A Two Sample Test for Mean Vectors with Unequal Covariance Matrices

Tamae Kawasaki[1] and Takashi Seo[2]

[1]Department of Mathematical Information Science
Graduate School of Science, Tokyo University of Science, Tokyo, Japan
[2]Department of Mathematical Information Science
Faculty of Science, Tokyo University of Science, Tokyo, Japan

Abstract

In this paper, we consider 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 T^2 type test statistic does not follow the F distribution, and it is also difficult to derive the exact distribution. In this study, we propose an approximate solution to the problem by adjusting the degrees of freedom of the F distribution. That is, we derive an extension of the results derived by Yanagihara and Yuan (2005). 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, and a new result of the approximate degrees of freedom is obtained. Finally, numerical comparison is presented by a Monte Carlo simulation.

Keywords Approximate degrees of freedom; F approximation; Hotelling's T^2 statistic; Multivariate Behrens-Fisher problem; Two sample problem.

1 Introduction

Let $\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{ij}, \ldots, \boldsymbol{x}_{in_i}$ 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:

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs.} \quad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2, \tag{1.1}$$

where $\Sigma_1 \neq \Sigma_2$. A natural statistic for testing (1.1) 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

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

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 (1.1), $(n-p-1)T/\{p(n-2)\}$ follows the F distribution with p and $n-p-1$ degrees of freedom, where $n = n_1 + n_2$.

To consider the tests for equality of two mean vectors is a fundamental problem. 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. The problem concerns the difference between the mean vectors of two normal populations with the case of a monotone missing pattern when $\Sigma_1 = \Sigma_2$, as suggested by Seko, Kawasaki and Seo (2011). 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. Our goal is to give a new approximate solution by an extension of Yanagihara and Yuan (2005).

The following section presents a derivation of the main result by approximate degrees of freedom and presents the proof. In Section 3, we compare four 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.

2 Approximate Degrees of Freedom

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

$$T = \boldsymbol{z}' W^{-1} \boldsymbol{z} = \frac{\boldsymbol{z}'\boldsymbol{z}}{U}, \tag{2.1}$$

where

$$\boldsymbol{z} = \sqrt{\frac{n_1 n_2}{n}} \overline{\Sigma}^{-1/2} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2), \quad W = \overline{\Sigma}^{-1/2} \left(\frac{n_2}{n} S_1 + \frac{n_1}{n} S_2 \right) \overline{\Sigma}^{-1/2},$$

$$\overline{\Sigma} = \frac{n_2}{n} \Sigma_1 + \frac{n_1}{n} \Sigma_2, \quad U = \frac{\boldsymbol{z}'\boldsymbol{z}}{\boldsymbol{z}' W^{-1} \boldsymbol{z}}.$$

By approximating the distribution of U as

$$U \approx \frac{\chi_\nu^2}{\phi}, \tag{2.2}$$

we have

$$\frac{\nu}{p\phi} T \approx \frac{\chi_p^2/p}{\chi_\nu^2/\nu} \sim F_{p,\nu}.$$

Note that when $\Sigma_1 = \Sigma_2$ and $n_1 = n_2$, U is exactly distributed as χ_ν^2/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.

Theorem 2.1 *Let $U = \boldsymbol{z}'\boldsymbol{z}/\boldsymbol{z}'W^{-1}\boldsymbol{z}$ be defined by (2.1). Then, an asymptotic expansion up to the term of order N^{-2} for $E[U]$ can be expanded as*

$$\mathrm{E}[U] = 1 - \frac{\theta_1}{N} + \frac{1}{N^2}(\theta_2 - \theta_3) + O(N^{-3}), \tag{2.3}$$

where

$$N = n - 2, \quad \theta_1 = \frac{1}{p(p+2)} \sum_{i=1}^{2} c_i \{ p(a_i^{(1)})^2 + (p-2) a_i^{(2)} \},$$

$$\theta_2 = \frac{1}{p(p+2)(p+4)} \sum_{i=1}^{2} d_i \{4p^2 a_i^{(3)} + (p-2)(3p+4) a_i^{(1)} a_i^{(2)} + p(p+2)(a_i^{(1)})^3\},$$

$$\begin{aligned}\theta_3 =& \frac{1}{p(p+2)(p+4)(p+6)} \Bigg\{ \sum_{i=1}^{2} c_i^2 \left\{p^2(5p+14) a_i^{(4)} + 4(p+3)(p+2)(p-2) a_i^{(1)} a_i^{(3)} \right. \\ & \left. + p(p+3)(p-2)(a_i^{(2)})^2 + 2(p^3+5p^2+7p+6) a_i^{(2)} (a_i^{(1)})^2 - p(p+4)(a_i^{(1)})^4 \right\} \\ & + 4(p+3)(p+2)(p-2)\psi_1 + 4p(p+2)(p-2)\psi_2 + 4p(p+4)(p+2)\psi_3 - 2p(p-2)\psi_4 \\ & - 2(p+3)(p-2)\psi_5 - 2p(p+4)(p-2)\psi_6 - 2p(p+4)\psi_7 + 2p(p+4)(3p+2)\psi_8 \Bigg\},\end{aligned}$$

with $\psi_k, k = 1, 2, \ldots, 7$ *given by*

$$\psi_1 = c_1 c_2 \{a_1^{(1)} b^{(1,2,1)} + a_2^{(1)} b^{(2,1,1)}\}, \quad \psi_2 = c_1 c_2 b^{(2,2,1)},$$

$$\psi_3 = c_1 c_2 a_1^{(1)} a_2^{(1)} b^{(1,1,1)}, \quad \psi_4 = c_1 c_2 a_1^{(2)} a_2^{(2)}, \quad \psi_5 = c_1 c_2 \{a_1^{(2)} (a_2^{(1)})^2 + (a_1^{(1)})^2 a_2^{(2)}\},$$

$$\psi_6 = c_1 c_2 (b^{(1,1,1)})^2, \quad \psi_7 = c_1 c_2 (a_1^{(1)})^2 (a_2^{(1)})^2, \quad \psi_8 = c_1 c_2 b^{(1,1,2)},$$

and

$$\begin{aligned}& c_i = \frac{(n-n_i)^2(n-2)}{n^2(n_i-1)}, \quad d_i = \frac{(n-n_i)^3(n-2)^2}{n^3(n_i-1)^2}, \quad a_i^{(\ell)} = \mathrm{tr}(\Sigma_i \overline{\Sigma}^{-1})^\ell, \ i = 1, 2, \ \ell = 1, 2, 3, 4, \\ & b^{(q,r,s)} = \mathrm{tr}\{(\Sigma_1 \overline{\Sigma}^{-1})^q (\Sigma_2 \overline{\Sigma}^{-1})^r\}^s, \ (q, r, s) = (1,1,1), (1,1,2), (1,2,1), (2,1,2), (2,2,1).\end{aligned}$$

Proof. Let

$$\rho_i = \sqrt{\frac{n_i - 1}{n-2}} \ (i = 1, 2), \quad \Omega_i = \sqrt{\frac{n - n_i}{n}} \overline{\Sigma}^{-\frac{1}{2}} \Sigma_i^{\frac{1}{2}}$$

and

$$V_i = \sqrt{n_i - 1}(\Sigma_i^{-1/2} S_i \Sigma_i^{-1/2} - I_p), \ i = 1, 2.$$

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

$$W^{-1} = I_p - \frac{1}{\sqrt{N}} \overline{V} + \frac{1}{N} \overline{V}^2 - \frac{1}{N\sqrt{N}} \overline{V}^3 + \frac{1}{N^2} \overline{V}^4 + O_p(N^{-5/2}), \tag{2.4}$$

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 (2.4) that we can expand U as

$$\begin{aligned}U =& 1 + \frac{1}{\sqrt{N}}Q_1 + \frac{1}{N}(Q_1^2 - Q_2) + \frac{1}{N\sqrt{N}}(Q_3 - 2Q_1Q_2 + Q_1^3) \\ &+ \frac{1}{N^2}(Q_1^4 - Q_4 + 2Q_1Q_3 + Q_2^2 - 3Q_1^2Q_2) + O_p(N^{-5/2}),\end{aligned}$$

where $Q_i = \boldsymbol{z}'\overline{V}^i\boldsymbol{z}/\boldsymbol{z}'\boldsymbol{z}, i = 1,2,3,4$. 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) $\mathrm{E}[Q_1] = 0$,

(ii) $\mathrm{E}[Q_2] = \frac{1}{p}\sum_{i=1}^{2} c_i\{(a_i^{(1)})^2 + a_i^{(2)}\}$,

$$\mathrm{E}[Q_1^2] = \frac{2}{p(p+2)}\sum_{i=1}^{2} c_i\{(a_i^{(1)})^2 + 2a_i^{(2)}\},$$

(iii) $\sqrt{N}\mathrm{E}[Q_3] = \frac{1}{p}\sum_{i=1}^{2} d_i\{4a_i^{(3)} + 3a_i^{(1)}a_i^{(2)} + (a_i^{(1)})^3\}$,

$$\sqrt{N}\mathrm{E}[Q_1Q_2] = \frac{2}{p(p+2)}\sum_{i=1}^{2} d_i\{6a_i^{(3)} + 5a_i^{(1)}a_i^{(2)} + (a_i^{(1)})^3\},$$

$$\sqrt{N}\mathrm{E}[Q_1^3] = \frac{8}{p(p+2)(p+4)}\sum_{i=1}^{2} d_i\{8a_i^{(3)} + 6a_i^{(1)}a_i^{(2)} + (a_i^{(1)})^3\},$$

(iv) $\mathrm{E}[Q_4] = \frac{1}{p}\left\{\sum_{i=1}^{2} c_i^2\{5a_i^{(4)} + 4a_i^{(1)}a_i^{(3)} + (a_i^{(2)})^2 + 2a_i^{(2)}(a_i^{(1)})^2\} + 2(2\psi_1 + 2\psi_2 + 2\psi_3 + \psi_6 + 3\psi_8)\right\} + O(N^{-1})$,

$$\begin{aligned}\mathrm{E}[Q_1^4] =& \frac{12}{p(p+2)(p+4)(p+6)}\left\{\sum_{i=1}^{2} c_i^2\{48a_i^{(4)} + 32a_i^{(1)}a_i^{(3)} + 12(a_i^{(2)})^2 + 12a_i^{(2)}(a_i^{(1)})^2 + (a_i^{(1)})^4\}\right. \\ &\left. + 2(16\psi_1 + 32\psi_2 + 8\psi_3 + 4\psi_4 + 2\psi_5 + 8\psi_6 + \psi_7 + 16\psi_8)\right\} + O(N^{-1}),\end{aligned}$$

$$\begin{aligned}\mathrm{E}[Q_1Q_3] =& \frac{2}{p(p+2)}\left\{\sum_{i=1}^{2} c_i^2\{8a_i^{(4)} + 7a_i^{(1)}a_i^{(3)} + (a_i^{(2)})^2 + 2a_i^{(2)}(a_i^{(1)})^2\}\right. \\ &\left. + (7\psi_1 + 10\psi_2 + 4\psi_3 + 2\psi_6 + 6\psi_8)\right\} + O(N^{-1}),\end{aligned}$$

$$\mathrm{E}[Q_2^2] = \frac{1}{p(p+2)}\Bigg\{\sum_{i=1}^{2} c_i^2\{14a_i^{(4)} + 8a_i^{(1)}a_i^{(3)} + 7(a_i^{(2)})^2 + (a_i^{(1)})^4 + 6a_i^{(2)}(a_i^{(1)})^2\}$$
$$+2(4\psi_1 + 4\psi_2 + 4\psi_3 + \psi_4 + \psi_5 + 6\psi_6 + \psi_7 + 10\psi_8)\Bigg\} + O(N^{-1}),$$
$$\mathrm{E}[Q_1^2 Q_2] = \frac{2}{p(p+2)(p+4)}\Bigg\{\sum_{i=1}^{2} c_i^2\{40a_i^{(4)} + 28a_i^{(1)}a_i^{(3)} + 10(a_i^{(2)})^2 + 11a_i^{(2)}(a_i^{(1)})^2 + (a_i^{(1)})^4\}$$
$$+28\psi_1 + 48\psi_2 + 16\psi_3 + 4\psi_4 + 3\psi_5 + 16\psi_6 + 2\psi_7 + 32\psi_8\Bigg\} + O(N^{-1}).$$

Using the above results, we can show (2.3). This completes the proof of Theorem 2.1.

In addition, we have the following result.

Corollary 2.1 *If $\Sigma_1 = \Sigma_2$, $n_1 = n_2$, then*

$$\mathrm{E}[U] = 1 - \frac{1}{N}(p-1) + \frac{1}{N^2}\cdot\frac{p^2(p-2)}{(p+2)(p+6)} + O(N^{-3}).$$

Similarly, as the result of asymptotic expansion for $\mathrm{E}[U^2]$,we have the following theorem.

Theorem 2.2 *Let $U = \boldsymbol{z}'\boldsymbol{z}/\boldsymbol{z}'W^{-1}\boldsymbol{z}$ be defined by (2.1). Then, an asymptotic expansion up to the term of order N^{-2} for $E[U^2]$ can be expanded as*

$$\mathrm{E}[U^2] = 1 - \frac{2}{N}(\theta_1 - \theta_4) + \frac{1}{N^2}(2\theta_5 - \theta_6) + O(N^{-3}), \qquad (2.5)$$

where

$$\theta_4 = \frac{1}{p(p+2)}\sum_{i=1}^{2} c_i\left\{(a_i^{(1)})^2 + 2a_i^{(2)}\right\},$$
$$\theta_5 = \frac{1}{p(p+2)(p+4)}\sum_{i=1}^{2} d_i\left\{4(p^2-3p+4)a_i^{(3)} + 3p(p-4)a_i^{(1)}a_i^{(2)} + p^2(a_i^{(1)})^3\right\},$$

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

Proof. In the same way, we can expand U^2 as

$$
\begin{aligned}
U^2 =& 1 + \frac{2}{\sqrt{N}}Q_1 + \frac{1}{N}(3Q_1^2 - 2Q_2) + \frac{2}{N\sqrt{N}}(Q_3 - 3Q_1Q_2 + 2Q_1^3) \\
&+ \frac{1}{N^2}(5Q_1^4 - 2Q_4 + 6Q_1Q_3 + 3Q_2^2 - 12Q_1^2Q_2) + O_p(N^{-5/2}).
\end{aligned}
$$

By calculating the expectations of the above results, we can show (2.5). This completes the proof of Theorem 2.2.

In addition, we have the following result.

Corollary 2.2 *If* $\Sigma_1 = \Sigma_2$, $n_1 = n_2$, *then*

$$
\mathrm{E}[U^2] = 1 - \frac{2}{N}(p-2) + \frac{1}{N^2} \cdot \frac{p^5 + 6p^4 - p^3 - 92p^2 - 60p + 144}{(p+2)(p+4)(p+6)} + O(N^{-3}).
$$

It follows from (2.2) that

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

Therefore, using the asymptotic expansions of $\mathrm{E}[U]$ and $\mathrm{E}[U^2]$ in Theorems 2.1 and 2.2, the new approximation to the values of ν and ϕ are given by

$$
\nu_{KS} = \frac{2(N^2 - N\theta_1 + \theta_2 - \theta_3)^2}{N^2(N^2 - 2N\theta_1 + 2N\theta_4 + 2\theta_5 - \theta_6) - (N^2 - N\theta_1 + \theta_2 - \theta_3)^2}, \tag{2.6}
$$

$$
\phi_{KS} = \frac{N^2\nu}{N^2 - N\theta_1 + \theta_2 - \theta_3}, \tag{2.7}
$$

respectively. We can propose a new procedure as follows.

(I) High Order Procedure

$$T_{KS} = \frac{\nu_{KS}}{p\phi_{KS}} T \overset{a}{\sim} F_{p,\nu_{KS}}$$

where ν_{KS} and ϕ_{KS} are given by (2.6) and (2.7), respectively, and where "$\overset{a}{\sim}$" means "approximately following".

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

$$\nu_{KS} = \frac{(N - \theta_1)^2}{N\theta_4 - \theta_1^2/2} (= \nu_Y), \quad \phi_{KS} = \frac{\nu_Y N}{N - \theta_1} (= \phi_Y), \tag{2.8}$$

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

$$\nu_M = \frac{(N - \theta_1)^2}{N\theta_4 - \theta_1}, \tag{2.9}$$

which will be used to obtain the F approximation.

3 Numerical Studies

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

(II) Yanagihara and Yuan's (2005) Procedure

$$T_Y = \frac{\nu_Y}{p\phi_Y} T \overset{a}{\sim} F_{p,\nu_Y}$$

where ν_Y and ϕ_Y are given by (2.8).

(III) Modified Yanagihara and Yuan's (2005) Procedure

$$T_M = \frac{\nu_M}{p\phi_M} T \overset{a}{\sim} F_{p,\nu_M}$$

where

$$\phi_M = \frac{\nu_M N}{N - \theta_1}$$

and ν_M is given by (2.9).

(IV) Modified Bartlett Procedure (see Yanagihara and Yuan (2005))

$$T_{MB} = (N\widehat{\beta}_1 + \widehat{\beta}_2) \log \left(1 + \frac{T}{N\widehat{\beta}_1}\right) \overset{a}{\sim} \chi_p^2,$$

where

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

$$\widehat{\gamma}_1 = \frac{1}{p}\sum_{i=1}^{2} c_i\{(\widehat{a}_i^{(1)})^2 + \widehat{a}_i^{(2)}\}, \quad \widehat{\gamma}_2 = \frac{1}{p(p+2)}\sum_{i=1}^{2} c_i\{2(p+3)(\widehat{a}_i^{(1)})^2 + 2(p+4)\widehat{a}_i^{(2)}\},$$

$$\widehat{a}_i^{(\ell)} = \mathrm{tr}(S_i\overline{S}^{-1})^\ell, \quad \overline{S} = \frac{n_2}{n}S_1 + \frac{n_1}{n}S_2.$$

For each of parameter, 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 four procedures:

$$\text{(I)} \quad \alpha_1 = \mathrm{P}(T_{KS} > F_{\alpha;p,\nu_{KS}}), \qquad \text{(II)} \quad \alpha_2 = \mathrm{P}(T_Y > F_{\alpha;p,\nu_Y}),$$

$$\text{(III)} \quad \alpha_3 = \mathrm{P}(T_M > F_{\alpha;p,\nu_M}), \qquad \text{(IV)} \quad \alpha_4 = \mathrm{P}(T_{MB} > c_{\alpha;p}),$$

where $F_{\alpha;m,n}$ is the upper 100α percentile of the F distribution with m and n degrees of freedom and $c_{\alpha;p}$ 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), (50, 50), (50, 80), (80, 50), (80, 80) for (I)$\sim$(IV). We note that the second degree of freedom of F distribution for the test statistic (I)$\sim$(III) changes with each simulation of 1,000,000 trials. In practical use, we must estimate $a_i^{(\ell)}$ and $b^{(q,r,s)}$ for (I)$\sim$(III) since Σ_i and $\overline{\Sigma}$ are unknown. In this paper, we use the consistent estimators of $a_i^{(\ell)}$ and $b^{(q,r,s)}$, which are the same as that of procedure (IV).

Table 1 presents the empirical sizes $\widehat{\alpha}_j, j = 1, 2, 3, 4$ in the case of $\Sigma_1 = \mathrm{diag}(\eta, \eta^2, \cdots, \eta^p)$ and $\Sigma_2 = I$, where $\eta = 1$, 5(5)20. We note that $|\Sigma_1| \geq 1$ and the difference between Σ_1 and Σ_2 is large when η is large. Tables 2 and 3 present the empirical sizes $\widehat{\alpha}_j, j = 1, 2, 3, 4$ in the case of $\Sigma_1 = \sigma^2 I$ and $\Sigma_2 = I$, where $\sigma^2 = 0.1(0.2)0.9,1$ in Table 2, and $\sigma^2 = 2$, 5(5)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).

Table 1

Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_4$) when $p = 4, 8$, and $\eta = 1,\ 5(5)20$

				$\alpha = 0.05$				$\alpha = 0.01$			
p	n_1	n_2	η	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
4	10	10	1	0.047	0.048	0.044	0.046	0.009	0.010	0.008	0.009
			5	0.057	0.067	0.045	0.057	0.013	0.017	0.008	0.013
			10	0.055	0.069	0.042	0.057	0.012	0.018	0.007	0.013
			15	0.054	0.070	0.040	0.057	0.012	0.019	0.007	0.013
			20	0.054	0.070	0.039	0.057	0.012	0.019	0.007	0.013
	10	20	1	0.053	0.053	0.049	0.051	0.011	0.011	0.009	0.010
			5	0.057	0.070	0.039	0.059	0.013	0.019	0.007	0.013
			10	0.054	0.070	0.036	0.057	0.012	0.019	0.006	0.013
			15	0.053	0.070	0.034	0.057	0.012	0.019	0.005	0.013
			20	0.053	0.070	0.033	0.056	0.011	0.019	0.005	0.013
	20	10	1	0.053	0.053	0.049	0.051	0.011	0.011	0.009	0.010
			5	0.051	0.052	0.049	0.051	0.010	0.011	0.010	0.010
			10	0.051	0.053	0.049	0.051	0.010	0.011	0.010	0.011
			15	0.051	0.053	0.049	0.051	0.010	0.011	0.010	0.010
			20	0.050	0.053	0.049	0.051	0.010	0.011	0.009	0.010
	20	20	1	0.049	0.049	0.048	0.049	0.010	0.010	0.009	0.009
			5	0.051	0.053	0.049	0.052	0.010	0.011	0.009	0.010
			10	0.051	0.054	0.048	0.052	0.010	0.012	0.009	0.011
			15	0.051	0.053	0.048	0.051	0.010	0.012	0.009	0.010
			20	0.050	0.053	0.047	0.051	0.010	0.012	0.009	0.010
AAD				0.322	1.098	0.661	0.439	0.124	0.476	0.217	0.168
8	10	10	1	0.046	0.000	0.035	0.048	0.009	0.000	0.005	0.009
			5	0.077	0.000	0.030	0.149	0.018	0.000	0.004	0.052
			10	0.073	0.000	0.022	0.160	0.018	0.000	0.003	0.061
			15	0.071	0.000	0.019	0.165	0.018	0.000	0.003	0.064
			20	0.069	0.000	0.017	0.167	0.018	0.000	0.002	0.066
	10	20	1	0.060	0.028	0.048	0.060	0.012	0.013	0.008	0.012
			5	0.079	0.000	0.017	0.162	0.020	0.000	0.002	0.061
			10	0.073	0.000	0.012	0.169	0.020	0.000	0.001	0.067
			15	0.070	0.000	0.011	0.171	0.019	0.000	0.001	0.029
			20	0.068	0.000	0.010	0.172	0.019	0.000	0.001	0.071
	20	10	1	0.059	0.028	0.048	0.059	0.012	0.013	0.008	0.012
			5	0.053	0.000	0.046	0.058	0.011	0.000	0.009	0.013
			10	0.052	0.000	0.044	0.059	0.011	0.000	0.008	0.013
			15	0.052	0.000	0.043	0.058	0.011	0.000	0.007	0.013
			20	0.052	0.000	0.043	0.059	0.011	0.000	0.007	0.013
	20	20	1	0.050	0.068	0.047	0.049	0.010	0.019	0.009	0.010
			5	0.053	0.000	0.041	0.059	0.011	0.000	0.007	0.013
			10	0.052	0.000	0.038	0.059	0.011	0.000	0.006	0.013
			15	0.051	0.000	0.037	0.058	0.010	0.000	0.006	0.013
			20	0.051	0.000	0.037	0.058	0.010	0.000	0.006	0.013
AAD				0.011	0.046	0.018	0.050	0.004	0.009	0.005	0.021

Note : AAD $= \sum |100\widehat{\alpha} - 100\alpha|/20$

Table 2

Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_4$) when $p = 4, 8$ and $\sigma^2 = 0.1(0.2)0.9,\ 1$

				$\alpha = 0.05$				$\alpha = 0.01$			
p	n_1	n_2	σ^2	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
4	10	10	0.1	0.058	0.064	0.049	0.057	0.013	0.015	0.009	0.012
			0.3	0.052	0.054	0.047	0.051	0.010	0.012	0.009	0.010
			0.5	0.049	0.051	0.045	0.048	0.009	0.010	0.008	0.009
			0.7	0.047	0.049	0.045	0.047	0.009	0.010	0.008	0.009
			0.9	0.047	0.049	0.044	0.046	0.009	0.010	0.008	0.009
			1	0.047	0.048	0.044	0.046	0.009	0.010	0.008	0.009
	10	20	0.1	0.050	0.051	0.049	0.050	0.010	0.011	0.010	0.010
			0.3	0.048	0.049	0.047	0.048	0.009	0.010	0.009	0.009
			0.5	0.049	0.050	0.047	0.048	0.010	0.010	0.010	0.009
			0.7	0.051	0.051	0.048	0.050	0.010	0.010	0.009	0.010
			0.9	0.053	0.053	0.048	0.051	0.011	0.011	0.009	0.010
			1	0.053	0.053	0.049	0.051	0.011	0.011	0.009	0.010
	20	10	0.1	0.061	0.069	0.044	0.060	0.014	0.018	0.008	0.014
			0.3	0.061	0.063	0.049	0.058	0.014	0.015	0.009	0.013
			0.5	0.059	0.059	0.050	0.056	0.013	0.013	0.010	0.012
			0.7	0.056	0.056	0.050	0.054	0.012	0.012	0.010	0.011
			0.9	0.054	0.054	0.049	0.052	0.011	0.011	0.009	0.010
			1	0.053	0.053	0.049	0.051	0.011	0.011	0.009	0.010
	20	20	0.1	0.052	0.053	0.049	0.052	0.011	0.011	0.010	0.010
			0.3	0.051	0.051	0.049	0.050	0.010	0.011	0.010	0.010
			0.5	0.050	0.050	0.049	0.050	0.010	0.010	0.010	0.010
			0.7	0.049	0.050	0.049	0.049	0.010	0.010	0.009	0.010
			0.9	0.049	0.049	0.048	0.049	0.010	0.010	0.009	0.010
			1	0.049	0.049	0.048	0.049	0.010	0.010	0.009	0.009
AAD				0.335	0.387	0.227	0.263	0.112	0.140	0.096	0.089

Note : $\text{AAD} = \sum |100\widehat{\alpha} - 100\alpha|/24$

Table 2

(*Continued*)

p	n_1	n_2	σ^2	$\alpha = 0.05$ $\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\alpha = 0.01$ $\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
8	10	10	0.1	0.069	0.000	0.044	0.095	0.012	0.000	0.006	0.022
			0.3	0.054	0.000	0.040	0.062	0.010	0.000	0.006	0.013
			0.5	0.048	0.000	0.037	0.052	0.009	0.000	0.006	0.010
			0.7	0.046	0.000	0.036	0.049	0.009	0.000	0.006	0.010
			0.9	0.046	0.000	0.035	0.048	0.009	0.000	0.005	0.010
			1	0.046	0.000	0.035	0.048	0.009	0.000	0.005	0.009
	10	20	0.1	0.051	0.103	0.047	0.052	0.010	0.040	0.009	0.011
			0.3	0.049	0.092	0.044	0.048	0.010	0.033	0.008	0.009
			0.5	0.051	0.099	0.044	0.050	0.010	0.038	0.008	0.010
			0.7	0.055	0.081	0.046	0.054	0.011	0.034	0.008	0.011
			0.9	0.057	0.043	0.047	0.057	0.011	0.020	0.008	0.011
			1	0.060	0.028	0.048	0.060	0.012	0.013	0.008	0.012
	20	10	0.1	0.086	0.000	0.038	0.130	0.016	0.000	0.004	0.035
			0.3	0.078	0.000	0.050	0.093	0.015	0.000	0.007	0.021
			0.5	0.071	0.000	0.051	0.077	0.014	0.000	0.008	0.016
			0.7	0.065	0.003	0.050	0.067	0.013	0.002	0.008	0.014
			0.9	0.061	0.017	0.048	0.062	0.012	0.008	0.008	0.012
			1	0.059	0.028	0.048	0.059	0.012	0.013	0.008	0.012
	20	20	0.1	0.057	0.086	0.048	0.058	0.012	0.038	0.009	0.012
			0.3	0.054	0.083	0.049	0.053	0.011	0.026	0.009	0.011
			0.5	0.051	0.073	0.048	0.050	0.010	0.021	0.009	0.010
			0.7	0.050	0.069	0.047	0.049	0.010	0.019	0.009	0.010
			0.9	0.049	0.067	0.047	0.049	0.010	0.018	0.009	0.009
			1	0.050	0.068	0.047	0.049	0.010	0.019	0.009	0.010
AAD				0.809	3.758	0.547	1.211	0.157	1.265	0.260	0.322

Note : $\text{AAD} = \sum |100\widehat{\alpha} - 100\alpha|/24$

Table 3

Empirical sizes ($\widehat{\alpha}_1 \sim \widehat{\alpha}_4$) when $p = 4, 8$ and $\sigma^2 = 2,\ 5(5)30$

				$\alpha = 0.05$				$\alpha = 0.01$			
p	n_1	n_2	σ^2	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
4	10	10	2	0.048	0.050	0.045	0.048	0.009	0.010	0.008	0.009
			5	0.055	0.058	0.048	0.054	0.011	0.013	0.009	0.011
			10	0.057	0.064	0.048	0.057	0.013	0.015	0.009	0.012
			15	0.059	0.066	0.047	0.058	0.013	0.016	0.009	0.013
			20	0.059	0.068	0.047	0.059	0.013	0.017	0.009	0.013
			25	0.058	0.069	0.045	0.059	0.013	0.018	0.009	0.013
			30	0.058	0.069	0.045	0.058	0.013	0.018	0.008	0.013
	10	20	2	0.058	0.059	0.050	0.055	0.013	0.013	0.009	0.011
			5	0.062	0.066	0.048	0.060	0.014	0.016	0.009	0.013
			10	0.061	0.069	0.044	0.060	0.014	0.018	0.008	0.014
			15	0.059	0.070	0.042	0.059	0.014	0.019	0.007	0.014
			20	0.058	0.070	0.040	0.059	0.013	0.019	0.007	0.014
			25	0.057	0.070	0.038	0.058	0.013	0.019	0.007	0.014
			30	0.057	0.071	0.038	0.058	0.013	0.019	0.006	0.013
	20	10	2	0.049	0.050	0.047	0.048	0.010	0.010	0.009	0.009
			5	0.049	0.050	0.048	0.049	0.010	0.010	0.009	0.009
			10	0.050	0.051	0.049	0.050	0.010	0.010	0.009	0.010
			15	0.051	0.052	0.050	0.051	0.010	0.011	0.010	0.010
			20	0.051	0.053	0.050	0.051	0.011	0.011	0.010	0.010
			25	0.051	0.053	0.049	0.051	0.010	0.011	0.010	0.010
			30	0.051	0.053	0.050	0.051	0.010	0.011	0.010	0.010
	20	20	2	0.050	0.050	0.049	0.049	0.010	0.010	0.010	0.010
			5	0.052	0.052	0.050	0.051	0.011	0.011	0.010	0.011
			10	0.052	0.053	0.049	0.051	0.011	0.011	0.010	0.011
			15	0.052	0.053	0.049	0.052	0.011	0.011	0.010	0.011
			20	0.052	0.053	0.049	0.051	0.011	0.011	0.009	0.011
			25	0.051	0.053	0.048	0.051	0.011	0.011	0.009	0.011
			30	0.051	0.054	0.048	0.052	0.011	0.012	0.009	0.011
AAD				0.449	0.890	0.326	0.435	0.168	0.366	0.114	0.160

Note : $\text{AAD} = \sum |100\widehat{\alpha} - 100\alpha|/28$

Table 3

(*Continued*)

p	n_1	n_2	σ^2	$\alpha = 0.05$ $\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\alpha = 0.01$ $\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
8	10	10	2	0.049	0.000	0.037	0.053	0.009	0.000	0.006	0.010
			5	0.060	0.000	0.042	0.073	0.011	0.000	0.006	0.015
			10	0.069	0.000	0.044	0.095	0.013	0.000	0.006	0.022
			15	0.073	0.000	0.043	0.109	0.013	0.000	0.005	0.027
			20	0.075	0.000	0.041	0.119	0.014	0.000	0.005	0.031
			25	0.077	0.000	0.039	0.127	0.014	0.000	0.004	0.035
			30	0.077	0.000	0.037	0.133	0.014	0.000	0.004	0.038
	10	20	2	0.071	0.000	0.051	0.077	0.014	0.000	0.008	0.016
			5	0.083	0.000	0.047	0.107	0.016	0.000	0.006	0.026
			10	0.086	0.000	0.038	0.130	0.016	0.000	0.004	0.035
			15	0.073	0.000	0.043	0.109	0.013	0.000	0.005	0.027
			20	0.086	0.000	0.029	0.150	0.017	0.000	0.002	0.046
			25	0.087	0.000	0.026	0.156	0.018	0.000	0.002	0.050
			30	0.086	0.000	0.024	0.160	0.018	0.000	0.002	0.053
	20	10	2	0.051	0.099	0.044	0.050	0.010	0.039	0.008	0.010
			5	0.049	0.092	0.045	0.049	0.010	0.033	0.008	0.010
			10	0.051	0.103	0.047	0.052	0.010	0.040	0.009	0.011
			15	0.053	0.084	0.048	0.055	0.011	0.036	0.009	0.011
			20	0.053	0.040	0.048	0.056	0.011	0.018	0.009	0.012
			25	0.054	0.013	0.048	0.057	0.011	0.007	0.009	0.012
			30	0.054	0.004	0.048	0.057	0.011	0.002	0.009	0.012
	20	20	2	0.051	0.073	0.048	0.050	0.010	0.021	0.009	0.010
			5	0.056	0.096	0.049	0.055	0.011	0.034	0.009	0.011
			10	0.057	0.086	0.048	0.058	0.012	0.038	0.009	0.012
			15	0.056	0.015	0.046	0.059	0.012	0.008	0.008	0.013
			20	0.056	0.001	0.045	0.059	0.012	0.001	0.008	0.013
			25	0.055	0.000	0.043	0.059	0.011	0.000	0.007	0.013
			30	0.055	0.000	0.043	0.059	0.011	0.000	0.007	0.013
AAD				1.446	4.496	0.758	3.486	0.267	1.289	0.351	1.128

Note : $\text{AAD} = \sum |100\widehat{\alpha} - 100\alpha|/28$

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

In conclusion, when the difference between covariance matrices is large, the approximate upper percentile of the null distribution of T by the method (high order procedure) proposed in this paper is better than those of other procedures.

Appendix

In this Appendix, we present some results of expectation:

A.1

Let $\boldsymbol{u} \sim N_p(\mathbf{0}, I)$ 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 $S \sim W_p(n, \Sigma)$ 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) $\mathrm{E}[(\mathrm{tr}AVBV)] = (\mathrm{tr}AVBV) + (\mathrm{tr}A\Sigma)(\mathrm{tr}B\Sigma)$,

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

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

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

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

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

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

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

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

A.3

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

(1) $\mathrm{E}[\mathrm{tr}(AV_1)^2(BV_2)^2] = \mathrm{tr}(A\Sigma)^2(B\Sigma)^2 + (\mathrm{tr}A\Sigma)\{\mathrm{tr}A\Sigma(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2B\Sigma\}(\mathrm{tr}B\Sigma)$
$+(\mathrm{tr}A\Sigma)(\mathrm{tr}B\Sigma)(\mathrm{tr}A\Sigma B\Sigma)$,

(2) $\mathrm{E}[(\mathrm{tr}AV_1BV_2)^2] = 2\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + (\mathrm{tr}A\Sigma B\Sigma)^2$,

(3) $\mathrm{E}[\{\mathrm{tr}(AV_1)^2\}\{\mathrm{tr}(BV_2)^2\}] = \{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + \{\mathrm{tr}(A\Sigma)^2\}(\mathrm{tr}B\Sigma)^2$
$+(\mathrm{tr}A\Sigma)^2\{\mathrm{tr}(B\Sigma)^2\} + (\mathrm{tr}A\Sigma)^2(\mathrm{tr}B\Sigma)^2$,

(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) $\mathrm{E}[\{\mathrm{tr}(AV_1)^2\}(\mathrm{tr}BV_2)^2] = 2\{\mathrm{tr}(A\Sigma)^2\}\{\mathrm{tr}(B\Sigma)^2\} + 2(\mathrm{tr}A\Sigma)^2\{\mathrm{tr}(B\Sigma)^2\}$,

(6) $\mathrm{E}[(\mathrm{tr}AV_1)\{\mathrm{tr}AV_1(BV_2)^2\}] = 2\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\} + 2\{\mathrm{tr}(A\Sigma)^2B\Sigma\}(\mathrm{tr}B\Sigma)$,

(7) $\mathrm{E}[(\mathrm{tr}AV_1)(\mathrm{tr}BV_2)(\mathrm{tr}AV_1BV_2)] = 4\{\mathrm{tr}(A\Sigma)^2(B\Sigma)^2\}$,

(8) $\mathrm{E}[\mathrm{tr}AV_1BV_1CV_2DV_2] = (\mathrm{tr}A\Sigma B\Sigma C\Sigma D\Sigma) + (\mathrm{tr}A\Sigma B\Sigma C\Sigma)(\mathrm{tr}D\Sigma)$
$+(\mathrm{tr}A\Sigma C\Sigma D\Sigma)(\mathrm{tr}B) + (\mathrm{tr}A\Sigma C\Sigma)(\mathrm{tr}B\Sigma)(\mathrm{tr}D\Sigma)$.

A.4

The following results are presented as supplementary expectations:

(1) $\mathrm{E}[\mathrm{tr}(\Omega_2'\Omega_1V_1\Omega_1'\Omega_2V_2)^2] = 3\mathrm{tr}(\Omega_1\Omega_1'\Omega_2\Omega_2')^2 + (\mathrm{tr}\Omega_1\Omega_1'\Omega_2\Omega_2')^2$,

(2) $\mathrm{E}[(\mathrm{tr}\Omega_1'\Omega_1V_1)(\mathrm{tr}\Omega_2'\Omega_1V_1\Omega_1'\Omega_2V_2\Omega_2'\Omega_2V_2)] = 2\{\mathrm{tr}(\Omega_1'\Omega_1)^2\Omega_2'\Omega_2\}(\mathrm{tr}\Omega_2'\Omega_2)$
$+2\{\mathrm{tr}(\Omega_1'\Omega_1)^2(\Omega_2'\Omega_2)^2\}$,

(3) $\mathrm{E}[(\mathrm{tr}\Omega_2'\Omega_1V_1\Omega_1'\Omega_2V_2)^2] = 2\mathrm{tr}(\Omega_1\Omega_1'\Omega_2\Omega_2')^2 + 2(\mathrm{tr}\Omega_1\Omega_1'\Omega_2\Omega_2')^2$,

(4) $\mathrm{E}[(\mathrm{tr}\Omega_1'\Omega_1V_1)(\mathrm{tr}\Omega_2'\Omega_2V_2)(\mathrm{tr}\Omega_2'\Omega_1V_1\Omega_1'\Omega_2V_2)] = 4\{\mathrm{tr}(\Omega_1'\Omega_1)^2(\Omega_2'\Omega_2)^2\}$,

where the notations are defined by Section 2.

Acknowledgments

Second author's research was in part supported by Grant-Aid for Scientific Research (C) (23500360).

References

[1] Bennett, B. M. (1951). Note on a solution of the generalized Behrens-Fisher problem. *Ann. Inst. Statist. Math.*, **2**, 87–90.

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

[3] Fujikoshi, Y. (2000). Transformations with improved chi-square approximations. *J. Multivariate Anal.*, **72**, 249–263.

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

[5] 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.

[6] 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.

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

[8] Krishnamoorthy, K. and Yu, J. (2004). Modified Nel and Van der Merwe test for the multivariate Bahrens-Fisher problem. *Statist. Probab. Lett.*, **66**, 161–169.

[9] Krishnamoorthy, K. and Yu, J. (2012). Multivariate Behrens-Fisher problem with mussing data. *J. Multivariate Anal.*, **105**, 141–150.

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

[11] 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. *Comm. Statist., Theory Methods*, **19**, 279–298.

[12] Scheffé, H. (1943). On solutions of the Behrens-Fisher problem, based on the t-distribution. *Ann. Math. Statist.*, **14**, 35–44.

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

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

[15] Yanagihara, H. and Yuan, K. (2005). Three approximate solutions to the multivariate Behrens-Fisher problem. *Comm. Statist., Simulation Comput.*, **34**, 975–988.

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