbol{z}_1^{[1]} \end{pmatrix},$$

$$\boldsymbol{u}_2 = \begin{pmatrix} \left(\dfrac{1}{\sqrt{\gamma_1}} - 2\sqrt{\gamma_1}\right)\boldsymbol{z}_1^{[1]} + \left(\dfrac{1}{\sqrt{\gamma_2}} - 2\sqrt{\gamma_2}\right)\boldsymbol{z}^{[2]} \\ \dfrac{\gamma_2}{\gamma_1}V_{21}^{[1]}V_{11}^{[1]}\boldsymbol{z}_1^{[1]} - \sqrt{\dfrac{\gamma_2}{\gamma_1}}V_{21}^{[1]}V_{11}^{[1]}\boldsymbol{z}^{[2]} \end{pmatrix},$$

$$M_1 = \begin{pmatrix} \sqrt{\gamma_1}V_{11}^{[1]} + \sqrt{\gamma_2}V^{[2]} & V_{12}^{[1]} \\ V_{21}^{[1]} & \dfrac{1}{\sqrt{\gamma_1}}V_{22}^{[1]} \end{pmatrix}, \quad M_2 = \begin{pmatrix} M_{2,11} & M_{2,12} \\ M_{2,21} & M_{2,22} \end{pmatrix},$$

$$M_{2,11} = (\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})(\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})' - 4I_{p_1},$$

$$M_{2,12} = M_{2,21}' = \sqrt{\gamma_2}V^{[2]}V_{12}^{[1]} - \frac{\gamma_2}{\sqrt{\gamma_1}}V_{11}^{[1]}V_{12}^{[1]},$$

$$M_{2,22} = \frac{\gamma_2 p_1 - 2}{\gamma_1}I_{p_2} - \frac{\gamma_2}{\gamma_1}V_{21}^{[1]}V_{12}^{[1]}.$$

Further, since

$$\left\{\Delta^{\frac{1}{2}}N\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}})\Delta^{\frac{1}{2}'}\right\}^{-1} = I_p - \frac{1}{\sqrt{N}}M_1 + \frac{1}{N}M_3 + \mathrm{O}_p(N^{-\frac{3}{2}}),$$

where

$$M_3 = \begin{pmatrix} M_{3,11} & M_{3,12} \\ M_{3,21} & M_{3,22} \end{pmatrix},$$

$$\begin{aligned} M_{3,11} = {} & \left(\sqrt{\gamma_1}V_{11}^{[1]} + \sqrt{\gamma_2}V^{[2]}\right)^2 + V_{12}^{[1]}V_{21}^{[1]} \\ & - (\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})(\sqrt{\gamma_2}\boldsymbol{z}_1^{[1]} - \sqrt{\gamma_1}\boldsymbol{z}^{[2]})' + 4I_{p_1}, \end{aligned}$$

$$M_{3,12} = M_{3,21}' = \frac{1}{\sqrt{\gamma_1}}V_{11}^{[1]}V_{12}^{[1]} + \frac{1}{\sqrt{\gamma_1}}V_{12}^{[1]}V_{22}^{[1]},$$

$$M_{3,22} = \frac{1}{\gamma_1}\left(V_{21}^{[1]}V_{12}^{[1]} + \left(V_{22}^{[1]}\right)^2\right) - \frac{\gamma_2 p_1 - 2}{\gamma_1}I_{p_2},$$

we can obtain a stochastic expansion of T^2 given by

$$T^2 = Q_0 + \frac{1}{\sqrt{N}}Q_1 + \frac{1}{N}Q_2 + \mathrm{O}_p(N^{-\frac{3}{2}}),$$

where

$$
\begin{aligned}
&\Lambda = (\sqrt{\gamma_1}\boldsymbol{z}_1^{[1]} + \sqrt{\gamma_2}\boldsymbol{z}^{[2]})(\sqrt{\gamma_1}\boldsymbol{z}_1^{[1]} + \sqrt{\gamma_2}\boldsymbol{z}^{[2]})', \\
&Q_0 = \mathrm{tr}\Lambda + \boldsymbol{z}_2^{[1]'}\boldsymbol{z}_2^{[1]}, \\
&Q_1 = -\frac{1}{\sqrt{\gamma_1}}(\boldsymbol{z}_1^{[1]'}V_{12}^{[1]}\boldsymbol{z}_2^{[1]} + \boldsymbol{z}_2^{[1]'}V_{21}^{[1]}\boldsymbol{z}_1^{[1]} + \boldsymbol{z}_2^{[1]'}V_{22}^{[1]}\boldsymbol{z}_2^{[1]}) - \mathrm{tr}(\sqrt{\gamma_1}V_{11}^{[1]} + \sqrt{\gamma_2}V^{[2]})\Lambda, \\
&Q_2 = 2\left(\boldsymbol{z}_1^{[1]'}\boldsymbol{z}_1^{[1]} + \boldsymbol{z}^{[2]'}\boldsymbol{z}^{[2]} + \frac{1}{\sqrt{\gamma_1\gamma_2}}\boldsymbol{z}_1^{[1]'}\boldsymbol{z}^{[2]}\right) - \frac{\gamma_2 p_1 - 2}{\gamma_1}\boldsymbol{z}_2^{[1]'}\boldsymbol{z}_2^{[1]} \\
&\qquad + \mathrm{tr}\left(\sqrt{\gamma_1}V_{11}^{[1]} + \sqrt{\gamma_2}V^{[2]}\right)^2\Lambda \\
&\qquad - \gamma_1\gamma_2 \times \mathrm{tr}\left(\frac{1}{\sqrt{\gamma_1}}\boldsymbol{z}_1^{[1]} - \frac{1}{\sqrt{\gamma_2}}\boldsymbol{z}^{[2]}\right)\left(\frac{1}{\sqrt{\gamma_1}}\boldsymbol{z}_1^{[1]} - \frac{1}{\sqrt{\gamma_2}}\boldsymbol{z}^{[2]}\right)'\Lambda \\
&\qquad + \frac{1}{\gamma_1}\left(\boldsymbol{z}_1^{[1]'}V_{11}^{[1]}V_{12}^{[1]}\boldsymbol{z}_2^{[1]} + \boldsymbol{z}_1^{[1]'}V_{12}^{[1]}V_{21}^{[1]}\boldsymbol{z}_1^{[1]} + \boldsymbol{z}_1^{[1]'}V_{12}^{[1]}V_{22}^{[1]}\boldsymbol{z}_2^{[1]} + \boldsymbol{z}_2^{[1]'}V_{21}^{[1]}V_{11}^{[1]}\boldsymbol{z}_1^{[1]}\right. \\
&\qquad \left. + \boldsymbol{z}_2^{[1]'}V_{22}^{[1]}V_{21}^{[1]}\boldsymbol{z}_1^{[1]} + \boldsymbol{z}_2^{[1]'}V_{21}^{[1]}V_{12}^{[1]}\boldsymbol{z}_2^{[1]} + \boldsymbol{z}_2^{[1]'}(V_{22}^{[1]})^2\boldsymbol{z}_2^{[1]}\right).
\end{aligned}
$$

Similar to the asymptotic expansion of T^2, we have the following result. We can expand T^4 as

$$
T^4 = Q_0^2 + \frac{2}{\sqrt{N}}Q_0Q_1 + \frac{1}{N}\left(2Q_0Q_2 + Q_1^2\right) + \mathrm{O}_p(N^{-\frac{3}{2}}).
$$

By calculating the expectations of T^2 and T^4, we obtain the following theorem.

Theorem 2.1. (Kawasaki and Seo (2016))
Suppose that the data have a two-step monotone missing pattern. If $\gamma_1(= N_1/N)$ is a positive constant as $n_i's$ tend to infinity, then the expectations of T^2 and T^4 can be expanded as

$$
\mathrm{E}[T^2] = p + \frac{1}{N}c_1 + \mathrm{O}(N^{-2}), \quad \mathrm{E}[T^4] = p(p+2) + \frac{1}{N}c_2 + \mathrm{O}(N^{-2}),
$$

where

$$
\begin{aligned}
&c_1 = p_1(p+4) + \frac{p_2(p+3)}{\gamma_1}, \\
&c_2 = 2p_1\{4(p_1+2) + (p_1+1)(p+2) + p_2(p+7)\} + \frac{2p_2}{\gamma_1}\{p_2 + 2 + (p+2)(p+3)\}.
\end{aligned}
$$

Therefore, using Theorem 2.1, we have

$$\mathrm{Var}[T^2] = \mathrm{MSE}[T^2] = 2p + \frac{1}{N}(c_2 - 2pc_1) + \mathrm{O}(N^{-2}).$$

2.4 Transformation statistics

In this section, we propose two transformation statistics by modifying T^2 to follow χ_p^2. The first is the well-known Bartlett correction (Bartlett (1937)). The Bartlett corrected statistic for the T^2 type statistic is given by

$$T_B^2 = \left(1 - \frac{1}{Np}c_1\right) T^2.$$

For the statistic T_B^2, we have

$$\begin{aligned}
\mathrm{E}[T_B^2] &= p + \mathrm{O}(N^{-2}), \\
\mathrm{E}[T_B^4] &= p(p+2) + \frac{1}{N}\{c_2 - 2c_1(p+2)\} + \mathrm{O}(N^{-2}), \\
\mathrm{Var}[T_B^2] &= \mathrm{MSE}[T_B^2] = 2p + \frac{1}{N}\{c_2 - 2c_1(p+2)\} + \mathrm{O}(N^{-2}).
\end{aligned}$$

The second transformation statistic is the modified Bartlett correction by Fujikoshi (2000). The Bartlett correction was originally proposed so that its mean is coincident with that of χ_p^2 up to the order $\mathrm{O}(N^{-1})$ as T_B. Fujikoshi (2000) derived some monotone transformations so that first two moments of the transformed statistics are coincident with those of χ_p^2 up to $\mathrm{O}(N^{-1})$. For details, see Fujikoshi (2000). Note that these transformations can be applied to a wide class of statistics whether their asymptotic expansions are available or not. Applying Fujikoshi's modified Bartlett correction to T^2 leads to

$$T_{MB}^2 = (N\beta_1 + \beta_2) \log\left(1 + \frac{T^2}{N\beta_1}\right),$$

where β_1 and β_2 are given by

$$\beta_1 = \frac{2p(p+2)}{c_2 - 2c_1(p+2)}, \ \beta_2 = \frac{(p+2)\{c_2 - 2c_1(p+4)\}}{2\{c_2 - 2c_1(p+2)\}}.$$

For the statistic T^2_{MB}, we have

$$\mathrm{E}[T^2_{MB}] = p + \mathrm{O}(N^{-2}),$$
$$\mathrm{E}[T^4_{MB}] = p(p+2) + \mathrm{O}(N^{-2}),$$
$$\mathrm{Var}[T^2_{MB}] = \mathrm{MSE}[T^2_{MB}] = 2p + \mathrm{O}(N^{-2}).$$

Note that $\beta_1 > 0$ and $N\beta_1+\beta_2 > 0$ because $c_2-2c_1(p+2) > 0$ and $c_2-2c_1(p+4) > 0$.

2.5 Simulation studies

This section investigates the first and second moments, variance, MSE, upper 100α percentiles, and type I errors of T^2, T^2_B, and T^2_{MB} evaluated via a Monte Carlo simulation. We calculate the asymptotic results of the first moment ($\lambda_1 = p+N^{-1}c_1$), second moment ($\lambda_2 = p(p+2) + N^{-1}c_2$), variance, and MSE ($\lambda_3 = 2p + N^{-1}(c_2 - 2pc_1)$). Computations are carried out for $p = 4, 8, 20$, where the equal missing pattern has $p_1 = p_2$, and the various conditions of p, n_1, n_2, and $\alpha = 0.05, 0.01$. We generate two-step missing data from $N_p(\mathbf{0}, I_p)$ and, as in numerical studies, we carry out 1,000,000 replications. The simulation results are shown in Tables 1-3.

From these numerical studies, it can be observed that our asymptotic results (λ_1, λ_2, and λ_3) are considerably closer to the simulated values of T^2 than any transformation statistic stated in Tables 1 and 2. Meanwhile, the T^2_{MB} values are stable and closer to our asymptotic results than T^2_B when the sample sizes are not large. Table 3 presents the upper percentiles of the T^2 type statistic, T^2_B and T^2_{MB}, and the type I error rate under the simulated T^2 type statistic, T^2_B and T^2_{MB}, when the null hypothesis is rejected using χ^2_p. From Table 3, we note that T^2_{MB} is a very good value in any case, and T^2_B is a good value when n_1 and n_2 are large. It can be seen from Table 3 that the simulated upper percentiles of the T^2 type statistic are closer to the upper percentiles of the χ^2_p distribution as n_1 and n_2 get larger. On the other hand, the simulated upper percentiles of T_B and T_{MB} are much closer to the simulated upper percentiles of the T^2 type statistic than the upper percentiles of χ^2_p

distribution even when the sample size is not large. It can be observed from Table 3 that the null distribution of the transformation statistics is closer to the chi-square distribution.

2.6 Conclusion

In this paper, we considered the problem of testing for the mean vector when the data have a two-step monotone missing pattern. We obtained a stochastic expansion of Hotelling's T^2 type statistic with the use of MLEs developed by Kanda and Fujikoshi (1998), and we showed the asymptotic first two moments. Moreover, we proposed a Bartlett corrected statistic T_B^2 and a modified Bartlett correction statistic T_{MB}^2. We presented the first and second moments, variance, and MSE of T^2, T_B^2, and T_{MB}^2 through the asymptotic results and Monte Carlo simulated results. Additionally, we showed the upper 100α percentiles of T^2, T_B^2, and T_{MB}^2 and the type I errors when the null hypothesis was rejected using χ_p^2 under the simulated T^2, T_B^2, and T_{MB}^2. In conclusion, it may be noted that the asymptotic results of the first and second moments, variance, and MSE are considerably good approximations. Further, it may be also noted that the null distribution of the transformation statistics is closer to the chi-square distribution with p degrees of freedom, even when the sample size is small.

Table 1 First and second moments

p	p_1	p_2	n_1	n_2	$\mathrm{E}[T^2]$	$\mathrm{E}[T_B^2]$	$\mathrm{E}[T_{MB}^2]$	λ_1
4	2	2	40	20	4.58	3.85	3.86	4.63
$\mathrm{E}[\chi_4^2] = 4$			80	40	4.27	3.93	3.94	4.31
			160	80	4.13	3.97	3.97	4.16
			320	160	4.06	3.99	3.99	4.08
			640	320	4.03	3.99	3.99	4.04
			2560	1280	4.01	4.00	4.00	4.01
			20	20	5.25	3.73	3.76	5.16
			40	40	4.56	3.88	3.89	4.56
			80	80	4.24	3.94	3.95	4.28
			160	160	4.12	3.97	3.97	4.14
			240	240	4.08	3.98	3.98	4.09
			320	320	4.06	3.99	3.99	4.07
			20	40	5.11	3.82	3.83	5.01
			40	80	4.46	3.91	3.92	4.49
			80	160	4.21	3.96	3.96	4.24
			160	320	4.10	3.98	3.98	4.12
			320	640	4.05	3.99	3.99	4.06
			1280	2560	4.01	4.00	4.00	4.02
8	4	4	40	20	10.18	7.69	7.74	9.96
$\mathrm{E}[\chi_8^2] = 8$			80	40	8.94	7.87	7.88	8.96
			160	80	8.44	7.94	7.94	8.48
			320	160	8.22	7.97	7.97	8.24
			640	320	8.11	7.99	7.99	8.12
			2560	1280	8.03	8.00	8.00	8.03
			20	20	13.71	7.57	7.70	11.58
			40	40	9.95	7.78	7.81	9.74
			80	80	8.85	7.90	7.90	8.86
			160	160	8.40	7.95	7.95	8.43
			240	240	8.26	7.97	7.97	8.28
			320	320	8.19	7.97	7.98	8.21
			20	40	13.09	7.94	7.94	11.14
			40	80	9.74	7.87	7.88	9.53
			80	160	8.75	7.92	7.93	8.76
			160	320	8.35	7.96	7.96	8.38
			320	640	8.17	7.98	7.98	8.19
			1280	2560	8.04	8.00	8.00	8.05
20	10	10	80	40	26.19	19.72	19.83	24.95
$\mathrm{E}[\chi_{20}^2] = 20$			160	80	22.63	19.85	19.88	22.45
			320	160	21.23	19.93	19.93	21.22
			640	320	20.60	19.97	19.97	20.61
			2560	1280	20.15	19.99	19.99	20.16
			40	40	37.35	20.59	20.70	28.97
			80	80	25.55	19.89	19.95	24.43
			160	160	22.37	19.90	19.92	22.20
			240	240	21.51	19.93	19.94	21.46
			320	320	21.10	19.94	19.95	21.10
			40	80	35.47	21.40	21.24	27.93
			80	160	24.93	20.05	20.07	23.92
			160	320	22.10	19.95	19.96	21.95
			320	640	20.98	19.96	19.96	20.97
			1280	2560	20.23	19.99	19.99	20.24

Table 1 (*Continued*)

p	p_1	p_2	n_1	n_2	$\mathrm{E}[T^4]$	$\mathrm{E}[T_B^4]$	$\mathrm{E}[T_{MB}^4]$	λ_2
4	2	2	40	20	33.05	23.39	22.44	32.58
$\mathrm{E}[\chi_4^4] = 24$			80	40	27.93	23.73	23.26	28.23
			160	80	25.86	23.89	23.66	26.10
			320	160	24.89	23.94	23.82	25.05
			640	320	24.46	23.99	23.93	24.52
			2560	1280	24.13	24.00	23.99	24.13
			20	20	46.43	23.44	21.47	39.58
			40	40	32.02	23.62	22.78	31.59
			80	80	27.49	23.79	23.39	27.75
			160	160	25.64	23.90	23.70	25.86
			240	240	25.06	23.92	23.79	25.24
			320	320	24.79	23.96	23.84	24.93
			20	40	43.65	24.35	22.44	37.55
			40	80	31.03	23.83	23.10	30.62
			80	160	27.08	23.87	23.52	27.27
			160	320	25.44	23.92	23.75	25.63
			320	640	24.70	23.95	23.87	24.81
			1280	2560	24.19	24.00	23.98	24.20
8	4	4	40	20	136.94	78.17	75.15	122.28
$\mathrm{E}[\chi_8^4] = 80$			80	40	102.31	79.15	77.64	100.83
			160	80	90.02	79.58	78.82	90.34
			320	160	84.81	79.84	79.45	85.15
			640	320	82.37	79.94	79.74	82.57
			2560	1280	80.58	79.96	79.91	80.66
			20	20	283.17	86.48	76.14	157.05
			40	40	130.08	79.56	76.71	117.54
			80	80	99.87	79.54	78.22	98.53
			160	160	88.94	79.68	79.04	89.21
			240	240	85.80	79.81	79.38	86.13
			320	320	84.25	79.82	79.50	84.59
			20	40	255.93	94.32	82.12	147.33
			40	80	123.81	80.85	78.28	112.88
			80	160	97.42	79.82	78.71	96.25
			160	320	87.88	79.79	79.25	88.08
			320	640	83.78	79.89	79.62	84.03
			1280	2560	80.94	79.99	79.93	81.00
20	10	10	80	40	776.21	439.81	433.37	666.05
$\mathrm{E}[\chi_{20}^4] = 440$			160	80	569.84	438.53	434.88	552.22
			320	160	498.17	439.10	437.14	495.91
			640	320	467.81	439.69	438.68	467.90
			2560	1280	446.84	439.82	439.56	447.21
			40	40	1694.30	514.92	480.69	849.23
			80	80	736.31	446.23	439.37	642.03
			160	160	555.60	440.03	436.77	540.38
			240	240	511.97	439.74	437.53	506.78
			320	320	491.98	439.48	437.81	490.03
			40	80	1519.68	553.37	509.58	800.56
			80	160	699.11	451.93	445.27	618.20
			160	320	541.81	441.37	438.58	528.59
			320	640	485.89	439.83	438.44	484.17
			1280	2560	450.76	439.90	439.55	451.02

Table 2 Variances and MSEs

p	p_1	p_2	n_1	n_2	Var$[T^2]$	Var$[T_B^2]$	Var$[T_{MB}^2]$	λ_3
4	2	2	40	20	12.07	8.54	7.50	11.50
Var$[\chi_4^2] = 8$			80	40	9.71	8.25	7.76	9.72
			160	80	8.80	8.13	7.89	8.86
			320	160	8.38	8.05	7.94	8.43
			640	320	8.19	8.04	7.98	8.21
			2560	1280	8.05	8.01	8.00	8.05
			20	20	18.90	9.54	7.32	14.32
			40	40	11.57	8.53	7.63	11.08
			80	80	9.51	8.23	7.81	9.52
			160	160	8.69	8.10	7.90	8.75
			240	240	8.45	8.06	7.93	8.50
			320	320	8.33	8.05	7.95	8.38
			20	40	17.53	9.77	7.76	13.44
			40	80	11.10	8.52	7.76	10.66
			80	160	9.31	8.21	7.86	9.31
			160	320	8.61	8.09	7.92	8.65
			320	640	8.29	8.04	7.96	8.33
			1280	2560	8.07	8.01	7.99	8.08
8	4	4	40	20	33.35	19.04	15.23	26.99
Var$[\chi_8^2] = 16$			80	40	22.30	17.25	15.57	21.41
			160	80	18.75	16.57	15.77	18.69
			320	160	17.32	16.30	15.91	17.34
			640	320	16.63	16.14	15.94	16.67
			2560	1280	16.15	16.03	15.98	16.17
			20	20	95.30	29.11	16.88	35.79
			40	40	31.08	19.01	15.69	25.64
			80	80	21.55	17.16	15.73	20.76
			160	160	18.40	16.49	15.81	18.36
			240	240	17.56	16.33	15.89	17.57
			320	320	17.13	16.23	15.90	17.18
			20	40	84.69	31.21	19.06	33.03
			40	80	29.03	18.96	16.16	24.32
			80	160	20.79	17.03	15.86	20.11
			160	320	18.11	16.44	15.88	18.04
			320	640	17.01	16.22	15.95	17.02
			1280	2560	16.24	16.05	15.98	16.25
20	10	10	80	40	90.15	51.08	40.29	68.24
Var$[\chi_{20}^2] = 40$			160	80	57.68	44.39	39.76	54.02
			320	160	47.65	42.00	39.81	46.98
			640	320	43.59	40.97	39.90	43.49
			2560	1280	40.88	40.24	39.97	40.87
			40	40	298.98	90.86	52.00	90.26
			80	80	83.54	50.63	41.35	64.81
			160	160	55.41	43.88	40.00	52.33
			240	240	49.35	42.39	39.90	48.20
			320	320	46.66	41.68	39.85	46.14
			40	80	261.75	95.31	58.45	83.30
			80	160	77.47	50.08	42.41	61.41
			160	320	53.32	43.43	40.28	50.64
			320	640	45.81	41.47	39.99	45.31
			1280	2560	41.33	40.34	39.98	41.32

Table 2 (*Continued*)

p	p_1	p_2	n_1	n_2	MSE[T^2]	MSE[T_B^2]	MSE[T_{MB}^2]	λ_3
4	2	2	40	20	12.40	8.56	7.52	11.50
MSE[χ_4^2] = 8			80	40	9.79	8.26	7.77	9.72
			160	80	8.82	8.13	7.89	8.86
			320	160	8.38	8.05	7.94	8.43
			640	320	8.19	8.04	7.98	8.21
			2560	1280	8.06	8.01	8.00	8.05
			20	20	20.45	9.61	7.38	14.32
			40	40	11.84	8.55	7.64	11.08
			80	80	9.57	8.23	7.81	9.52
			160	160	8.71	8.10	7.90	8.75
			240	240	8.45	8.06	7.93	8.50
			320	320	8.34	8.05	7.95	8.38
			20	40	18.76	9.81	7.79	13.44
			40	80	11.31	8.53	7.77	10.66
			80	160	9.36	8.21	7.86	9.31
			160	320	8.62	8.09	7.92	8.65
			320	640	8.29	8.04	7.96	8.33
			1280	2560	8.07	8.01	7.99	8.08
8	4	4	40	20	38.10	19.14	15.30	26.99
MSE[χ_8^2] = 16			80	40	23.19	17.27	15.58	21.41
			160	80	18.94	16.57	15.77	18.69
			320	160	17.36	16.30	15.91	17.34
			640	320	16.64	16.14	15.94	16.67
			2560	1280	16.15	16.03	15.98	16.17
			20	20	127.87	29.29	16.97	35.79
			40	40	34.88	19.06	15.73	25.64
			80	80	22.27	17.17	15.74	20.76
			160	160	18.56	16.49	15.82	18.36
			240	240	17.63	16.33	15.89	17.57
			320	320	17.17	16.23	15.90	17.18
			20	40	110.55	31.21	19.06	33.03
			40	80	32.05	18.98	16.17	24.32
			80	160	21.36	17.04	15.86	20.11
			160	320	18.23	16.44	15.89	18.04
			320	640	17.04	16.22	15.95	17.02
			1280	2560	16.24	16.05	15.98	16.25
20	10	10	80	40	128.50	51.16	40.32	68.24
MSE[χ_{20}^2] = 40			160	80	64.60	44.41	39.77	54.02
			320	160	49.15	42.00	39.82	46.98
			640	320	43.95	40.97	39.90	43.49
			2560	1280	40.90	40.24	39.97	40.87
			40	40	600.14	91.22	52.49	90.26
			80	80	114.33	50.64	41.35	64.81
			160	160	61.00	43.89	40.00	52.33
			240	240	51.63	42.39	39.90	48.20
			320	320	47.88	41.69	39.86	46.14
			40	80	500.99	97.28	59.99	83.30
			80	160	101.80	50.08	42.42	61.41
			160	320	57.73	43.44	40.28	50.64
			320	640	46.77	41.47	39.99	45.31
			1280	2560	41.39	40.34	39.98	41.32

Table 3 Upper percentiles and type I errors

						$\alpha = 0.05$	
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
4	2	2	40	20	11.29	9.51	9.18
					(0.088)	(0.050)	(0.044)
$\chi^2_{4,0.05} = 9.45$			80	40	10.31	9.51	9.34
					(0.067)	(0.050)	(0.047)
			160	80	9.89	9.50	9.42
					(0.058)	(0.050)	(0.049)
			320	160	9.68	9.49	9.45
					(0.054)	(0.050)	(0.049)
			640	320	9.59	9.49	9.47
					(0.052)	(0.050)	(0.050)
			2560	1280	9.52	9.49	9.49
					(0.051)	(0.050)	(0.050)
			20	20	13.44	9.55	8.98
					(0.130)	(0.051)	(0.040)
			40	40	11.09	9.53	9.24
					(0.084)	(0.051)	(0.045)
			80	80	10.22	9.51	9.37
					(0.065)	(0.050)	(0.048)
			160	160	9.84	9.50	9.43
					(0.057)	(0.050)	(0.049)
			240	240	9.71	9.49	9.44
					(0.055)	(0.050)	(0.049)
			320	320	9.66	9.49	9.46
					(0.054)	(0.050)	(0.049)
			20	40	12.98	9.70	9.19
					(0.121)	(0.054)	(0.044)
			40	80	10.90	9.55	9.31
					(0.080)	(0.051)	(0.046)
			80	160	10.13	9.51	9.39
					(0.063)	(0.050)	(0.048)
			160	320	9.79	9.49	9.43
					(0.056)	(0.050)	(0.049)
			320	640	9.63	9.49	9.46
					(0.053)	(0.050)	(0.049)
			1280	2560	9.53	9.49	9.48
					(0.051)	(0.050)	(0.050)

Table 3 (*Continued*)

					$\alpha = 0.05$		
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
8	4	4	40	20	21.10 (0.154)	15.94 (0.056)	15.07 (0.043)
$\chi^2_{8,0.05} = 15.51$			80	40	17.87 (0.093)	15.71 (0.053)	15.29 (0.046)
			160	80	16.60 (0.069)	15.60 (0.052)	15.40 (0.048)
			320	160	16.04 (0.059)	15.56 (0.051)	15.46 (0.049)
			640	320	15.77 (0.054)	15.54 (0.050)	15.48 (0.050)
			2560	1280	15.57 (0.051)	15.51 (0.050)	15.50 (0.050)
			20	20	31.42 (0.307)	17.36 (0.071)	15.44 (0.049)
			40	40	20.48 (0.142)	16.02 (0.057)	15.25 (0.046)
			80	80	17.61 (0.088)	15.72 (0.053)	15.35 (0.048)
			160	160	16.47 (0.067)	15.59 (0.051)	15.42 (0.048)
			240	240	16.14 (0.061)	15.56 (0.051)	15.45 (0.049)
			320	320	15.97 (0.058)	15.55 (0.051)	15.46 (0.049)
			20	40	29.73 (0.280)	18.05 (0.080)	16.17 (0.060)
			40	80	19.89 (0.131)	16.08 (0.058)	15.42 (0.049)
			80	160	17.34 (0.083)	15.70 (0.053)	15.40 (0.048)
			160	320	16.36 (0.065)	15.59 (0.051)	15.44 (0.049)
			320	640	15.92 (0.057)	15.55 (0.051)	15.47 (0.049)
			1280	2560	15.61 (0.052)	15.52 (0.050)	15.50 (0.050)

Table 3 (*Continued*)

					$\alpha = 0.05$		
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
20	10	10	80	40	43.68	32.88	31.31
					(0.253)	(0.066)	(0.049)
$\chi^2_{20,0.05} = 31.41$			160	80	36.47	31.99	31.26
					(0.125)	(0.057)	(0.049)
			320	160	33.73	31.67	31.31
					(0.082)	(0.053)	(0.049)
			640	320	32.54	31.55	31.37
					(0.065)	(0.052)	(0.050)
			2560	1280	31.69	31.44	31.39
					(0.053)	(0.050)	(0.050)
			40	40	69.51	38.32	33.95
					(0.572)	(0.115)	(0.081)
			80	80	42.35	32.97	31.60
					(0.229)	(0.067)	(0.052)
			160	160	35.91	31.96	31.34
					(0.116)	(0.056)	(0.049)
			240	240	34.26	31.75	31.34
					(0.090)	(0.054)	(0.049)
			320	320	33.48	31.64	31.34
					(0.078)	(0.053)	(0.049)
			40	80	65.52	39.54	35.35
					(0.526)	(0.130)	(0.098)
			80	160	41.08	33.03	31.88
					(0.206)	(0.068)	(0.056)
			160	320	35.37	31.93	31.42
					(0.107)	(0.056)	(0.050)
			320	640	33.24	31.63	31.38
					(0.074)	(0.053)	(0.050)
			1280	2560	31.84	31.46	31.40
					(0.055)	(0.051)	(0.050)

Table 3 (*Continued*)

						$\alpha = 0.01$	
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
4	2	2	40	20	16.52	13.90	12.87
					(0.027)	(0.013)	(0.008)
$\chi^2_{4,0.01} = 13.28$			80	40	14.73	13.57	13.08
					(0.017)	(0.011)	(0.009
			160	80	13.98	13.43	13.19
					(0.013)	(0.011)	(0.010)
			320	160	13.60	13.34	13.22
					(0.011)	(0.010)	(0.010)
			640	320	13.45	13.32	13.26
					(0.011)	(0.010)	(0.010)
			2560	1280	13.32	13.29	13.27
					(0.010)	(0.010)	(0.010)
			20	20	20.88	14.83	12.78
					(0.052)	(0.016)	(0.008)
			40	40	16.15	13.88	12.98
					(0.025)	(0.012)	(0.009)
			80	80	14.56	13.54	13.12
					(0.016)	(0.011)	(0.009)
			160	160	13.87	13.39	13.19
					(0.013)	(0.010)	(0.010)
			240	240	13.67	13.36	13.22
					(0.012)	(0.010)	(0.010)
			320	320	13.57	13.34	13.24
					(0.011)	(0.010)	(0.010)
			20	40	20.07	14.99	13.11
					(0.047)	(0.016)	(0.010)
			40	80	15.82	13.86	13.11
					(0.023)	(0.012)	(0.010)
			80	160	14.40	13.52	13.17
					(0.015)	(0.011)	(0.010)
			160	320	13.81	13.39	13.21
					(0.012)	(0.010)	(0.010)
			320	640	13.54	13.33	13.25
					(0.011)	(0.010)	(0.010)
			1280	2560	13.34	13.29	13.27
					(0.010)	(0.010)	(0.010)

Table 3 (*Continued*)

					$\alpha = 0.01$		
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
8	4	4	40	20	29.09	21.98	19.63
					(0.061)	(0.016)	(0.008)
$\chi^2_{8,0.01} = 20.09$			80	40	23.74	20.88	19.82
					(0.027)	(0.013)	(0.009)
			160	80	21.75	20.45	19.95
					(0.017)	(0.011)	(0.010)
			320	160	20.90	20.28	20.03
					(0.013)	(0.011)	(0.010)
			640	320	20.47	20.17	20.05
					(0.011)	(0.010)	(0.010)
			2560	1280	20.18	20.11	20.08
					(0.010)	(0.010)	(0.010)
			20	20	49.03	27.10	20.83
					(0.177)	(0.031)	(0.012)
			40	40	28.11	21.99	19.95
					(0.054)	(0.017)	(0.010)
			80	80	23.34	20.82	19.93
					(0.025)	(0.013)	(0.009)
			160	160	21.56	20.40	19.98
					(0.016)	(0.011)	(0.010)
			240	240	21.04	20.29	20.01
					(0.014)	(0.011)	(0.010)
			320	320	20.79	20.23	20.02
					(0.013)	(0.011)	(0.010)
			20	40	46.43	28.19	22.24
					(0.156)	(0.035)	(0.018)
			40	80	27.17	21.96	20.26
					(0.048)	(0.017)	(0.011)
			80	160	22.93	20.76	20.02
					(0.023)	(0.012)	(0.010)
			160	320	21.37	20.36	20.01
					(0.015)	(0.011)	(0.010)
			320	640	20.71	20.23	20.05
					(0.012)	(0.010)	(0.010)
			1280	2560	20.23	20.11	20.07
					(0.011)	(0.010)	(0.010)

Table 3 (*Continued*)

					$\alpha = 0.01$		
p	p_1	p_2	n_1	n_2	T^2	T_B^2	T_{MB}^2
20	10	10	80	40	54.49	41.02	37.61
					(0.117)	(0.020)	(0.010)
$\chi^2_{20,0.01} = 37.57$			160	80	44.38	38.93	37.43
					(0.040)	(0.014)	(0.010)
			320	160	40.68	38.19	37.47
					(0.021)	(0.012)	(0.010)
			640	320	39.05	37.86	37.51
					(0.015)	(0.011)	(0.010)
			2560	1280	37.93	37.63	37.54
					(0.011)	(0.010)	(0.010)
			40	40	95.44	52.62	42.05
					(0.406)	(0.055)	(0.025)
			80	80	52.68	41.01	38.08
					(0.101)	(0.020)	(0.011)
			160	160	43.61	38.81	37.54
					(0.036)	(0.014)	(0.010)
			240	240	41.37	38.34	37.52
					(0.024)	(0.012)	(0.010)
			320	320	40.31	38.10	37.50
					(0.020)	(0.012)	(0.010)
			40	80	89.80	54.19	44.35
					(0.359)	(0.063)	(0.034)
			80	160	50.94	40.96	38.52
					(0.086)	(0.020)	(0.013)
			160	320	42.87	38.70	37.67
					(0.032)	(0.013)	(0.010)
			320	640	39.97	38.03	37.54
					(0.018)	(0.011)	(0.010)
			1280	2560	38.13	37.66	37.55
					(0.012)	(0.010)	(0.010)

Chapter 3

Multivariate Behrens-Fisher problem

In this chapter, a new approximate solution is proposed by deriving the asymptotic expansions up to the term of order N^{-2} for the first and second moments of U statistic, where N is total sample size minus 2. We note that the asymptotic expansions up to the term of order N^{-1} for the moments of U are obtained by Yanagihara and Yuan (2005). We obtain the asymptotic expansions of the moments of U, and then the bias correction for the asymptotic expansions using the consistent estimators instead of the unknown parameters are also given since the result of the asymptotic expansions of the moments of U includes some unknown parameters. As a result, two new approximate solutions to the multivariate Behrens-Fisher problem by the approximate moments of U are proposed, moreover, the proofs are presented. We compare five approximate procedures by Monte Carlo simulation and evaluate the advantages of the proposed procedures. In the Appendix, we present certain formulas used to derive the main result.

3.1 Behrens-Fisher problem

Let $\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(i)}, \ldots, \boldsymbol{x}_{n^{(i)}}^{(i)}$ be p-dimensional random vectors from $N_p(\boldsymbol{\mu}^{(i)}, \Sigma^{(i)})$, $i = 1, 2$, $j = 1, 2, \ldots, n^{(i)}$. We consider the following hypothesis test problem:

$$H_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} \ \text{ vs. } \ H_1 : \boldsymbol{\mu}^{(1)} \neq \boldsymbol{\mu}^{(2)}, \tag{3.1}$$

where $\Sigma^{(1)} \neq \Sigma^{(2)}$. A natural statistic for testing (3.1) is

$$T = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})' \left(\frac{S^{(1)}}{n^{(1)}} + \frac{S^{(2)}}{n^{(2)}} \right)^{-1} (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}),$$

where

$$\overline{\boldsymbol{x}}^{(i)} = \frac{1}{n^{(i)}} \sum_{j=1}^{n^{(i)}} \boldsymbol{x}_j^{(i)}, \quad S^{(i)} = \frac{1}{n^{(i)} - 1} \sum_{j=1}^{n^{(i)}} (\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})(\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})'.$$

When $n^{(1)} = n^{(2)}$ and $\Sigma^{(1)} = \Sigma^{(2)}$, the T statistic is reduced to the two sample Hotelling's T^2 statistic. Then, under the null hypothesis in (3.1), $(n-p-1)T/\{p(n-2)\}$ follows the F distribution with p and $n - p - 1$ degrees of freedom, where $n = n^{(1)} + n^{(2)}$.

Mean comparison with unequal variances is intrinsically difficult, and is well known as the Behrens-Fisher problem. Welch (1938) and Scheffé (1943) proposed approximate solutions for the univariate case. One of the earliest methods for solving the multivariate Behrens-Fisher problem was derived by Bennett (1951) based on an extension of Scheffé's (1943) univariate solution. Some approximate solutions were considered by James (1954), Yao (1965), Johansen (1980), Nel et al. (1990), and Kim (1992). Nel et al. (1986) obtained the exact null distribution of T, and Krishnamoorthy and Yu (2004) proposed a modification to the solution. Recently, Krishnamoorthy and Yu (2012) proposed a solution extending the modified Nel and Van der Merwe's test procedure in their earlier study to the case of incomplete data with a monotone pattern. Girón and del Castillo (2010) studied the multivariate Behrens-Fisher distribution, which is defined as the convolution of two independent multivariate Student t distributions. Yanagihara and Yuan (2005) provided three approximate solutions to the multivariate Behrens-Fisher problem that are two F approximations with approximate degrees of freedom and modified Bartlett corrected statistic. However, these solutions are not good approximations when the difference between the covariance matrices is large.

3.2 Approximate degrees of freedom

Assuming the standard regularity condition $n^{(i)}/n = \mathrm{O}(1), i = 1, 2,$ then, as in Yanagihara and Yuan (2005), we can write

$$T = \boldsymbol{z}'W^{-1}\boldsymbol{z} = \frac{\boldsymbol{z}'\boldsymbol{z}}{U}, \tag{3.2}$$

where

$$\boldsymbol{z} = \sqrt{\frac{n^{(1)}n^{(2)}}{n}}\overline{\Sigma}^{-\frac{1}{2}}(\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}), \quad W = \overline{\Sigma}^{-\frac{1}{2}}\left(\frac{n^{(2)}}{n}S^{(1)} + \frac{n^{(1)}}{n}S^{(2)}\right)\overline{\Sigma}^{-\frac{1}{2}},$$

$$\overline{\Sigma} = \frac{n^{(2)}}{n}\Sigma^{(1)} + \frac{n^{(1)}}{n}\Sigma^{(2)}, \quad U = \frac{\boldsymbol{z}'\boldsymbol{z}}{\boldsymbol{z}'W^{-1}\boldsymbol{z}}.$$

By approximating the distribution of U as

$$U \overset{a}{\sim} \frac{\chi^2_\nu}{\phi}, \tag{3.3}$$

we have

$$\frac{\nu}{p\phi}T \overset{a}{\sim} \frac{\chi^2_p/p}{\chi^2_\nu/\nu} \sim F_{p,\nu},$$

where "$\overset{a}{\sim}$" means "approximately following". Note that when $\Sigma^{(1)} = \Sigma^{(2)}$ and $n^{(1)} = n^{(2)}$, U is exactly distributed as χ^2_ν/ϕ, where $\nu = n - p - 1$ and $\phi = n - 2$. In general, the constants ν and ϕ can be given using the following theorems for the first and second moments of U. The proof of the following theorem is provided in Section 3.3.

Theorem 3.1. (Kawasaki and Seo (2015))
Let $U = \boldsymbol{z}'\boldsymbol{z}/\boldsymbol{z}'W^{-1}\boldsymbol{z}$ be defined by (3.2). Then, an asymptotic expansion up to the term of order N^{-2} for $E[U]$ can be expanded as

$$\xi_1 = \mathrm{E}[U] = 1 - \frac{\theta_1}{N} + \frac{1}{N^2}(\theta_2 - \theta_3) + \mathrm{O}(N^{-3}), \tag{3.4}$$

where

$$\begin{aligned}
N =& n-2, \quad \theta_1 = \frac{1}{p(p+2)} \sum_{i=1}^{2} c^{(i)} \left\{ p(a_{(1)}^{(i)})^2 + (p-2)a_{(2)}^{(i)} \right\}, \\
\theta_2 =& \frac{1}{p(p+2)(p+4)} \sum_{i=1}^{2} d^{(i)} \left\{ 4p^2 a_{(3)}^{(i)} + (p-2)(3p+4)a_{(1)}^{(i)}a_{(2)}^{(i)} + p(p+2)(a_{(1)}^{(i)})^3 \right\}, \\
\theta_3 =& \frac{1}{p(p+2)(p+4)(p+6)} \left\{ \sum_{i=1}^{2} (c^{(i)})^2 \left\{ p^2(5p+14)a_{(4)}^{(i)} \right.\right. \\
&+ 4(p+3)(p+2)(p-2)a_{(1)}^{(i)}a_{(3)}^{(i)} + p(p+3)(p-2)(a_{(2)}^{(i)})^2 \\
&\left. +2(p^3+5p^2+7p+6)a_{(2)}^{(i)}(a_{(1)}^{(i)})^2 - p(p+4)(a_{(1)}^{(i)})^4 \right\} \\
&+ 4(p+3)(p+2)(p-2)\psi_1 + 4p(p+2)(p-2)\psi_2 + 4p(p+4)(p+2)\psi_3 \\
&- 2p(p-2)\psi_4 - 2(p+3)(p-2)\psi_5 + 2p(p+4)(p-2)\psi_6 \\
&\left. - 2p(p+4)\psi_7 + 2p(p+4)(3p+2)\psi_8 \right\}
\end{aligned}$$

with $\psi_k, k = 1, 2, \ldots, 8$ given by

$$\begin{aligned}
&\psi_1 = c^{(1)}c^{(2)} \left\{ a_{(1)}^{(1)} b_{(1,2,1)} + a_{(1)}^{(2)} b_{(2,1,1)} \right\}, \quad \psi_2 = c^{(1)}c^{(2)} b_{(2,2,1)}, \\
&\psi_3 = c^{(1)}c^{(2)} a_{(1)}^{(1)} a_{(1)}^{(2)} b_{(1,1,1)}, \quad \psi_4 = c^{(1)}c^{(2)} a_{(2)}^{(1)} a_{(2)}^{(2)}, \\
&\psi_5 = c^{(1)}c^{(2)} \left\{ a_{(2)}^{(1)} (a_{(1)}^{(2)})^2 + (a_{(1)}^{(1)})^2 a_{(2)}^{(2)} \right\}, \quad \psi_6 = c^{(1)}c^{(2)} (b_{(1,1,1)})^2, \\
&\psi_7 = c^{(1)}c^{(2)} (a_{(1)}^{(1)})^2 (a_{(1)}^{(2)})^2, \quad \psi_8 = c^{(1)}c^{(2)} b_{(1,1,2)},
\end{aligned}$$

and

$$\begin{aligned}
&c^{(i)} = \frac{(n-n^{(i)})^2(n-2)}{n^2(n^{(i)}-1)}, \quad d^{(i)} = \frac{(n-n^{(i)})^3(n-2)^2}{n^3(n^{(i)}-1)^2}, \\
&a_{(\ell)}^{(i)} = \mathrm{tr}\left(\Sigma^{(i)}\overline{\Sigma}^{-1}\right)^{\ell}, \ i = 1, 2, \ \ell = 1, 2, 3, 4,
\end{aligned}$$

$$b_{(q,r,s)} = \mathrm{tr}\left\{\left(\Sigma^{(1)}\overline{\Sigma}^{-1}\right)^q\left(\Sigma^{(2)}\overline{\Sigma}^{-1}\right)^r\right\}^s,$$
$$(q,r,s) = (1,1,1),(1,1,2),(1,2,1),(2,1,1),(2,2,1).$$

Similarly, as the result of asymptotic expansion for $\mathrm{E}[U^2]$,we have the following theorem. The proof of the following theorem is provided in Section 3.3.

Theorem 3.2. (Kawasaki and Seo (2015))
Let $U = \boldsymbol{z}'\boldsymbol{z}/\boldsymbol{z}'W^{-1}\boldsymbol{z}$ be defined by (3.2). Then, an asymptotic expansion up to the term of order N^{-2} for $E[U^2]$ can be expanded as

$$\xi_2 = \mathrm{E}[U^2] = 1 - \frac{2}{N}(\theta_1 - \theta_4) + \frac{1}{N^2}(2\theta_5 - \theta_6) + \mathrm{O}(N^{-3}), \tag{3.5}$$

where

$$\theta_4 = \frac{1}{p(p+2)}\sum_{i=1}^{2} c^{(i)}\left\{(a_{(1)}^{(i)})^2 + 2a_{(2)}^{(i)}\right\},$$
$$\theta_5 = \frac{1}{p(p+2)(p+4)}\sum_{i=1}^{2} d^{(i)}\left\{4(p^2-3p+4)a_{(3)}^{(i)} + 3p(p-4)a_{(1)}^{(i)}a_{(2)}^{(i)} + p^2(a_{(1)}^{(i)})^3\right\},$$
$$\begin{aligned}\theta_6 =& \frac{1}{p(p+2)(p+4)(p+6)}\left\{\sum_{i=1}^{2}(c^{(i)})^2\left\{2(p+1)(5p^2-14p+24)a_{(4)}^{(i)}\right.\right.\\
&+ 4(p-4)(2p^2+5p+6)a_{(1)}^{(i)}a_{(3)}^{(i)} + (p-2)(p-4)(2p+3)(a_{(2)}^{(i)})^2\\
&\left.+ 2(p+2)(2p^2-p+12)a_{(2)}^{(i)}(a_{(1)}^{(i)})^2 - 3(p^2+2p-4)(a_{(1)}^{(i)})^4\right\}\\
&+ 4(p-4)(2p^2+5p+6)\psi_1 + 8p(p-2)(p-4)\psi_2 + 8p(p^2+4p+2)\psi_3\\
&- 6(p-2)(p-4)\psi_4 - 6(p-4)(p+2)\psi_5 + 4(p+3)(p-2)(p-4)\psi_6\\
&\left.- 6(p^2+2p-4)\psi_7 + 12(p^3+p^2-2p+8)\psi_8\right\}.\end{aligned}$$

It follows from (3.3) that

$$\mathrm{E}[U] \approx \frac{\nu}{\phi}, \quad \mathrm{E}[U^2] \approx \frac{\nu(\nu+2)}{\phi^2},$$

where "≈" means "approximate equality". Therefore, using the asymptotic expansions of E$[U]$ and E$[U^2]$ in Theorems 3.1 and 3.2, the new approximation to the values of ν and ϕ are given by

$$\nu_S = \frac{2(N^2 - N\theta_1 + \theta_2 - \theta_3)^2}{N^2(N^2 - 2N\theta_1 + 2N\theta_4 + 2\theta_5 - \theta_6) - (N^2 - N\theta_1 + \theta_2 - \theta_3)^2}, \tag{3.6}$$

$$\phi_S = \frac{N^2\nu_S}{N^2 - N\theta_1 + \theta_2 - \theta_3}, \tag{3.7}$$

respectively. If $\theta_2 = \theta_3 = \theta_5 = \theta_6 = 0$, then

$$\nu_S = \frac{(N - \theta_1)^2}{N\theta_4 - \theta_1^2/2}(= \nu_F), \quad \phi_S = \frac{\nu_S N}{N - \theta_1}(= \phi_F), \tag{3.8}$$

and these values are the same as the results of Yanagihara and Yuan (2005). In addition, they slightly adjust the coefficient ν_F to

$$\nu_Y = \frac{(N - \theta_1)^2}{N\theta_4 - \theta_1}, \tag{3.9}$$

which will be used to obtain the F approximation.

In practical use, we must estimate $a_{(\ell)}^{(i)}$ and $b_{(q,r,s)}$ since $\Sigma^{(i)}$ and $\overline{\Sigma}$ are unknown. In this paper, using the consistent estimators of them, we obtain the following results. The proof of the following theorem is provided in Section 3.4.

Theorem 3.3. (Kawasaki and Seo (2015))
Let the estimators for (3.4) and (3.5) be

$$\widehat{\xi}_1 = \widehat{\mathrm{E}}[U] = 1 - \frac{\widehat{\theta}_1}{N} + \frac{1}{N^2}(\widehat{\theta}_2 - \widehat{\theta}_3) + \mathrm{O}_p(N^{-3}),$$
$$\widehat{\xi}_2 = \widehat{\mathrm{E}}[U^2] = 1 - \frac{2}{N}(\widehat{\theta}_1 - \widehat{\theta}_4) + \frac{1}{N^2}(2\widehat{\theta}_5 - \widehat{\theta}_6) + \mathrm{O}_p(N^{-3}),$$

respectively. Then

$$\mathrm{E}[\widehat{\xi}_1] = 1 - \frac{\theta_1}{N} + \frac{1}{N^2}(\theta_2 - \theta_3 - \theta_1^*) + \mathrm{o}(N^{-2}),$$
$$\mathrm{E}[\widehat{\xi}_2] = 1 - \frac{2}{N}(\theta_1 - \theta_4) + \frac{1}{N^2}(2\theta_5 - \theta_6 - 2\theta_1^* + 2\theta_4^*) + \mathrm{o}(N^{-2}),$$

where

$$\theta_1^* = \eta_1 - \eta_2 + \eta_3,\ \theta_4^* = \eta_4 - \eta_5 + \eta_6$$

with

$$\eta_1 = \frac{1}{p(p+2)} \sum_{i=1}^{2} \frac{c^{(i)}}{(\rho^{(i)})^2} \left\{ (p-2)(a_{(1)}^{(i)})^2 + (3p-2)a_{(2)}^{(i)} \right\},$$

$$\eta_2 = \frac{1}{p(p+2)} \sum_{i=1}^{2} d^{(i)} \left\{ (7p-6)a_{(3)}^{(i)} + (5p-6)a_{(1)}^{(i)}a_{(2)}^{(i)} + 2p(a_{(1)}^{(i)})^3 \right\},$$

$$\eta_3 = \frac{1}{p(p+2)} \times \left[\sum_{i=1}^{2} (c^{(i)})^2 \left\{ 4(p-1)a_{(4)}^{(i)} + (3p-2)a_{(1)}^{(i)}a_{(3)}^{(i)} + (p-2)(a_{(2)}^{(i)})^2 + 2p(a_{(1)}^{(i)})^2 a_{(2)}^{(i)} \right\} + (3p-2)\psi_1 + 2(p-2)\psi_2 + 4p\psi_3 + 2(p-2)\psi_6 + 2(3p-2)\psi_8 \right],$$

$$\eta_4 = \frac{1}{p(p+2)} \sum_{i=1}^{2} \frac{c^{(i)}}{(\rho^{(i)})^2} \left\{ 2(a_{(1)}^{(i)})^2 + 4a_{(2)}^{(i)} \right\},$$

$$\eta_5 = \frac{1}{p(p+2)} \sum_{i=1}^{2} d^{(i)} \left\{ 10a_{(3)}^{(i)} + 8a_{(1)}^{(i)}a_{(2)}^{(i)} + 2(a_{(1)}^{(i)})^3 \right\},$$

$$\eta_6 = \frac{1}{p(p+2)} \left[\sum_{i=1}^{2} (c^{(i)})^2 \left\{ 6a_{(4)}^{(i)} + 4a_{(1)}^{(i)}a_{(3)}^{(i)} + 2(a_{(2)}^{(i)})^2 + 2(a_{(1)}^{(i)})^2 a_{(2)}^{(i)} \right\} + 4(\psi_1 + \psi_2 + \psi_3 + \psi_6 + 2\psi_8) \right].$$

As with derivation of (3.6) and (3.7), using the asymptotic expansions of $\mathrm{E}[\widehat{\xi}_1]$ and $\mathrm{E}[\widehat{\xi}_2]$ in Theorems 3.3, we have the following ν and ϕ. Consequently

$$\nu_{BC} = \frac{2\ \vartheta^2}{N^2\left(N^2 - 2N\theta_1 + 2N\theta_4 + 2\theta_5 - \theta_6 - 2\theta_1^* + 2\theta_4^*\right) - \vartheta^2}, \tag{3.10}$$

$$\phi_{BC} = \frac{N^2 \nu_{BC}}{\vartheta}, \tag{3.11}$$

where $\vartheta = N^2 - N\theta_1 + \theta_2 - \theta_3 - \theta_1^*$.

Then we can propose procedures as follows.

(I) Second Order Procedure (S-Procedure)

$$T_S = \frac{\widehat{\nu}_S}{p\widehat{\phi}_S} T \overset{a}{\sim} F_{p,\widehat{\nu}_S},$$

where $\widehat{\nu}_S$ and $\widehat{\phi}_S$ are the consistent estimators of (3.6) and (3.7), respectively.

(II) Bias Correction Procedure (BC-Procedure)

$$T_{BC} = \frac{\widehat{\nu}_{BC}}{p\widehat{\phi}_{BC}} T \overset{a}{\sim} F_{p,\widehat{\nu}_{BC}},$$

where $\widehat{\nu}_{BC}$ and $\widehat{\phi}_{BC}$ are the consistent estimators of (3.10) and (3.11), respectively.

3.3 Proofs of Theorem 3.1 and 3.2

Let

$$\rho^{(i)} = \sqrt{\frac{n^{(i)}-1}{n-2}}, \quad \Omega^{(i)} = \sqrt{\frac{n-n^{(i)}}{n}} \overline{\Sigma}^{-\frac{1}{2}} \left(\Sigma^{(i)}\right)^{\frac{1}{2}},$$

and

$$V^{(i)} = \sqrt{n^{(i)}-1} \left\{ \left(\Sigma^{(i)}\right)^{-\frac{1}{2}} S^{(i)} \left(\Sigma^{(i)}\right)^{-\frac{1}{2}} - I_p \right\}, \ i = 1, 2.$$

Then, W^{-1} can be expanded as

$$W^{-1} = I_p - \frac{1}{\sqrt{N}}\overline{V} + \frac{1}{N}\overline{V}^2 - \frac{1}{N\sqrt{N}}\overline{V}^3 + \frac{1}{N^2}\overline{V}^4 + \mathrm{O}_p(N^{-5/2}), \tag{3.12}$$

where

$$\overline{V} = \sum_{i=1}^{2} \rho_i^{-1} \Omega_i V_i \Omega_i'.$$

Note that $U = \boldsymbol{z}'\boldsymbol{z}/\boldsymbol{z}'W^{-1}\boldsymbol{z}$. It follows from (3.12) that we can expand U as

$$\begin{aligned} U =& 1 + \frac{1}{\sqrt{N}}\Theta_1 + \frac{1}{N}(\Theta_1^2 - \Theta_2) + \frac{1}{N\sqrt{N}}(\Theta_3 - 2\Theta_1\Theta_2 + \Theta_1^3) \\ &+ \frac{1}{N^2}(\Theta_1^4 - \Theta_4 + 2\Theta_1\Theta_3 + \Theta_2^2 - 3\Theta_1^2\Theta_2) + \mathrm{O}_p(N^{-5/2}), \end{aligned}$$

where $\Theta_i = \boldsymbol{z}'\overline{V}^i\boldsymbol{z}/\boldsymbol{z}'\boldsymbol{z}, i = 1,2,3,4$. Note that $\overline{V}$ and $\boldsymbol{z}$ are independent, and so are $\boldsymbol{z}'\overline{V}\boldsymbol{z}/\boldsymbol{z}'\boldsymbol{z}$ and $\boldsymbol{z}'\boldsymbol{z}$, as well as $\boldsymbol{z}'\overline{V}^2\boldsymbol{z}/\boldsymbol{z}'\boldsymbol{z}$ and $\boldsymbol{z}'\boldsymbol{z}$ (see Fang et al., 1990, p.30). In the same way as in Yanagihara and Yuan (2005), the following results can be obtained after a good deal of calculation:

(i) $\mathrm{E}[\Theta_1] = 0$,

(ii)
$$\mathrm{E}[\Theta_2] = \frac{1}{p}\sum_{i=1}^{2} c^{(i)}\left\{(a_{(1)}^{(i)})^2 + a_{(2)}^{(i)}\right\},$$
$$\mathrm{E}[\Theta_1^2] = \frac{2}{p(p+2)}\sum_{i=1}^{2} c^{(i)}\left\{(a_{(1)}^{(i)})^2 + 2a_{(2)}^{(i)}\right\},$$

(iii)
$$\sqrt{N}\mathrm{E}[\Theta_3] = \frac{1}{p}\sum_{i=1}^{2} d^{(i)}\left\{4a_{(3)}^{(i)} + 3a_{(1)}^{(i)}a_{(2)}^{(i)} + (a_{(1)}^{(i)})^3\right\},$$
$$\sqrt{N}\mathrm{E}[\Theta_1\Theta_2] = \frac{2}{p(p+2)}\sum_{i=1}^{2} d^{(i)}\left\{6a_{(3)}^{(i)} + 5a_{(1)}^{(i)}a_{(2)}^{(i)} + (a_{(1)}^{(i)})^3\right\},$$
$$\sqrt{N}\mathrm{E}[\Theta_1^3] = \frac{8}{p(p+2)(p+4)}\sum_{i=1}^{2} d^{(i)}\left\{8a_{(3)}^{(i)} + 6a_{(1)}^{(i)}a_{(2)}^{(i)} + (a_{(1)}^{(i)})^3\right\},$$

(iv)
$$\mathrm{E}[\Theta_4] = \frac{1}{p}\left\{\sum_{i=1}^{2}(c^{(i)})^2\left\{5a_{(4)}^{(i)} + 4a_{(1)}^{(i)}a_{(3)}^{(i)} + (a_{(2)}^{(i)})^2 + 2a_{(2)}^{(i)}(a_{(1)}^{(i)})^2\right\}\right.$$
$$\left. +2(2\psi_1 + 2\psi_2 + 2\psi_3 + \psi_6 + 3\psi_8)\right\} + \mathrm{O}(N^{-1}),$$
$$\mathrm{E}[\Theta_1^4] = \frac{12}{p(p+2)(p+4)(p+6)}$$
$$\times\left\{\sum_{i=1}^{2}(c^{(i)})^2\left\{48a_{(4)}^{(i)} + 32a_{(1)}^{(i)}a_{(3)}^{(i)} + 12(a_{(2)}^{(i)})^2 + 12a_{(2)}^{(i)}(a_{(1)}^{(i)})^2 + (a_{(1)}^{(i)})^4\right\}\right.$$
$$\left. +2(16\psi_1 + 32\psi_2 + 8\psi_3 + 4\psi_4 + 2\psi_5 + 8\psi_6 + \psi_7 + 16\psi_8)\right\} + \mathrm{O}(N^{-1}),$$
$$\mathrm{E}[\Theta_1\Theta_3] = \frac{2}{p(p+2)}\left\{\sum_{i=1}^{2}(c^{(i)})^2\left\{8a_{(4)}^{(i)} + 7a_{(1)}^{(i)}a_{(3)}^{(i)} + (a_{(2)}^{(i)})^2 + 2a_{(2)}^{(i)}(a_{(1)}^{(i)})^2\right\}\right.$$
$$\left. +7\psi_1 + 10\psi_2 + 4\psi_3 + 2\psi_6 + 6\psi_8\right\} + \mathrm{O}(N^{-1}),$$

$$\mathrm{E}[\Theta_2^2] = \frac{1}{p(p+2)}$$
$$\times\left\{ \sum_{i=1}^{2}(c^{(i)})^2 \left\{14a_{(4)}^{(i)} + 8a_{(1)}^{(i)}a_{(3)}^{(i)} + 7(a_{(2)}^{(i)})^2 + (a_{(1)}^{(i)})^4 + 6a_{(2)}^{(i)}(a_{(1)}^{(i)})^2\right\}\right.$$
$$\left. +2(4\psi_1 + 4\psi_2 + 4\psi_3 + \psi_4 + \psi_5 + 6\psi_6 + \psi_7 + 10\psi_8)\right\} + \mathrm{O}(N^{-1}),$$
$$\mathrm{E}[\Theta_1^2\Theta_2] = \frac{2}{p(p+2)(p+4)}$$
$$\times\left\{ \sum_{i=1}^{2}(c^{(i)})^2 \left\{40a_{(4)}^{(i)} + 28a_{(1)}^{(i)}a_{(3)}^{(i)} + 10(a_{(2)}^{(i)})^2 + 11a_{(2)}^{(i)}(a_{(1)}^{(i)})^2 + (a_{(1)}^{(i)})^4\right\}\right.$$
$$\left. +28\psi_1 + 48\psi_2 + 16\psi_3 + 4\psi_4 + 3\psi_5 + 16\psi_6 + 2\psi_7 + 32\psi_8\right\} + \mathrm{O}(N^{-1}).$$

Using the above results, we can show (3.4). This completes the proof of Theorem 3.1.

In the same as Theorem 3.1, we can expand U^2 as

$$U^2 = 1 + \frac{2}{\sqrt{N}}\Theta_1 + \frac{1}{N}(3\Theta_1^2 - 2\Theta_2) + \frac{2}{N\sqrt{N}}(\Theta_3 - 3\Theta_1\Theta_2 + 2\Theta_1^3)$$
$$+ \frac{1}{N^2}(5\Theta_1^4 - 2\Theta_4 + 6\Theta_1\Theta_3 + 3\Theta_2^2 - 12\Theta_1^2\Theta_2) + \mathrm{O}_p(N^{-5/2}).$$

By calculating the expectations of the above results, we can show (3.5). This completes the proof of Theorem 3.2.

3.4 Proof of Theorem 3.3

Let

$$S^{(i)} = \Sigma^{(i)} + \frac{1}{\sqrt{n^{(i)}-1}}\left(\Sigma^{(i)}\right)^{\frac{1}{2}} V^{(i)} \left(\Sigma^{(i)}\right)^{\frac{1}{2}},$$
$$\overline{S} = \frac{n^{(2)}}{n}S^{(1)} + \frac{n^{(1)}}{n}S^{(2)},$$

where $S^{(i)}$ $(i = 1, 2)$ is consistent estimator. Then $\overline{S}^{-1}$ can be written by

$$\overline{S}^{-1} = \overline{\Sigma}^{-\frac{1}{2}}\left(I_p + \frac{n^{(2)}}{n\sqrt{n^{(1)}-1}}\overline{\Sigma}^{-\frac{1}{2}}\left(\Sigma^{(1)}\right)^{-\frac{1}{2}} V^{(1)} \left(\Sigma^{(1)}\right)^{-\frac{1}{2}}\overline{\Sigma}^{-\frac{1}{2}}\right.$$
$$\left. + \frac{n^{(1)}}{n\sqrt{n^{(2)}-1}}\overline{\Sigma}^{-\frac{1}{2}}\left(\Sigma^{(1)}\right)^{-\frac{1}{2}} V^{(2)} \left(\Sigma^{(2)}\right)^{-\frac{1}{2}}\overline{\Sigma}^{-\frac{1}{2}}\right)^{-1}\overline{\Sigma}^{-\frac{1}{2}}.$$

Moreover above $S^{(i)}$ and $\overline{S}$ are substituted into $\widehat{\theta}_1$ and $\widehat{\theta}_4$, and calculated the expected value of them as

$$
\begin{aligned}
\mathrm{E}\left[\widehat{\theta}_1\right] &= \theta_1 + \frac{1}{p(p+2)}\left[\sum_{i=1}^{2} c^{(i)}\left\{\frac{1}{n^{(i)}-1}\left\{(p-2)(a_{(1)}^{(i)})^2 + (3p-2)a_{(2)}^{(i)}\right\}\right.\right. \\
&\quad - \frac{n-n^{(i)}}{n(n^{(i)}-1)}\left\{(7p-6)a_{(3)}^{(i)} + (5p-6)a_{(1)}^{(i)}a_{(2)}^{(i)} + 2p(a_{(1)}^{(i)})^3\right\} \\
&\quad + \frac{(n-n^{(i)})^2}{n^2(n^{(i)}-1)} \\
&\quad \left.\times\left\{4(p-1)a_{(4)}^{(i)} + (3p-2)a_{(1)}^{(i)}a_{(3)}^{(i)} + (p-2)(a_{(2)}^{(i)})^2 + 2p(a_{(1)}^{(i)})^2a_{(2)}^{(i)}\right\}\right\} \\
&\quad \left. + \frac{1}{N}\left\{(3p-2)\psi_1 + 2(p-2)\psi_2 + 4p\psi_3 + 2(p-2)\psi_6 + 2(3p-2)\psi_8\right\}\right] \\
&\quad + \mathrm{o}(N^{-1}) \\
&= \theta_1 + \frac{1}{N}(\eta_1 - \eta_2 + \eta_3) + \mathrm{o}(N^{-1}) \\
&= \theta_1 + \frac{1}{N}\theta_1^* + \mathrm{o}(N^{-1}), \\
\mathrm{E}\left[\widehat{\theta}_4\right] &= \theta_4 + \frac{1}{p(p+2)}\left[\sum_{i=1}^{2} c^{(i)}\left\{\frac{1}{n^{(i)}-1}\left\{2(a_{(1)}^{(i)})^2 + 4a_{(2)}^{(i)}\right\}\right.\right. \\
&\quad - \frac{n-n^{(i)}}{n(n^{(i)}-1)}\left\{10a_{(3)}^{(i)} + 8a_{(1)}^{(i)}a_{(2)}^{(i)} + 2(a_{(1)}^{(i)})^3\right\} \\
&\quad \left. + \frac{(n-n^{(i)})^2}{n^2(n^{(i)}-1)}\left\{6a_{(4)}^{(i)} + 4a_{(1)}^{(i)}a_{(3)}^{(i)} + 2(a_{(2)}^{(i)})^2 + 2(a_{(1)}^{(i)})^2a_{(2)}^{(i)}\right\}\right\} \\
&\quad \left. + \frac{4}{N}\left\{\psi_1 + \psi_2 + \psi_3 + \psi_6 + 2\psi_8\right\}\right] + \mathrm{o}(N^{-1}) \\
&= \theta_4 + \frac{1}{N}(\eta_4 - \eta_5 + \eta_6) + \mathrm{o}(N^{-1}) \\
&= \theta_4 + \frac{1}{N}\theta_4^* + \mathrm{o}(N^{-1}).
\end{aligned}
$$

Using the above results, we can calculate as

$$
\begin{aligned}
\mathrm{E}[\widehat{\xi}_1] &= 1 - \frac{1}{N}\mathrm{E}[\widehat{\theta}_1] + \frac{1}{N^2}(\mathrm{E}[\widehat{\theta}_2] - \mathrm{E}[\widehat{\theta}_3]) + \mathrm{O}(N^{-3}) \\
&= 1 - \frac{1}{N}\left(\theta_1 + \frac{1}{N}\theta_1^*\right) + \frac{1}{N^2}(\theta_2 - \theta_3) + \mathrm{o}(N^{-2}) \\
&= 1 - \frac{1}{N}\theta_1 + \frac{1}{N^2}(\theta_2 - \theta_3 - \theta_1^*) + \mathrm{o}(N^{-2}), \\
\mathrm{E}[\widehat{\xi}_2] &= 1 - \frac{2}{N}(\mathrm{E}[\widehat{\theta}_1] - \mathrm{E}[\widehat{\theta}_4]) + \frac{1}{N^2}(2\mathrm{E}[\widehat{\theta}_5] - \mathrm{E}[\widehat{\theta}_6]) + \mathrm{O}(N^{-3}) \\
&= 1 - \frac{2}{N}\left(\theta_1 + \frac{1}{N}\theta_1^* - \theta_4 - \frac{1}{N}\theta_4^*\right) + \frac{1}{N^2}(2\theta_5 - \theta_6) + \mathrm{o}(N^{-2}) \\
&= 1 - \frac{2}{N}(\theta_1 - \theta_4) + \frac{1}{N^2}(2\theta_5 - \theta_6 - 2\theta_1^* + 2\theta_4^*) + \mathrm{o}(N^{-2}).
\end{aligned}
$$

This completes the proof of Theorem 3.3.

3.5 Simulation studies

In this section, we perform a Monte Carlo simulation in order to investigate the accuracy of our procedures (I) and (II), and to compare it with the following three procedures:

(III) First Order Procedure (F-Procedure)

$$
T_F = \frac{\widehat{\nu}_F}{p\widehat{\phi}_F} T \overset{a}{\sim} F_{p,\widehat{\nu}_F},
$$

where $\widehat{\nu}_F$ and $\widehat{\phi}_F$ are consistent estimators of (3.8), respectively.

(IV) Yanagihara and Yuan's Procedure (Y-Procedure)

$$
T_Y = \frac{\widehat{\nu}_Y}{p\widehat{\phi}_Y} T \overset{a}{\sim} F_{p,\widehat{\nu}_Y},
$$

where

$$
\widehat{\phi}_Y = \frac{\widehat{\nu}_Y N}{N - \widehat{\theta}_1}
$$

and $\widehat{\nu}_Y$ is a consistent estimator of (3.9). See Yanagihara and Yuan (2005).

(V) Modified Bartlett Procedure (MB-Procedure)

$$T_{MB} = (N\widehat{\beta}_3 + \widehat{\beta}_4)\log\left(1 + \frac{T}{N\widehat{\beta}_3}\right) \overset{a}{\sim} \chi_p^2,$$

where

$$\widehat{\beta}_3 = \frac{2}{\widehat{\kappa}_2 - 2\widehat{\kappa}_1}, \quad \widehat{\beta}_4 = \frac{(p+2)\widehat{\kappa}_2 - 2(p+4)\widehat{\kappa}_1}{2(\widehat{\kappa}_2 - 2\widehat{\kappa}_1)},$$

$$\widehat{\kappa}_1 = \frac{1}{p}\sum_{i=1}^{2} c^{(i)}\left\{(\widehat{a}_{(1)}^{(i)})^2 + \widehat{a}_{(2)}^{(i)}\right\},$$

$$\widehat{\kappa}_2 = \frac{1}{p(p+2)}\sum_{i=1}^{2} c_i\left\{2(p+3)(\widehat{a}_i^{(1)})^2 + 2(p+4)\widehat{a}_i^{(2)}\right\}.$$

See Yanagihara and Yuan (2005).

For each of parameters, the simulation was carried out for 1,000,000 trials based on normal random vectors. Without loss of generality, we can assume that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \mathbf{0}$.

We compare the following type I errors for five procedures:

(I) $\alpha_1 = \mathrm{P}(T_S > F_{\alpha;p,\widehat{\nu}_S})$, (II) $\alpha_2 = \mathrm{P}(T_{BC} > F_{\alpha;p,\widehat{\nu}_{BC}})$,

(III) $\alpha_3 = \mathrm{P}(T_F > F_{\alpha;p,\widehat{\nu}_F})$, (IV) $\alpha_4 = \mathrm{P}(T_Y > F_{\alpha;p,\widehat{\nu}_Y})$,

(V) $\alpha_5 = \mathrm{P}(T_{MB} > \chi^2_{\alpha;p})$,

where $F_{\alpha;f,g}$ is the upper 100α percentile of the F distribution with f and g degrees of freedom and $\chi^2_{\alpha;p}$ is the upper 100α percentile of the chi-square distribution with p degrees of freedom. We choose $\alpha = 0.05,\ 0.01$, $p = 4, 8$, and the sample sizes $(n^{(1)}, n^{(2)}) =$ (10, 10), (10, 20), (20, 10), (20, 20) for (I)$\sim$(V). We note that the second degree of freedom of F distribution for the test statistic (I)$\sim$(IV) changes with each simulation of 1,000,000 trials.

Table 1 presents the empirical sizes $\widehat{\alpha}_j, j = 1, 2, 3, 4, 5$ in the case of $\Sigma^{(1)} = \mathrm{diag}(\epsilon, \epsilon^2, \cdots, \epsilon^p)$ and $\Sigma^{(2)} = I$, where $\epsilon =$ 1, 5, 10, 20. We note that $|\Sigma^{(1)}| \geq 1$ and the difference between $\Sigma^{(1)}$ and $\Sigma^{(2)}$ is large when ϵ is large. Tables 2 and 3 present the empirical sizes $\widehat{\alpha}_j, j = 1, 2, 3, 4, 5$ in the case of $\Sigma^{(1)} = \sigma^2 I$ and $\Sigma^{(2)} = I$, where $\sigma^2 = 0.1(0.2)0.7$ in Table 2, and $\sigma^2 = 2$, 5, 10, 20, 30 in Table 3. We note that the empirical sizes for the case of $|\Sigma^{(1)}| \leq 1$ and $|\Sigma^{(1)}| > 1$ are given in Tables 2 and 3, respectively. The last row of each of these Tables indicates the average absolute discrepancy (AAD). In this context, see Yanagihara and Yuan (2005).

From Table 1, it is seen that the proposed approximations $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ are very good for cases when ϵ is large. In contrast, it seems that other $\widehat{\alpha}_j$ are farther from α as ϵ becomes large. It may also be noted that $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ are stable and good approximation to α when n_1 and n_2 are large. From Table 2, we can see that $\widehat{\alpha}_4$'s AAD and $\widehat{\alpha}_5$'s AAD are lower than the others, and their approximations are good for the case of $p = 4$, and that $\widehat{\alpha}_1$ are a good approximation for the case of $p = 8$. On the other hand, it seems from Table 2 that the empirical sizes are almost unchanged except for $\widehat{\alpha}_3$. It can be seen from Table 3 that $\widehat{\alpha}_4$ are closer to α when $p = 4$. In addition, it is seen from Table 3 that $\widehat{\alpha}_1$ and $\widehat{\alpha}_4$ are good approximations when $p = 8$.

Table 1 Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_5$) when $p = 4, 8$, and $\epsilon = 1,\ 5{,}10{,}20$

			$\alpha = 0.05$				
p	$(n^{(1)}, n^{(2)})$	ϵ	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	1	0.047	0.043	0.048	0.044	0.046
		5	0.057	0.061	0.067	0.045	0.057
		10	0.055	0.059	0.069	0.042	0.057
		20	0.054	0.056	0.070	0.039	0.057
	(10,20)	1	0.053	0.056	0.053	0.049	0.051
		5	0.057	0.062	0.070	0.039	0.059
		10	0.054	0.058	0.070	0.036	0.057
		20	0.053	0.054	0.070	0.033	0.056
	(20,10)	1	0.053	0.055	0.053	0.049	0.051
		5	0.051	0.050	0.052	0.049	0.051
		10	0.051	0.051	0.053	0.049	0.051
		20	0.050	0.051	0.053	0.049	0.051
	(20,20)	1	0.049	0.048	0.049	0.048	0.049
		5	0.051	0.052	0.053	0.049	0.052
		10	0.051	0.051	0.054	0.048	0.052
		20	0.050	0.051	0.053	0.047	0.051
	AAD		0.289	0.473	0.896	0.531	0.369
8	(10,10)	1	0.046	0.044	0.000	0.035	0.048
		5	0.077	0.061	0.000	0.030	0.149
		10	0.073	0.056	0.000	0.022	0.160
		20	0.069	0.052	0.000	0.017	0.167
	(10,20)	1	0.060	0.061	0.028	0.048	0.060
		5	0.079	0.061	0.000	0.017	0.162
		10	0.073	0.055	0.000	0.012	0.169
		20	0.068	0.051	0.000	0.010	0.172
	(20,10)	1	0.059	0.061	0.028	0.048	0.059
		5	0.053	0.053	0.000	0.046	0.058
		10	0.052	0.053	0.000	0.044	0.059
		20	0.052	0.052	0.000	0.043	0.059
	(20,20)	1	0.050	0.048	0.068	0.047	0.049
		5	0.053	0.054	0.000	0.041	0.059
		10	0.052	0.052	0.000	0.038	0.059
		20	0.051	0.051	0.000	0.037	0.058
	AAD		1.098	0.514	4.447	1.651	4.697

Note : $\text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/16,\ j = 1, 2, 3, 4, 5$

Table 1 *(Continued)*

			$\alpha = 0.01$				
p	$(n^{(1)}, n^{(2)})$	ϵ	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	1	0.009	0.008	0.010	0.008	0.009
		5	0.013	0.014	0.017	0.008	0.013
		10	0.012	0.013	0.018	0.007	0.013
		20	0.012	0.012	0.019	0.007	0.013
	(10,20)	1	0.011	0.012	0.011	0.009	0.010
		5	0.013	0.014	0.019	0.007	0.013
		10	0.012	0.013	0.019	0.006	0.013
		20	0.011	0.012	0.019	0.005	0.013
	(20,10)	1	0.011	0.011	0.011	0.009	0.010
		5	0.010	0.010	0.011	0.010	0.010
		10	0.010	0.010	0.011	0.010	0.011
		20	0.010	0.010	0.011	0.009	0.010
	(20,20)	1	0.010	0.009	0.010	0.009	0.009
		5	0.010	0.011	0.011	0.009	0.010
		10	0.010	0.011	0.012	0.009	0.011
		20	0.010	0.010	0.012	0.009	0.010
	AAD		0.108	0.174	0.384	0.181	0.139
8	(10,10)	1	0.009	0.008	0.000	0.005	0.009
		5	0.018	0.012	0.000	0.004	0.052
		10	0.018	0.012	0.000	0.003	0.061
		20	0.018	0.011	0.000	0.002	0.066
	(10,20)	1	0.012	0.012	0.013	0.008	0.012
		5	0.020	0.012	0.000	0.002	0.061
		10	0.020	0.012	0.000	0.001	0.067
		20	0.019	0.012	0.000	0.001	0.071
	(20,10)	1	0.012	0.012	0.013	0.008	0.012
		5	0.011	0.011	0.000	0.009	0.013
		10	0.011	0.011	0.000	0.008	0.013
		20	0.011	0.010	0.000	0.007	0.013
	(20,20)	1	0.010	0.010	0.019	0.009	0.010
		5	0.011	0.011	0.000	0.007	0.013
		10	0.011	0.011	0.000	0.006	0.013
		20	0.010	0.010	0.000	0.006	0.013
	AAD		0.384	0.135	0.908	0.460	2.122

Note : $\text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/16, \; j = 1, 2, 3, 4, 5$

Table 2 Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_5$) when $p = 4, 8$ and $\sigma^2 = 0.1(0.2)0.7$

			$\alpha = 0.05$				
p	$(n^{(1)}, n^{(2)})$	σ^2	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	0.1	0.058	0.063	0.064	0.049	0.057
		0.3	0.052	0.052	0.054	0.047	0.051
		0.5	0.049	0.047	0.051	0.045	0.048
		0.7	0.047	0.045	0.049	0.045	0.047
	(10,20)	0.1	0.050	0.049	0.051	0.049	0.050
		0.3	0.048	0.047	0.049	0.047	0.048
		0.5	0.049	0.049	0.050	0.047	0.048
		0.7	0.051	0.051	0.051	0.048	0.050
	(20,10)	0.1	0.061	0.068	0.069	0.044	0.060
		0.3	0.061	0.068	0.063	0.049	0.058
		0.5	0.059	0.064	0.059	0.050	0.056
		0.7	0.056	0.060	0.056	0.050	0.054
	(20,20)	0.1	0.052	0.053	0.053	0.049	0.052
		0.3	0.051	0.051	0.051	0.049	0.050
		0.5	0.050	0.049	0.050	0.049	0.050
		0.7	0.049	0.048	0.050	0.049	0.049
	AAD		0.362	0.613	0.471	0.211	0.298
8	(10,10)	0.1	0.069	0.063	0.000	0.044	0.095
		0.3	0.054	0.052	0.000	0.040	0.062
		0.5	0.048	0.047	0.000	0.037	0.052
		0.7	0.046	0.045	0.000	0.036	0.049
	(10,20)	0.1	0.051	0.050	0.103	0.047	0.052
		0.3	0.049	0.047	0.092	0.044	0.048
		0.5	0.051	0.050	0.099	0.044	0.050
		0.7	0.055	0.054	0.081	0.046	0.054
	(20,10)	0.1	0.086	0.080	0.000	0.038	0.130
		0.3	0.078	0.081	0.000	0.050	0.093
		0.5	0.071	0.075	0.000	0.051	0.077
		0.7	0.065	0.069	0.003	0.050	0.067
	(20,20)	0.1	0.057	0.059	0.086	0.048	0.058
		0.3	0.054	0.055	0.083	0.049	0.053
		0.5	0.051	0.051	0.073	0.048	0.050
		0.7	0.050	0.049	0.069	0.047	0.049
	AAD		0.921	0.959	4.266	0.533	1.541

Note : $\text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/16,\ j = 1, 2, 3, 4, 5$

Table 2 *(Continued)*

p	$(n^{(1)}, n^{(2)})$	σ^2	$\alpha = 0.01$				
			$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	0.1	0.013	0.014	0.015	0.009	0.012
		0.3	0.010	0.010	0.012	0.009	0.010
		0.5	0.009	0.009	0.010	0.008	0.009
		0.7	0.009	0.008	0.010	0.008	0.009
	(10,20)	0.1	0.010	0.010	0.011	0.010	0.010
		0.3	0.009	0.009	0.010	0.009	0.009
		0.5	0.010	0.009	0.010	0.010	0.009
		0.7	0.010	0.010	0.010	0.009	0.010
	(20,10)	0.1	0.014	0.017	0.018	0.008	0.014
		0.3	0.014	0.016	0.015	0.009	0.013
		0.5	0.013	0.014	0.013	0.010	0.012
		0.7	0.012	0.013	0.012	0.010	0.011
	(20,20)	0.1	0.011	0.011	0.011	0.010	0.010
		0.3	0.010	0.010	0.011	0.010	0.010
		0.5	0.010	0.010	0.010	0.010	0.010
		0.7	0.010	0.009	0.010	0.009	0.010
	AAD		0.124	0.206	0.176	0.091	0.104
8	(10,10)	0.1	0.012	0.010	0.000	0.006	0.022
		0.3	0.010	0.009	0.000	0.006	0.013
		0.5	0.009	0.009	0.000	0.006	0.010
		0.7	0.009	0.009	0.000	0.006	0.010
	(10,20)	0.1	0.010	0.010	0.040	0.009	0.011
		0.3	0.010	0.009	0.033	0.008	0.009
		0.5	0.010	0.010	0.038	0.008	0.010
		0.7	0.011	0.011	0.034	0.008	0.011
	(20,10)	0.1	0.016	0.012	0.000	0.004	0.035
		0.3	0.015	0.014	0.000	0.007	0.021
		0.5	0.014	0.014	0.000	0.008	0.016
		0.7	0.013	0.013	0.002	0.008	0.014
	(20,20)	0.1	0.012	0.012	0.038	0.009	0.012
		0.3	0.011	0.011	0.026	0.009	0.011
		0.5	0.010	0.010	0.021	0.009	0.010
		0.7	0.010	0.010	0.019	0.009	0.010
	AAD		0.174	0.144	1.552	0.266	0.423

$$\text{Note}: \ \text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/16, \ j = 1, 2, 3, 4, 5$$

Table 3 Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_5$) when $p = 4, 8$ and $\sigma^2 = 2$, 5,10,20,30

p	$(n^{(1)}, n^{(2)})$	σ^2	$\alpha = 0.05$				
			$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	2	0.048	0.047	0.050	0.045	0.048
		5	0.055	0.057	0.058	0.048	0.054
		10	0.057	0.063	0.064	0.048	0.057
		20	0.059	0.064	0.068	0.047	0.059
		30	0.058	0.063	0.069	0.045	0.058
	(10,20)	2	0.058	0.064	0.059	0.050	0.055
		5	0.062	0.070	0.066	0.048	0.060
		10	0.061	0.068	0.069	0.044	0.060
		20	0.058	0.063	0.070	0.040	0.059
		30	0.057	0.060	0.071	0.038	0.058
	(20,10)	2	0.049	0.049	0.050	0.047	0.048
		5	0.049	0.047	0.050	0.048	0.049
		10	0.050	0.049	0.051	0.049	0.050
		20	0.051	0.051	0.053	0.050	0.051
		30	0.051	0.052	0.053	0.050	0.051
	(20,20)	2	0.050	0.049	0.050	0.049	0.049
		5	0.052	0.052	0.052	0.050	0.051
		10	0.052	0.054	0.053	0.049	0.051
		20	0.052	0.053	0.053	0.049	0.051
		30	0.051	0.053	0.054	0.048	0.052
	AAD		0.437	0.727	0.816	0.299	0.415
8	(10,10)	2	0.049	0.047	0.000	0.037	0.053
		5	0.060	0.056	0.000	0.042	0.073
		10	0.069	0.062	0.000	0.044	0.095
		20	0.075	0.065	0.000	0.041	0.119
		30	0.077	0.064	0.000	0.037	0.133
	(10,20)	2	0.071	0.075	0.000	0.051	0.077
		5	0.083	0.083	0.000	0.047	0.107
		10	0.086	0.080	0.000	0.038	0.130
		20	0.086	0.074	0.000	0.029	0.150
		30	0.086	0.071	0.000	0.024	0.160
	(20,10)	2	0.051	0.050	0.099	0.044	0.050
		5	0.049	0.047	0.092	0.045	0.049
		10	0.051	0.050	0.103	0.047	0.052
		20	0.053	0.054	0.040	0.048	0.056
		30	0.054	0.055	0.004	0.048	0.057
	(20,20)	2	0.051	0.051	0.073	0.048	0.050
		5	0.056	0.058	0.096	0.049	0.055
		10	0.057	0.059	0.086	0.048	0.058
		20	0.056	0.058	0.001	0.045	0.059
		30	0.055	0.057	0.000	0.043	0.059
	AAD		1.395	1.144	4.516	0.739	3.229

Note : $\text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/20,\ j = 1, 2, 3, 4, 5$

Table 3 (*Continued*)

p	$(n^{(1)}, n^{(2)})$	σ^2	$\alpha = 0.01$				
			$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
4	(10,10)	2	0.009	0.009	0.010	0.008	0.009
		5	0.011	0.012	0.013	0.009	0.011
		10	0.013	0.014	0.015	0.009	0.012
		20	0.013	0.015	0.017	0.009	0.013
		30	0.013	0.015	0.018	0.008	0.013
	(10,20)	2	0.013	0.014	0.013	0.009	0.011
		5	0.014	0.017	0.016	0.009	0.013
		10	0.014	0.017	0.018	0.008	0.014
		20	0.013	0.015	0.019	0.007	0.014
		30	0.013	0.015	0.019	0.006	0.013
	(20,10)	2	0.010	0.009	0.010	0.009	0.009
		5	0.010	0.009	0.010	0.009	0.009
		10	0.010	0.010	0.010	0.009	0.010
		20	0.011	0.010	0.011	0.010	0.010
		30	0.010	0.010	0.011	0.010	0.010
	(20,20)	2	0.010	0.010	0.010	0.010	0.010
		5	0.011	0.011	0.011	0.010	0.011
		10	0.011	0.011	0.011	0.010	0.011
		20	0.011	0.011	0.011	0.009	0.011
		30	0.011	0.011	0.012	0.009	0.011
	AAD		0.161	0.263	0.331	0.110	0.151
8	(10,10)	2	0.009	0.009	0.000	0.006	0.010
		5	0.011	0.009	0.000	0.006	0.015
		10	0.013	0.010	0.000	0.006	0.022
		20	0.014	0.009	0.000	0.005	0.031
		30	0.014	0.009	0.000	0.004	0.038
	(10,20)	2	0.014	0.014	0.000	0.008	0.016
		5	0.016	0.014	0.000	0.006	0.026
		10	0.016	0.012	0.000	0.004	0.035
		20	0.017	0.011	0.000	0.002	0.046
		30	0.018	0.011	0.000	0.002	0.053
	(20,10)	2	0.051	0.050	0.099	0.044	0.050
		5	0.049	0.047	0.092	0.045	0.049
		10	0.051	0.050	0.103	0.047	0.052
		20	0.053	0.054	0.040	0.048	0.056
		30	0.054	0.055	0.004	0.048	0.057
	(20,20)	2	0.051	0.051	0.073	0.048	0.050
		5	0.056	0.058	0.096	0.049	0.055
		10	0.057	0.059	0.086	0.048	0.058
		20	0.056	0.058	0.001	0.045	0.059
		30	0.055	0.057	0.000	0.043	0.059
	AAD		1.395	1.144	4.516	0.739	3.229

Note : $\text{AAD} = \sum |100\widehat{\alpha}_j - 100\alpha|/20,\ j = 1, 2, 3, 4, 5$

3.6 Conclusion

We considered testing the equality of two mean vectors with unequal covariance matrices. In the case of unequal covariance matrices, we proposed two approximate solutions to the problem by adjusting the degrees of freedom of the F distribution. Asymptotic expansions up to the term of order N^{-2} for the first and second moments of U are given, where N is the total sample size minus two. Also, the bias correction for the asymptotic expansions using the consistent estimators instead of the unknown parameters are also obtained. As for the asymptotic expansions of the moments of U, we note that the result in this paper is an extension of Yanagihara and Yuan (2005) in the meaning of asymptotic expansions up to the term of higher order. However, it should be noted that S and BC procedures do not always yield the improved approximate solution since the S and BC procedures do not make use of asymptotic expansions of the T statistic itself. In conclusion, it may be seen from simulations that the approximate upper percentiles of the null distribution of T by S and BC Procedure are better than those of other procedures. When the difference between covariance matrices is large, and $n^{(1)}$ and $n^{(2)}$ are small, it seems that BC-Procedure yields good approximations. Finally, S and BC Procedure are expected to yield good approximations when the other cases, although it must be checked.

Appendix

In this Appendix, we present some results of expectation.

A.1.

Let $\boldsymbol{u} \sim N_p(\boldsymbol{0}, I)$ with A and B are $p \times p$ symmetric matrices, then

(1) $\mathrm{E}[\boldsymbol{u}'A\boldsymbol{u}] = \mathrm{tr}A$,

(2) $\mathrm{E}[\boldsymbol{u}'\boldsymbol{u}(\boldsymbol{u}'A\boldsymbol{u})] = (p+2)(\mathrm{tr}A)$,

(3) $\mathrm{E}[(\boldsymbol{u}'A\boldsymbol{u})^2] = 2(\mathrm{tr}A^2) + (\mathrm{tr}A)^2$,

(4) $\mathrm{E}[(\boldsymbol{u}'A\boldsymbol{u})(\boldsymbol{u}'B\boldsymbol{u})] = 2(\mathrm{tr}AB) + (\mathrm{tr}A)(\mathrm{tr}B)$,

(5) $\mathrm{E}[(\boldsymbol{u}'A\boldsymbol{u})^3] = 8(\mathrm{tr}A^3) + 6(\mathrm{tr}A^2)(\mathrm{tr}A) + (\mathrm{tr}A)^3$,

(6) $\mathrm{E}[(\boldsymbol{u}'A\boldsymbol{u})^2(\boldsymbol{u}'B\boldsymbol{u})] = 8(\mathrm{tr}A^2B) + 4(\mathrm{tr}AB)(\mathrm{tr}A) + 2(\mathrm{tr}A^2)(\mathrm{tr}B) + (\mathrm{tr}A)^2(\mathrm{tr}B)$,

(7) $\mathrm{E}[(\boldsymbol{u}'A\boldsymbol{u})^4] = 48(\mathrm{tr}A^4) + 32(\mathrm{tr}A^3)(\mathrm{tr}A) + 12(\mathrm{tr}A^2)(\mathrm{tr}A)^2 + 12(\mathrm{tr}A^2)^2 + (\mathrm{tr}A)^4$.

A.2.

Let $nS \sim W_p(n, \Sigma)$ and $V = \sqrt{n}(S - \Sigma)$ with A, B and C are $p \times p$ symmetric matrices, then

(1) $\mathrm{E}[(\mathrm{tr}AV)^2] = 2\ \mathrm{tr}(A\Sigma)^2$,

(2) $\mathrm{E}[\mathrm{tr}(AV)^2] = \mathrm{tr}(A\Sigma)^2 + (\mathrm{tr}A\Sigma)^2$,

(3) $\mathrm{E}[(\mathrm{tr}AVBV)] = (\mathrm{tr}A\Sigma B\Sigma) + (\mathrm{tr}A\Sigma)(\mathrm{tr}B\Sigma)$,

(4) $\mathrm{E}[\{\mathrm{tr}(AV)\}^3] = \dfrac{8}{\sqrt{n}}\ \mathrm{tr}(A\Sigma)^3,$

(5) $\mathrm{E}[\mathrm{tr}(AV)^3] = \dfrac{1}{\sqrt{n}}\ [4\ \mathrm{tr}(A\Sigma)^3 + 3\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}A\Sigma) + (\mathrm{tr}A\Sigma)^3],$

(6) $\mathrm{E}[(\mathrm{tr}AV)\{\mathrm{tr}(AV)^2\}] = \dfrac{4}{\sqrt{n}}\ [\mathrm{tr}(A\Sigma)^3 + (\mathrm{tr}A\Sigma)\{\mathrm{tr}(A\Sigma)^2\}],$

(7) $\mathrm{E}[(\mathrm{tr}AV)^2(\mathrm{tr}BV)] = \dfrac{8}{\sqrt{n}}\ \{\mathrm{tr}(A\Sigma)^2 B\Sigma\},$

(8) $\mathrm{E}[(\mathrm{tr}AV)^4] = 12\{\mathrm{tr}(A\Sigma)^2\}^2 + \mathrm{O}(N^{-1})$

(9) $\mathrm{E}[\mathrm{tr}(AV)^4]$
$= 5\ \mathrm{tr}(A\Sigma)^4 + 4\{\mathrm{tr}(A\Sigma)^3\}(\mathrm{tr}A\Sigma) + 2\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}A\Sigma)^2 + \{\mathrm{tr}(A\Sigma)^2\}^2 + \mathrm{O}(N^{-1}),$

(10) $\mathrm{E}[(\mathrm{tr}AV)\{\mathrm{tr}(AV)^3\}] = 6[\ \mathrm{tr}(A\Sigma)^4 + (\mathrm{tr}A\Sigma)\{\mathrm{tr}(A\Sigma)^3\}] + \mathrm{O}(N^{-1}),$

(11) $\mathrm{E}[\{\mathrm{tr}(AV)^2\}^2]$
$= 5\{\mathrm{tr}(A\Sigma)^2\}^2 + 4\ \mathrm{tr}(A\Sigma)^4 + 2\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}A\Sigma)^2 + (\mathrm{tr}A\Sigma)^4 + \mathrm{O}(N^{-1}),$

(12) $\mathrm{E}[(\mathrm{tr}AV)^2\{\mathrm{tr}(AV)^2\}]$
$= 8\ \mathrm{tr}(A\Sigma)^4 + 2\{\mathrm{tr}(A\Sigma)^2\}^2 + 2\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}A\Sigma)^2 + \mathrm{O}(N^{-1}),$

(13) $\mathrm{E}[(\mathrm{tr}AV)^3(\mathrm{tr}BV)] = 12\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}A\Sigma B\Sigma) + \mathrm{O}(N^{-1}),$

(14) $\mathrm{E}[(\mathrm{tr}AV)^2(\mathrm{tr}BV)^2] = 8(\mathrm{tr}A\Sigma B\Sigma)^2 + 4\{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + \mathrm{O}(N^{-1}),$

(15) $\mathrm{E}[(\mathrm{tr}AV)^2(\mathrm{tr}BV)(\mathrm{tr}CV)]$
$= 8(\mathrm{tr}A\Sigma B\Sigma)(\mathrm{tr}A\Sigma C\Sigma) + 4\{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}B\Sigma C\Sigma) + \mathrm{O}(N^{-1}).$

A.3.

Let $n^{(i)}S^{(i)} \sim W_p(n^{(i)}, \Sigma)$ and $V^{(i)} = \sqrt{n^{(i)}}\left(S^{(i)} - \Sigma\right)$ with A, B, C and D are $p \times p$ symmetric matrices where $i = 1, 2$, then

(1) $\mathrm{E}[\mathrm{tr}\left(AV^{(1)}\right)^2\left(BV^{(2)}\right)^2]$
$= \mathrm{tr}(A\Sigma)^2(B\Sigma)^2 + (\mathrm{tr}A\Sigma)\{\mathrm{tr}A\Sigma(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2B\Sigma\}(\mathrm{tr}B\Sigma)$
$+(\mathrm{tr}A\Sigma)(\mathrm{tr}B\Sigma)(\mathrm{tr}A\Sigma B\Sigma),$

(2) $\mathrm{E}[(\mathrm{tr}AV^{(1)}BV^{(2)})^2] = 2\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + (\mathrm{tr}A\Sigma B\Sigma)^2,$

(3) $\mathrm{E}[\{\mathrm{tr}\left(AV^{(1)}\right)^2\}\{\mathrm{tr}\left(BV^{(2)}\right)^2\}]$
$= \{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}B\Sigma)^2 + (\mathrm{tr}A\Sigma)^2\{\mathrm{tr}(B\Sigma)^2\}$
$+(\mathrm{tr}A\Sigma)^2(\mathrm{tr}B\Sigma)^2,$

(4) $\mathrm{E}[(\mathrm{tr}AV^{(1)})^2(\mathrm{tr}BV^{(2)})^2] = 4\{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\},$

(5) $\mathrm{E}[\{\mathrm{tr}\left(AV^{(1)}\right)^2\}(\mathrm{tr}BV^{(2)})^2] = 2\{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + 2(\mathrm{tr}A\Sigma)^2\{\mathrm{tr}(B\Sigma)^2\},$

(6) $\mathrm{E}[(\mathrm{tr}AV^{(1)})\{\mathrm{tr}AV^{(1)}\left(BV^{(2)}\right)^2\}] = 2\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\} + 2\{\mathrm{tr}(A\Sigma)^2B\Sigma\}(\mathrm{tr}B\Sigma),$

(7) $\mathrm{E}[(\mathrm{tr}AV^{(1)})(\mathrm{tr}BV^{(2)})(\mathrm{tr}AV^{(1)}BV^{(2)})] = 4\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\},$

(8) $\mathrm{E}[\mathrm{tr}AV^{(1)}BV^{(1)}CV^{(2)}DV^{(2)}]$
$= (\mathrm{tr}A\Sigma B\Sigma C\Sigma D\Sigma) + (\mathrm{tr}A\Sigma B\Sigma C\Sigma)(\mathrm{tr}D\Sigma) + (\mathrm{tr}A\Sigma C\Sigma D\Sigma)(\mathrm{tr}B\Sigma)$
$+(\mathrm{tr}A\Sigma C\Sigma)(\mathrm{tr}B\Sigma)(\mathrm{tr}D\Sigma).$

A.4.

The following results are presented as supplementary expectations:

(1) $\mathrm{E}[\mathrm{tr}\left(\Omega^{(2)\prime}\Omega^{(1)}V^{(1)}\Omega^{(1)\prime}\Omega^{(2)}V^{(2)}\right)^2]$
$= 3\mathrm{tr}\left(\Omega^{(1)}\Omega^{(1)\prime}\Omega^{(2)}\Omega^{(2)\prime}\right)^2 + (\mathrm{tr}\Omega^{(1)}\Omega^{(1)\prime}\Omega^{(2)}\Omega^{(2)\prime})^2,$

(2) $\mathrm{E}[(\mathrm{tr}\Omega^{(1)\prime}\Omega^{(1)}V^{(1)})(\mathrm{tr}\Omega^{(2)\prime}\Omega^{(1)}V^{(1)}\Omega^{(1)\prime}\Omega^{(2)}V^{(2)}\Omega^{(2)\prime}\Omega^{(2)}V^{(2)})]$
$= 2\{\mathrm{tr}\left(\Omega^{(1)\prime}\Omega^{(1)}\right)^2\Omega^{(2)\prime}\Omega^{(2)}\}(\mathrm{tr}\Omega^{(2)\prime}\Omega^{(2)}) + 2\{\mathrm{tr}\left(\Omega^{(1)\prime}\Omega^{(1)}\right)^2\left(\Omega^{(2)\prime}\Omega^{(2)}\right)^2\},$

(3) $\mathrm{E}[(\mathrm{tr}\Omega^{(2)\prime}\Omega^{(1)}V^{(1)}\Omega^{(1)\prime}\Omega^{(2)}V^{(2)})^2]$
$= 2\mathrm{tr}\left(\Omega^{(1)}\Omega^{(1)\prime}\Omega^{(2)}\Omega^{(2)\prime}\right)^2 + 2(\mathrm{tr}\Omega^{(1)}\Omega^{(1)\prime}\Omega^{(2)}\Omega^{(2)\prime})^2,$

(4) $\mathrm{E}[(\mathrm{tr}\Omega^{(1)\prime}\Omega^{(1)}V^{(1)})(\mathrm{tr}\Omega^{(2)\prime}\Omega^{(2)}V^{(2)})(\mathrm{tr}\Omega^{(2)\prime}\Omega^{(1)}V^{(1)}\Omega^{(1)\prime}\Omega^{(2)}V^{(2)})]$
$= 4\{\mathrm{tr}\left(\Omega^{(1)\prime}\Omega^{(1)}\right)^2\left(\Omega^{(2)\prime}\Omega^{(2)}\right)^2\},$

where the notations are defined by Section 3.3.

Bibliography

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

[2] 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.

[3] Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **160**, 268–282.

[4] Bennett, B. M. (1951). Note on a solution of the generalized Behrens-Fisher problem. *Annals of the Institute of Statistical Mathematics*, **2**, 87–90.

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

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

[7] Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric multivariate and related distributions*. London; Chapman & Hall / CRC.

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

[9] Girón, F. J. and del Castillo, C. (2010). The multivariate Behrens-Fisher distribution. *Journal of Multivariate Analysis*, **101**, 2091–2102.

[10] James, G. S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. *Biometrika*, **41**, 19–43.

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

[12] Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. *Biometrika*, **67**, 85–92.

[13] 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.

[14] Kawasaki, T. and Seo, T. (2015). A two sample test for mean vectors with unequal covariance matrices. *Communications in Statistics - Simulation and Computation*, **44**, 1850–1866.

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

[16] Kim, S. J. (1992). A practical solution to the multivariate Behrens-Fisher problem. *Biometrika*, **79**, 171–176.

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

[18] Krishnamoorthy, K. and Yu, J. (2004). Modified Nel and Van der Merwe test for the multivariate Bahrens-Fisher problem. *Statistics & Probability Letters*, **66**, 161–169.

[19] Krishnamoorthy, K. and Yu, J. (2012). Multivariate Behrens-Fisher problem with missing data. *Journal of Multivariate Analysis*, **105**, 141–150.

[20] Little, R. J. and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. Wiley Series in Probability and Statistics.

[21] Nel, D. G. and van der Merwe, C. A. (1986). A solution to the multivariate Behrens-Fisher problem. *Communications in Statistics - Theory and Methods*, **15**, 3719–3735.

[22] Nel, D. G., van der Merwe, C. A. and Moser, B. K. (1990). The exact distributions of the univariate and multivariate Behrens-Fisher statistics with a comparison of several solutions in the univariate case. *Communications in Statistics - Theory and Methods*, **19**, 279–298.

[23] Rubin, D. B. (1976). Inference and missing data. *Biometrika*, **63**, 581–592.

[24] Scheffé, H. (1943). On solutions of the Behrens-Fisher problem, based on the t-distribution. *The Annals of Mathematical Statistics*, **14**, 35–44.

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

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

[27] 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.

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

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

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

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

[32] Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **29**, 350–362.

[33] Yanagihara, H. and Yuan, K. (2005). Three approximate solutions to the multivariate Behrens-Fisher problem. *Communications in Statistics - Simulation and Computation*, **34**, 975–988.

[34] Yao, Y. (1965). An approximate degrees of freedom solution to the multivariate Behrens-Fisher problem. *Biometrika*, **52**, 139–147.

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

[36] Wu, Y., Genton, M. G. and Stefanski, A. (2006). A multivariate two-sample mean test for small sample size and missing data. *Biometrics*, **62**, 877–885.