Hiroshima Math. J.
32 (2002), 17–50

Asymptotic expansions of the null distributions of test statistics for multivariate linear hypothesis under nonnormality

Hirofumi Wakaki, Hirokazu Yanagihara and Yasunori Fujikoshi

(Received May 2, 2001)
(Revised August 21, 2001)

Abstract. This paper is concerned with the distributions of some test statistics for a multivariate linear hypothesis under nonnormality. The test statistics considered include the likelihood ratio statistic, the Lawley-Hotelling trace criterion and the Bartlett-Nanda-Pillai trace criterion, under normality. We derive asymptotic expansions of the null distributions of these test statistics up to the order n^{-1}, where n is the sample size, under nonnormality. It is shown that our general results can be effectively obtained by deriving an asymptotic expansion of the distribution of a multivariate t-statistic. As special cases of our general results our asymptotic expansions are given for Hotelling's T^2 statistic, one-way MANOVA test statistics, etc. Numerical accuracies of asymptotic expansion approximations are examined. The validity of the expansions is also discussed. Moreover, we will find conditions such that the Bartlett correction in the normal case implies an improved χ^2-approximation, even under nonnormality.

1. Introduction

We consider a multivariate linear model

$$Y = X\Xi + \mathscr{E},$$

where $Y = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_n)'$ is an $n \times p$ observation matrix of p response variables, $X = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)'$ is an $n \times k$ design matrix of k explanatory variables with full rank k $(< n)$, Ξ is a $k \times p$ unknown parameter matrix and $\mathscr{E} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n)'$ is an $n \times p$ error matrix. It is assumed that each vector $\boldsymbol{\varepsilon}_j$ is independently and identically distributed with $\mathrm{E}(\boldsymbol{\varepsilon}_j) = \mathbf{0}$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}_j) = \Sigma$.

For testing a linear hypothesis

$$H_0 : H\Xi = 0,$$

where H is a known $h \times k$ matrix with rank h $(\le k)$, let S_h and S_e be the variation matrices due to the hypothesis and the error, respectively, i.e.,

$$S_h = \hat{\Xi}' H' \{H(X'X)^{-1} H'\}^{-1} H\hat{\Xi}, \qquad S_e = Y'\{I_n - X(X'X)^{-1}X'\}Y,$$

2000 *Mathematics Subject Classification*. 62H10, 62E20.
Key words and phrases. Bartlett correction, Chi-squared approximation, Edgeworth expansion, Multivariate t-statistic, Robustness, Validity of expansion.

where $\hat{\Xi} = (X'X)^{-1}X'Y$. Then the following three criteria have been used, in particular, under normality.

(i) the likelihood ratio statistic:

$$T_{LR} = -(n - k + d_1) \log(|S_e|/|S_e + S_h|),$$

(ii) the Lawley-Hotelling trace criterion:

$$T_{HL} = (n - k + d_2) \operatorname{tr}(S_h S_e^{-1}),$$

(iii) the Bartlett-Nanda-Pillai trace criterion:

$$T_{BNP} = (n - k + d_3) \operatorname{tr}\{S_h(S_h + S_e)^{-1}\},$$

where the constants d_j's are the Bartlett corrections in the normal case, and they are given as follows:

$$d_1 = -\frac{p - h + 1}{2}, \qquad d_2 = -(p + 1), \qquad d_3 = h.$$

Under normality, the distributions of these statistics have been extensively studied, see e.g., Anderson [1] and Siotani, Hayakawa and Fujikoshi [23]. Under nonnormality it is shown that the null distributions of these statistics converge to χ^2_{ph} as the sample size n tends to infinity under an appropriate regularity condition on the design matrix (see Huber [15]). Our aim is to obtain asymptotic expansions of the null distributions of these statistics up to the order n^{-1} under a general condition.

As for the results of the usual asymptotic expansions under nonnormality, Kano [17] and Fujikoshi [10] independently derived an asymptotic expansion for the distribution of Hotelling's T^2 statistic. Fujikoshi, Ohmae and Yanagihara [13] obtained an asymptotic expansion of the null distributions for one-way ANOVA test statistics. Recently, Fujikoshi [12] derived such expansions in the cases of one-way MANOVA test statistics. For a univariate linear model, Qumsiyeh [22] derived an asymptotic expansion for the least squares estimate of regression coefficients. Using this result, Yanagihara [24] derived an asymptotic expansion of the null distribution of the likelihood ratio statistic for testing a linear hypothesis about regression coefficients. Our work is a generalization of these results.

One of the approaches for solving our problem will be to use an asymptotic expansion of the joint distribution of

$$Z = (X'X)^{1/2}(\hat{\Xi} - \Xi) \qquad \text{and} \qquad \sqrt{n}\left(\frac{1}{n}S_e - \Sigma\right).$$

This approach was used by Fujikoshi, Ohmae and Yanagihara [13] for one-way ANOVA test statistic, by Fujikoshi [12] for one-way MANOVA test statistics, and by Yanagihara [24] for a univariate linear model. However, for a multivariate linear hypothesis this approach leads to a prohibited calculation. In order to solve our problem effectively, we consider the distribution of a key statistic

$$U = Z(n^{-1}S_e)^{-1/2}.$$

The statistic U may be a multivariate t-statistic.

Note that the three statistics can be expressed in terms of U as

$$T_G = \mathrm{tr}(U'\Omega U) + \frac{1}{n}[(r_1 - k)\ \mathrm{tr}(U'\Omega U) + r_2\ \mathrm{tr}\{(U'\Omega U)^2\}] + \mathrm{O}_p(n^{-3/2}), \quad (1.1)$$

where

$$\Omega = (X'X)^{-1/2}H'\{H(X'X)^{-1}H'\}^{-1}H(X'X)^{-1/2}. \quad (1.2)$$

Here, the constants r_1 and r_2 are defined as follows;

$$\begin{aligned} &\text{(i)} \quad T_{LR} : r_1 = d_1, \qquad r_2 = -1/2, \\ &\text{(ii)} \quad T_{HL} : r_1 = d_2, \qquad r_2 = 0, \\ &\text{(iii)} \quad T_{BNP} : r_1 = d_3, \qquad r_2 = -1. \end{aligned} \quad (1.3)$$

Needless to say, we have $\Omega^2 = \Omega$ and $\mathrm{rank}(\Omega) = h$. Using an expansion of the distribution function of U, we will obtain an asymptotic expansion of the null distribution of T_G as

$$\mathrm{P}(T_G \le x) = G_{ph}(x) + \frac{1}{n}\sum_{j=0}^{3} b_j G_{ph+2j}(x) + \mathrm{o}(n^{-1}), \quad (1.4)$$

where G_f is the distribution function of a central chi-squared distribution with f degrees of freedom. In other words, our main purposes are to get a formula for b_j's in asymptotic expansion (1.4) and conditions for valid expansion up to the order n^{-1}.

On the other hand, by using (1.4), the expectation of T_G can be expanded as

$$\mathrm{E}(T_G) = ph\left(1 + \frac{c_1}{n}\right) + \mathrm{o}(n^{-1}).$$

Note that under normality, $c_1 = 0$ from the definitions of d_1, d_2 and d_3. Furthermore, although the error vectors are distributed as a certain nonnormal distribution, we shall obtain some conditions of X and H such that $c_1 = 0$. This means that even under nonnormality, the Bartlett correction in the normal case has given an improvement for the mean in a χ^2-approximation. These conditions are made more clear by using the coefficients b_j's in the asymptotic expansion (1.4). Based on such conditions, it is possible to see whether a test statistic is robust for nonnormality or not.

The present paper is organized in the following way. In §2, we state a main result on an asymptotic expansion of the null distribution of T_G. Some applications for our result are given in §3. In §4, numerical accuracies are studied for Hotelling's T^2 statistic. In §5, by using coefficients of an asymptotic expansion, we will obtain conditions such that the Bartlett correction in the normal case implies an improved χ^2-approximation, even under nonnormality. This gives an advantage for obtaining an asymptotic expansion. Note that such a result cannot be expected for Bootstrap method. Our derivation and its validity are discussed in Appendix. In Appendix 1, we prepare some basic results for the validity of asymptotic expansions for the distribution functions of U and T_G up to the order n^{-1}. We derive an asymptotic expansion of the distribution of U in Appendix 2. Based on the expansion of the distribution of U, we obtain an expansion of the null distribution of T_G, by expanding the characteristic function of T_G. An outline of the computation is given in Appendix 3.

2. Asymptotic expansion of distribution function of T_G

In this section, we state a main result on an asymptotic expansion of the null distribution of T_G in (1.1) up to the order n^{-1}. Without loss of generality, we may replace $\mathscr{E}$ by $\mathscr{E}\Sigma^{-1/2}$, and then $\mathrm{E}[\mathrm{vec}(\mathscr{E})] = \mathbf{0}$ and $\mathrm{Cov}[\mathrm{vec}(\mathscr{E})] = I_{np}$, where $\mathrm{vec}(A) = (\boldsymbol{a}_1', \dots, \boldsymbol{a}_m')'$ for any $n \times m$ matrix $A = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_m)$, since T_G is invariant under the transformation from Y to $Y\Sigma^{-1/2}$. Let $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_1$, and let $\mu_{i_1\dots i_m}$ be a moment of $\boldsymbol{\varepsilon}$ defined by

$$\mu_{i_1\dots i_m} = \mathrm{E}[\varepsilon_{i_1} \dots \varepsilon_{i_m}],$$

where ε_j denotes the jth element of $\boldsymbol{\varepsilon}$. Similarly the corresponding cumulant of $\boldsymbol{\varepsilon}$ is denoted by $\kappa_{i_1\dots i_m}$, e.g.,

$$\kappa_{abc} = \mu_{abc}, \qquad \kappa_{abcd} = \mu_{abcd} - \delta_{ab}\delta_{cd} - \delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc},$$

where δ_{ab} is the Kronecker delta, i.e., $\delta_{aa} = 1$ and $\delta_{ab} = 0$ for $a \neq b$.

Let λ_n be the smallest eigenvalue of $X'X$ and $M_n = \max\{\|\boldsymbol{x}_j\| : j = 1, \dots, n\}$, where $\|\cdot\|$ denotes the Euclidean norm. Furthermore, $\boldsymbol{t}$ is a p-dimensional vector and $T_2 = [t_{ab}^{(2)}(1 + \delta_{ab})/2]$ is a $p \times p$ symmetric matrix whose norm is defined by $\|T_2\| = \left[\sum_{a=1}^p \sum_{b=1}^p \{t_{ab}^{(2)}(1 + \delta_{ab})\}^2/4\right]^{1/2}$. Suppose that X and the distribution of $\boldsymbol{\varepsilon}$ satisfy the following assumptions A1, A2, A3, B1 and B2.

A1. $\limsup_{n\to\infty} \dfrac{1}{n} \sum_{j=1}^n \|\boldsymbol{x}_j\|^4 < \infty$, A2. $\liminf_{n\to\infty} \dfrac{\lambda_n}{n} > 0$,

A3. For some constant $\delta > 0$, $M_n = \mathrm{O}(n^{1/2-\delta})$,

B1. $\mathrm{E}(\|\boldsymbol{\varepsilon}\|^8) < \infty$,

B2. Cramér's condition for the joint distribution of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'$ hold, that is, for any $b > 0$,

$$\sup_{\|\boldsymbol{t}\|+\|T_2\|>b} |\mathrm{E}[\exp\{i\boldsymbol{t}'\boldsymbol{\varepsilon} + i\ \mathrm{tr}(\boldsymbol{\varepsilon}'T_2\boldsymbol{\varepsilon})\}]| < 1.$$

Then the distribution function of T_G can be expanded as in Theorem 2.1.

Theorem 2.1. *Under the assumptions* A1, A2, A3, B1 *and* B2, *the null distribution of* T_G *can be expanded as* (1.4), *where the coefficients* b_j *are given by*

$$\begin{aligned}
b_0 &= a_1\kappa_4^{(1)} - \{a_2 + (h-2)(h+1)a_5\}\kappa_3^{(1)} \\
&\quad - \{a_3 - 3ha_4 - (h-2)a_5\}\kappa_3^{(2)} + \frac{1}{4}ph\{(h-p-1) - 2r_1\}, \\
b_1 &= -2a_1\kappa_4^{(1)} + \{3a_2 - 6a_4 + (3h^2+h-6)a_5\}\kappa_3^{(1)} \\
&\quad + \{3a_3 - 3(3h+2)a_4 + (h+6)a_5\}\kappa_3^{(2)} \\
&\quad - \frac{1}{2}ph\{h - r_1 + r_2(h+p+1)\}, \qquad (2.1) \\
b_2 &= a_1\kappa_4^{(1)} - \{3a_2 - 12a_4 + (h+2)(3h-1)a_5\}\kappa_3^{(1)} \\
&\quad - \{3a_3 - 3(3h+4)a_4 + 5(h+2)a_5\}\kappa_3^{(2)} + \frac{1}{4}ph(h+p+1)(1+2r_2), \\
b_3 &= \{a_2 - 6a_4 + (h+1)(h+2)a_5\}\kappa_3^{(1)} \\
&\quad + \{a_3 - 3(h+2)a_4 + 3(h+2)a_5\}\kappa_3^{(2)}.
\end{aligned}$$

Note that B2 is equivalent to the usual Cramér's condition:

$$\limsup_{\|\boldsymbol{t}\|+\|T_2\|\to\infty} |\mathrm{E}[\exp\{i\boldsymbol{t}'\boldsymbol{\varepsilon} + i\ \mathrm{tr}(\boldsymbol{\varepsilon}'T_2\boldsymbol{\varepsilon})\}]| < 1$$

(see Bhattacharya and Ranga Rao [5], page 207).

Set

$$\Psi = X(X'X)^{-1}H'\{H(X'X)^{-1}H'\}^{-1}H(X'X)^{-1}X',$$

and let ψ_{ab} denote its (a,b)th element. Furthermore, $D_\Psi = \mathrm{diag}(\psi_{11},\ldots,\psi_{nn})$

and $\Psi_{(3)}$ is an $n \times n$ matrix whose (a,b)th element is denoted by ψ_{ab}^3. Then the coefficients a_j's in (2.1) are defined by

$$a_1 = \frac{1}{8}\{n\,\mathrm{tr}(D_\Psi^2) - h(h+2)\}, \qquad a_2 = \frac{n}{12}\mathbf{1}_n'\Psi_{(3)}\mathbf{1}_n,$$

$$a_3 = \frac{n}{8}\mathbf{1}_n' D_\Psi \Psi D_\Psi \mathbf{1}_n, \qquad a_4 = \frac{1}{12}\mathbf{1}_n' D_\Psi \Psi \mathbf{1}_n, \qquad a_5 = \frac{1}{8n}\mathbf{1}_n'\Psi\mathbf{1}_n,$$

where $\mathbf{1}_n$ is an n-dimensional vector all of whose elements are 1. Moreover, $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(1)}$ in (2.1) are the multivariate skewnesses and kurtosis (see Mardia [19] and Isogai [16]) which are defined by

$$\kappa_3^{(1)} = \sum_{abc}^{p} \kappa_{abc}^2, \qquad \kappa_3^{(2)} = \sum_{abc}^{p} \kappa_{aab}\kappa_{bcc}, \qquad \kappa_4^{(1)} = \sum_{ab}^{p} \kappa_{aabb}.$$

Note that the final result depends on the cumulants up to the fourth order. From the result (Hall [14]) on the univariate t-statistic, it is expected that the assumption B1 may be weakened as

B1′. $\mathrm{E}(\|\varepsilon\|^4) < \infty$.

Before concluding this section, we state an alternative expression of (1.4).

COROLLARY 2.2. *Under the same assumptions as in Theorem* 2.1, *the asymptotic expansion* (1.4) *can be written as*

$$\mathrm{P}(T_G \le x) = G_{ph}(x) - \frac{2x}{nph} g_{ph}(x)\left\{ b_1 + b_2 + b_3 + \frac{(b_2+b_3)x}{ph+2} + \frac{b_3 x^2}{(ph+2)(ph+4)} \right\} + \mathrm{o}(n^{-1}), \tag{2.2}$$

where $g_f(x)$ is the density function of a central chi-squared distribution with f degrees of freedom and the coefficients b_j are given by (2.1).

3. Some applications

In this section, we obtain asymptotic expansions of the null distribution for several test statistics by applying Theorem 2.1.

3.1. The multivariate normal case

When each error vector ε_j is independently and identically distributed as $\mathrm{N}_p(\mathbf{0}, \Sigma)$, the multivariate skewnesses and kurtosis are zero, respectively. Therefore, the coefficients b_j's are given by

$$b_0 = \frac{1}{4} ph(h - p - 1 - 2r_1), \qquad b_1 = -\frac{1}{2} ph\{h - r_1 + r_2(h + p + 1)\},$$

$$b_2 = \frac{1}{4} ph(h + p + 1)(1 + 2r_2), \qquad b_3 = 0.$$

These results correspond to the well known formulas (see, e.g., Anderson [1] and Siotani, Hayakawa and Fujikoshi [23]).

3.2. The univariate case

When $p = 1$, T_G corresponds to a perturbation expansion of three test statistics for linear hypothesis in nonnormal univariate linear model. Note that the three tests are essentially the same. In this case, $\kappa_3^{(1)} = \kappa_3^{(2)} = \kappa_3^2$ and $\kappa_4^{(1)} = \kappa_4$. Therefore the coefficients b_j's are given by

$$\begin{aligned}
b_0 &= -\kappa_3^2\{s_2 - hs_3 + h(h-2)s_4\} + \kappa_4 s_1 + \frac{1}{4} h\{(h-2) - 2r_1\}, \\
b_1 &= \kappa_3^2\{3s_2 - (3h+4)s_3 + h(3h+2)s_4\} \\
&\quad - 2\kappa_4 s_1 - \frac{1}{2} h\{h - r_1 + r_2(h+2)\}, \\
b_2 &= -\kappa_3^2\{3s_2 - (3h+8)s_3 + (h+2)(3h+4)s_4\} \\
&\quad + \kappa_4 s_1 + \frac{1}{4} h(h+2)(1+2r_2), \\
b_3 &= \kappa_3^2\{s_2 - (h+4)s_3 + (h+2)(h+4)s_4\},
\end{aligned}$$

where

$$s_1 = a_1, \qquad s_2 = a_2 + a_3, \qquad s_3 = 3a_4, \qquad s_4 = a_5.$$

The coefficients b_j's and s_j's in the case $r_1 = r_2 = 0$ are the same ones as in Yanagihara [24].

3.3. Hotelling's T^2 statistic

If we specify the design matrix as $X = \mathbf{1}_n$ and the constraint matrix as $H = 1$ and $r_1 = 0$, then the Lawley-Hotelling trace criterion becomes to Hotelling's T^2 statistic. Since $\Omega = 1$, we have

$$a_1 = -\frac{1}{4}, \qquad a_2 = \frac{1}{12}, \qquad a_3 = \frac{1}{8}, \qquad a_4 = \frac{1}{12}, \qquad a_5 = \frac{1}{8}.$$

Using these coefficients, we obtain b_j's as

$$b_0 = -\frac{1}{4}\kappa_4^{(1)} + \frac{1}{6}\kappa_3^{(1)} - \frac{1}{4}p^2, \qquad b_1 = \frac{1}{2}\kappa_4^{(1)} - \frac{1}{2}\kappa_3^{(1)} - \frac{1}{2}p,$$

$$b_2 = -\frac{1}{4}\kappa_4^{(1)} - \frac{1}{2}\kappa_3^{(2)} + \frac{1}{4}p(p+2), \qquad b_3 = \frac{1}{3}\kappa_3^{(1)} + \frac{1}{2}\kappa_3^{(2)}.$$

These coefficients b_j's correspond to those in Kano [17] and Fujikoshi [10].

3.4. One-way MANOVA test statistics for equality of means

Set

$$X = \begin{pmatrix} \mathbf{1}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{n_k} \end{pmatrix} \qquad (n \times k \ \text{matrix}), \tag{3.1}$$

$$H = \begin{pmatrix} 1 & \cdots & 0 & -1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -1 \end{pmatrix} \qquad ((k-1) \times k \ \text{matrix}), \tag{3.2}$$

and $\Omega = I_k - \boldsymbol{\rho}\boldsymbol{\rho}'$, where

$$\boldsymbol{\rho} = (\rho_1, \dots, \rho_k)' = \left(\sqrt{\frac{n_1}{n}}, \dots, \sqrt{\frac{n_k}{n}}\right)'.$$

Then T_G becomes a perturbation expansion of one-way MANOVA test statistics for testing an equality of mean vectors of k populations with each sample size n_i $(1 \le i \le k)$. It is easily seen that $\text{rank}(\Omega) = h = k - 1$. Set

$$D' = X(X'X)^{-1/2} = \begin{pmatrix} n_1^{-1/2}\mathbf{1}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & n_2^{-1/2}\mathbf{1}_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & n_k^{-1/2}\mathbf{1}_{n_k} \end{pmatrix},$$

then

$$\Psi = D'\Omega D = \begin{pmatrix} \omega_{11} n_1^{-1}\mathbf{1}_{n_1}\mathbf{1}_{n_1}' & \cdots & \omega_{1k} n_1^{-1/2} n_k^{-1/2}\mathbf{1}_{n_1}\mathbf{1}_{n_k}' \\ \vdots & \ddots & \vdots \\ \omega_{k1} n_k^{-1/2} n_1^{-1/2}\mathbf{1}_{n_k}\mathbf{1}_{n_1}' & \cdots & \omega_{kk} n_k^{-1}\mathbf{1}_{n_k}\mathbf{1}_{n_k}' \end{pmatrix}, \tag{3.3}$$

where ω_{ab} is the (a,b)th element of Ω in (1.2). Furthermore, using $\omega_{ab} = \delta_{ab} - \rho_a\rho_b$, we have

$$a_1 = \frac{1}{8}\left(\sum_{a=1}^{k}\rho_a^{-2} - k^2 - 2k + 2\right),$$

$$a_2 = \frac{1}{12}\left(\sum_{a=1}^{k}\rho_a^{-2} - 3k + 2\right),$$

$$a_3 = \frac{1}{8}\left(\sum_{a=1}^{k}\rho_a^{-2} - k^2\right), \qquad a_4 = a_5 = a_6 = 0.$$

Next, we consider the assumptions A1, A2 and A3. It is easily shown that all $\|\boldsymbol{x}_j\| = 1$, $n^{-1}\sum_{j=1}^{n}\|\boldsymbol{x}_j\|^4 = 1$ and $n/n_j \le n/\lambda_n$. Therefore A1, A2 and A3 are replaced by

$$n/n_j = \mathrm{O}(1) \qquad (j = 1, 2, \dots, k). \tag{3.4}$$

So we have

$$b_0 = a_1\kappa_4^{(1)} - (a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) + \frac{1}{4}p(k-1)(k-p-2-2r_1),$$

$$b_1 = -2a_1\kappa_4^{(1)} + 3(a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) - \frac{1}{2}p(k-1)\{k-1-r_1+r_2(k+p)\},$$

$$b_2 = a_1\kappa_4^{(1)} - 3(a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) + \frac{1}{4}p(k-1)(k+p)(1+2r_2),$$

$$b_3 = a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}.$$

The coefficients b_j coincide with those in Fujikoshi [12].

3.5. One-way MANOVA test statistics for linear hypothesis in k populations

Next, we consider test statistics for linear hypothesis in k populations, that is, the design matrix X is the same as (3.1). Then, Ψ can be written as (3.3). So we can rewrite the coefficients a_j in Theorem 2.1 in simpler forms as

$$a_1 = \frac{1}{8}\left\{\sum_{a=1}^{k}\rho_a^{-2}\omega_{aa}^2 - h(h+2)\right\}, \qquad a_2 = \frac{1}{12}\sum_{ab}^{k}\rho_a^{-1}\rho_b^{-1}\omega_{ab}^3,$$

$$a_3 = \frac{1}{8}\sum_{ab}^{k}\rho_a^{-1}\rho_b^{-1}\omega_{aa}\omega_{ab}\omega_{bb}, \qquad a_4 = \frac{1}{12}\sum_{ab}^{k}\rho_a^{-1}\rho_b\omega_{aa}\omega_{ab}, \qquad a_5 = \frac{1}{8}\sum_{ab}^{k}\rho_a\rho_b\omega_{ab},$$

where $\sum_{a_1\dots a_j}^{p}$ means $\sum_{a_1=1}^{p}\cdots\sum_{a_j=1}^{p}$. Furthermore, by the same reason as in §3.4, A1, A2 and A3 are replaced by (3.4).

3.6. Two-way MANOVA test statistics with interaction

Finally, we consider the model in which we observe independently $\boldsymbol{y}_{ijl}$ with

$$\boldsymbol{y}_{ijl} = \boldsymbol{\eta}_{ij} + \boldsymbol{\varepsilon}_{ijl}, \qquad (i = 1, 2, \dots, r; j = 1, 2, \dots, s; l = 1, 2 \dots, n_{ij}).$$

Here the cell mean is decomposed in the following way.

$$\boldsymbol{\eta}_{ij} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + \boldsymbol{\gamma}_{ij}.$$

Set $n = \sum_{i=1}^{r} \sum_{j=1}^{s} n_{ij}$, $n_{i\cdot} = \sum_{j=1}^{s} n_{ij}$ and $n_{\cdot j} = \sum_{i=1}^{r} n_{ij}$. Suppose that n_{ij} satisfies the proportional sampling in which

$$n_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n}.$$

To define $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_j$ and $\boldsymbol{\gamma}_{ij}$ uniquely, we need to have some constraints

$$\sum_{i=1}^{r} n_{i\cdot} \boldsymbol{\alpha}_i = \mathbf{0}, \qquad \sum_{j=1}^{s} n_{\cdot j} \boldsymbol{\beta}_j = \mathbf{0}, \qquad \sum_{i=1}^{r} n_{i\cdot} \boldsymbol{\gamma}_{ij} = \mathbf{0}, \qquad \sum_{j=1}^{s} n_{\cdot j} \boldsymbol{\gamma}_{ij} = \mathbf{0}.$$

Our hypothesis H_0 is

$$H_0 : \boldsymbol{\gamma}_{ij} = \mathbf{0} \qquad (i = 1, 2, \dots, r; j = 1, 2, \dots, s).$$

Set

$$X = \begin{pmatrix} \mathbf{1}_{n_{11}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_{12}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{n_{rs}} \end{pmatrix}, \qquad (n \times rs \text{ matrix}),$$

$$\Xi = (\boldsymbol{\eta}_{11}, \boldsymbol{\eta}_{12}, \dots, \boldsymbol{\eta}_{rs})', \qquad (rs \times p \text{ matrix}),$$

$$\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_{rs})' = \left(\sqrt{\frac{n_{11}}{n}}, \sqrt{\frac{n_{12}}{n}}, \dots, \sqrt{\frac{n_{rs}}{n}}\right)', \qquad (rs \times 1 \text{ vector}),$$

$$G_A = \begin{pmatrix} G_{11}^{(A)} & \cdots & G_{1r}^{(A)} \\ \vdots & \ddots & \vdots \\ G_{r1}^{(A)} & \cdots & G_{rr}^{(A)} \end{pmatrix}, \qquad (rs \times rs \text{ matrix}),$$

$$G_B = \begin{pmatrix} G_{11}^{(B)} & \cdots & G_{1r}^{(B)} \\ \vdots & \ddots & \vdots \\ G_{r1}^{(B)} & \cdots & G_{rr}^{(B)} \end{pmatrix}, \qquad (rs \times rs \text{ matrix}),$$

$$G_C = I_{rs} - G_A - G_B + \boldsymbol{\rho}\boldsymbol{\rho}',$$

where

$$G_{ii'}^{(A)} = \frac{\delta_{ii'}}{\sqrt{n_{i\cdot}}\sqrt{n_{i'\cdot}}}\begin{pmatrix} \sqrt{n_{i1}}\sqrt{n_{i'1}} & \cdots & \sqrt{n_{i1}}\sqrt{n_{i's}} \\ \vdots & \ddots & \vdots \\ \sqrt{n_{is}}\sqrt{n_{i'1}} & \cdots & \sqrt{n_{is}}\sqrt{n_{i's}} \end{pmatrix},$$

$$G_{ii'}^{(B)} = \mathrm{diag}\left(\frac{\sqrt{n_{i1}}\sqrt{n_{i'1}}}{n_{\cdot 1}}, \ldots, \frac{\sqrt{n_{is}}\sqrt{n_{i's}}}{n_{\cdot s}}\right).$$

Note that $G_C^2 = G_C$ in the case of proportional sampling. Then, by using $(X'X)^{-1/2}G_C(X'X)^{1/2}$, an unknown parameter matrix Ξ can be transformed into $(\gamma_{11}, \gamma_{12}, \ldots, \gamma_{rs})'$. So, our hypothesis H_0 can be rewritten as

$$H_0 : (X'X)^{1/2}G_C(X'X)^{1/2}\Xi = 0.$$

Since G_C is a projection matrix, there exists an $(r-1)(s-1) \times rs$ matrix L such that

$$L'L = G_C, \qquad LL' = I_{(r-1)(s-1)}.$$

Then an $(r-1)(s-1) \times rs$ restricted matrix H for hypothesis H_0 can be defined by

$$H = L(X'X)^{-1/2}.$$

Therefore, our theorem can be applied for

$$\Omega = G_C = I_{rs} - G_A - G_B + \boldsymbol{\rho}\boldsymbol{\rho}'.$$

Moreover, in order to simplify the coefficients a_j's, we define the indicator matrix C by

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1s} \\ c_{21} & c_{22} & \cdots & c_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ c_{r1} & c_{r2} & \cdots & c_{rs} \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & s \\ s+1 & s+2 & \cdots & 2s \\ \vdots & \vdots & \ddots & \vdots \\ (r-1)s+1 & (r-1)s+2 & \cdots & rs \end{pmatrix}.$$

Then the (a, b)th element of Ω can be rewritten as

$$\omega_{ab} = \omega_{c_{ij}c_{i'j'}} = \delta_{ii'}\delta_{jj'} - \delta_{ii'}\frac{\sqrt{n_{ij}}\sqrt{n_{i'j'}}}{\sqrt{n_{i\cdot}}\sqrt{n_{i'\cdot}}} - \delta_{jj'}\frac{\sqrt{n_{ij}}\sqrt{n_{i'j'}}}{\sqrt{n_{\cdot j}}\sqrt{n_{\cdot j'}}} + \frac{\sqrt{n_{ij}}\sqrt{n_{i'j'}}}{n}.$$

As $\Psi = D'\Omega D$, where

$$D' = X(X'X)^{-1/2} = \begin{pmatrix} n_{11}^{-1/2}\mathbf{1}_{n_{11}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & n_{12}^{-1/2}\mathbf{1}_{n_{12}} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & n_{rs}^{-1/2}\mathbf{1}_{n_{rs}} \end{pmatrix},$$

we have

$$a_1 = \frac{1}{8}\left[\sum_{a=1}^{rs}\rho_a^{-2}\omega_{aa}^2 - (s-1)(r-1)\{(s-1)(r-1)+2\}\right]$$

$$= \frac{1}{8}\left[\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n}{n_{ij}}(\omega_{c_{ij}c_{ij}})^2 - (s-1)(r-1)\{(s-1)(r-1)+2\}\right]$$

$$= \frac{1}{8}\left\{\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n}{n_{ij}} + (1-2s)\sum_{i=1}^{r}\frac{n}{n_{i\cdot}} + (1-2r)\sum_{j=1}^{s}\frac{n}{n_{\cdot j}} - (1-2rs) - (r-1)^2(s-1)^2\right\},$$

$$a_2 = \frac{1}{12}\sum_{ab}^{rs}\rho_a^{-1}\rho_b^{-1}\omega_{ab}^3 = \frac{1}{12}\sum_{ii'}^{r}\sum_{jj'}^{s}\frac{n}{\sqrt{n_{ij}}\sqrt{n_{i'j'}}}(\omega_{c_{ij}c_{i'j'}})^3$$

$$= \frac{1}{12}\left\{\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n}{n_{ij}} - (3s-2)\sum_{i=1}^{r}\frac{n}{n_{i\cdot}} - (3r-2)\sum_{j=1}^{s}\frac{n}{n_{\cdot j}} + 9rs - 6r - 6s + 4\right\},$$

$$a_3 = \frac{1}{8}\sum_{ab}^{rs}\rho_a^{-1}\rho_b^{-1}\omega_{aa}\omega_{ab}\omega_{bb} = \frac{1}{8}\sum_{ii'}^{r}\sum_{jj'}^{s}\frac{n}{\sqrt{n_{ij}}\sqrt{n_{i'j'}}}\omega_{c_{ij}c_{ij}}\omega_{c_{ij}c_{i'j'}}\omega_{c_{i'j'}c_{i'j'}}$$

$$= \frac{1}{8}\left\{\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n}{n_{ij}} - s^2\sum_{i=1}^{r}\frac{n}{n_{i\cdot}} - r^2\sum_{j=1}^{s}\frac{n}{n_{\cdot j}} + r^2s^2\right\},$$

and $a_4 = a_5 = a_6 = 0$. Therefore, the coefficients b_j's become

$$b_0 = a_1\kappa_4^{(1)} - (a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) + \frac{1}{4}p(r-1)(s-1)\{(r-1)(s-1) - p - 1 - 2r_1\},$$

$$b_1 = -2a_1\kappa_4^{(1)} + 3(a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) - \frac{1}{2}p(r-1)(s-1)[(r-1)(s-1) - r_1 + r_2\{(r-1)(s-1) + p + 1\}],$$

$$b_2 = a_1\kappa_4^{(1)} - 3(a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}) + \frac{1}{4}p(r-1)(s-1)\{(r-1)(s-1) + p + 1\}(1 + 2r_2),$$

$$b_3 = a_2\kappa_3^{(1)} + a_3\kappa_3^{(2)}.$$

Furthermore, as in §3.4 and 3.5, A1, A2 and A3 are replaced by

$$n/n_{ij} = \mathrm{O}(1) \qquad (i = 1, 2, \dots, r;\ j = 1, 2, \dots, s).$$

For a general case, i.e., non-proportional sampling, Fujikoshi [9] studied test statistics in the two-way AMOVA model. It may be necessary to devise an application of our formula to this general case.

4. Numerical accuracies

In this section, numerical accuracies are studied for the actual test sizes of some multivariate tests under four distributions considered by Everitt [8].

First, Hotelling's T^2 statistic, which is denoted by T_G, is taken up. Some effects of T_G to nonnormality have been pointed out by Chase and Bulgren [7] and Everitt [8], based on Monte Carlo experiment. Our purpose is to see how close the actual test size is to the nominal one by using the asymptotic expansion approximations.

Generally, the Cornish-Fisher expansion is used as an approximation to the true percentage point. Let $t(u)$ and u denote the true percentage point and the percentage point of limiting distribution of T_G respectively, that is

$$\mathrm{P}(T_G \le t(u)) = \mathrm{P}(\chi^2_{ph} \le u),$$

where χ^2_{ph} is a variate of a central chi-squared distribution with degrees of freedom ph. Then from (2.2), $t(u)$ can be expanded as

$$\begin{aligned} t(u) &= u + \frac{2u}{nph}\left\{ b_1 + b_2 + b_3 + \frac{(b_2 + b_3)u}{ph + 2} + \frac{b_3 u^2}{(ph+2)(ph+4)} \right\} + \mathrm{o}(n^{-1}) \\ &= t_{\mathrm{E}}(u) + \mathrm{o}(n^{-1}). \end{aligned} \tag{4.1}$$

In actual use, we use $\hat{t}_{\mathrm{E}}(u)$, which is defined from $t_{\mathrm{E}}(u)$ by replacing the unknown parameters $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(2)}$ by their estimators, respectively. Set

$$\tilde{\boldsymbol{y}}_j = \hat{\Sigma}^{-1/2}(\boldsymbol{y}_j - \bar{\boldsymbol{y}}),$$

where

$$\bar{\boldsymbol{y}} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{y}_j, \qquad \hat{\Sigma} = \frac{1}{n-1}\sum_{j=1}^{n} (\boldsymbol{y}_j - \bar{\boldsymbol{y}})(\boldsymbol{y}_j - \bar{\boldsymbol{y}})',$$

then the unknown parameters $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(1)}$ can be estimated as

$$\hat{\kappa}_3^{(1)} = \left(\frac{n}{(n-1)(n-2)} \right)^2 \sum_{ij}^{n} (\tilde{\boldsymbol{y}}_i' \tilde{\boldsymbol{y}}_j)^3,$$

$$\hat{\kappa}_3^{(2)} = \left(\frac{n}{(n-1)(n-2)}\right)^2 \sum_{ij}^{n} (\tilde{\boldsymbol{y}}_i'\tilde{\boldsymbol{y}}_i)(\tilde{\boldsymbol{y}}_i'\tilde{\boldsymbol{y}}_j)(\tilde{\boldsymbol{y}}_j'\tilde{\boldsymbol{y}}_j),$$

$$\hat{\kappa}_4^{(1)} = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{j=1}^{n} (\tilde{\boldsymbol{y}}_j'\tilde{\boldsymbol{y}}_j)^2 - p(p+2).$$

For estimations of the multivariate skewnesses and kurtosis, see, e.g., Kaplan [18], Mardia [19] and Isogai [16].

On the other hand, in the case of Hotelling's T^2 statistic, we can use a modified Cornish-Fisher expansion, which gives an exact percentage point in the normal error case. Such a modification is obtained by using the result that $(n-p)T_G/p(n-1)$ is distributed as F-distribution with degrees of freedom p and $n-p$ under normality. Then, we can modify $t_{\mathrm{E}}(u)$ as

$$\begin{aligned} t(u) &= \frac{p(n-1)}{n-p} u_F - \frac{2u}{np} \left\{ b_0' - \frac{(b_2' + b_3')u}{p+2} - \frac{b_3' u^2}{(p+2)(p+4)} \right\} + \mathrm{o}(n^{-1}) \\ &= t_{\mathrm{E}}(u) + \mathrm{o}(n^{-1}), \end{aligned}$$

where u_F is the percentage point of F-distribution with degrees of freedom p and $n-p$ and

$$b_0' = -\frac{1}{4}\kappa_4^{(1)} + \frac{1}{6}\kappa_3^{(1)}, \qquad b_1' = \frac{1}{2}\kappa_4^{(1)} - \frac{1}{2}\kappa_3^{(1)},$$

$$b_2' = -\frac{1}{4}\kappa_4^{(1)} - \frac{1}{2}\kappa_3^{(2)}, \qquad b_3' = \frac{1}{3}\kappa_3^{(1)} + \frac{1}{2}\kappa_3^{(2)}.$$

If ε is distributed as a normal distribution, then its expansion gives an exact percentage point.

The error distributions considered are the same ones as in Everitt [8], i.e.,

1. *Multivariate Normal Distribution*,
2. *Uniform Distribution*: Each of the p variables is generated independently from a uniform (0,1) distribution,
3. *Exponential Distribution*: Each of the p variables is generated independently from an exponential distribution with a mean of unity,
4. *Lognormal Distribution*: Each of the p variables is generated independently from a lognormal distribution such that $\log x \sim \mathrm{N}(0,1)$.

Table 4.1 gives the actual test sizes for the nominal 10%, 5% and 1% test in several cases of p and n. For each cell in Table 4.1, the top figure expresses the actual test sizes based on the percentage point of F-distribution, the middle and bottom figures show the actual sizes by using $t_{\mathrm{E}}(u)$ and $\hat{t}_{\mathrm{E}}(u)$, respectively. From Table 4.1, it seems that $t_{\mathrm{E}}(u)$ gives a considerable improvement for the

TABLE 4.1: Actual test sizes of Hotelling's T^2 test.

		Normal Nominal Sizes			Uniform Nominal Sizes			Exponential Nominal Sizes			Log-Normal Nominal Sizes		
n	p	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
5	2	10.0	5.1	0.8	11.8	6.4	1.8	20.9	14.5	4.9	30.2	21.5	7.8
		10.0	5.1	0.8	11.7	6.2	1.8	16.3	10.0	3.3	8.7	4.7	1.1
		8.8	4.1	0.7	10.2	5.6	1.6	18.7	12.4	4.2	27.3	19.1	6.8
10	2	10.1	5.0	0.9	10.1	6.0	1.8	17.9	12.4	5.4	28.1	21.9	11.8
		10.1	5.0	0.9	10.0	5.8	1.6	12.0	6.6	1.6	5.4	2.1	0.3
		8.6	4.0	0.6	8.9	4.8	1.2	15.1	9.3	3.5	23.8	16.4	7.0
	4	10.2	5.3	1.3	11.1	5.8	1.6	22.1	14.8	5.9	37.2	27.5	13.6
		10.2	5.3	1.3	10.9	5.7	1.6	15.2	9.1	2.3	7.7	3.2	0.5
		7.6	3.7	0.8	8.7	4.2	1.2	17.3	10.9	3.4	29.9	20.9	8.5
	6	9.9	5.1	1.0	10.8	5.4	1.1	21.8	12.9	3.7	39.3	27.0	10.2
		9.9	5.1	1.0	10.7	5.3	1.1	17.6	10.2	2.7	14.7	8.5	2.3
		8.0	4.0	0.8	8.4	4.2	0.9	18.1	10.5	2.8	34.0	22.9	8.1
15	2	10.1	5.3	0.8	10.8	5.5	1.1	17.1	11.7	5.5	26.4	20.7	11.9
		10.1	5.3	0.8	10.7	5.4	1.1	11.8	6.7	1.1	5.4	2.0	0.2
		9.5	4.7	0.4	9.5	4.5	0.9	13.8	8.9	2.8	22.4	14.8	6.3
	4	9.1	4.5	0.7	11.6	6.2	1.4	20.8	14.8	5.5	36.3	27.8	14.3
		9.1	4.5	0.7	11.3	5.9	1.3	13.7	6.9	1.5	5.3	2.0	0.1
		6.9	2.6	0.3	9.4	4.0	0.6	15.9	9.0	2.3	27.2	17.9	7.3
	6	11.1	5.5	1.1	9.0	4.2	0.9	20.2	13.3	4.9	41.3	31.6	16.3
		11.1	5.5	1.1	8.9	4.2	0.9	13.5	7.4	2.5	7.0	3.5	0.5
		7.6	3.5	0.4	6.5	2.9	0.4	14.2	7.9	2.8	31.3	22.0	8.1
	8	10.1	5.1	1.1	10.3	5.5	1.1	21.7	13.0	3.8	42.4	32.0	14.1
		10.1	5.1	1.1	10.2	5.5	1.1	16.0	8.6	2.4	10.4	5.4	1.1
		6.4	3.0	0.8	7.7	3.4	0.5	15.6	8.5	2.3	34.7	23.2	8.1

(cont'd on p. 32)

actual test size. However, there is a tendency that the approximation tends to be bad as p tends to be large. Moreover, though the estimation problem for $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(1)}$ is left over, as these cumulants tend to be large, it becomes difficult to obtain good estimators even when the sample size is not so small as that.

Next, we compare the asymptotic expansion method with other ones. Bootstrap is one of the powerful methods when error's distribution is general. To construct the bootstrap approximations, let $\chi^* = \{\boldsymbol{y}_1^*, \ldots, \boldsymbol{y}_n^*\}$ denote a resample drawn randomly, with replacement, from $\chi = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$. Then the

Table 4.1: (Continued)

n	p	Normal Nominal Sizes 10%	5%	1%	Uniform Nominal Sizes 10%	5%	1%	Exponential Nominal Sizes 10%	5%	1%	Log-Normal Nominal Sizes 10%	5%	1%
20	2	11.0	5.7	0.9	10.4	5.6	1.1	15.0	10.0	4.2	26.6	20.1	11.4
		11.0	5.7	0.9	10.4	5.3	1.0	10.8	5.7	1.0	4.8	2.0	0.1
		10.2	5.0	0.8	9.8	4.9	0.8	12.4	6.8	2.0	21.1	13.7	4.8
	4	9.0	4.2	0.7	9.6	5.0	1.1	19.1	12.0	4.7	33.0	25.4	14.1
		9.0	4.2	0.7	9.4	4.8	1.0	11.2	5.8	1.0	4.3	1.5	0.1
		7.3	3.1	0.3	8.5	3.9	0.6	12.9	7.7	1.8	23.4	15.2	5.8
	6	10.1	5.4	1.5	10.9	5.5	1.2	20.8	13.7	5.5	38.1	29.0	15.9
		10.1	5.4	1.5	10.7	5.3	1.2	12.8	7.0	2.0	5.0	2.1	0.5
		7.6	3.5	0.4	8.2	3.5	0.5	14.1	7.6	2.5	26.6	17.5	6.5
	8	10.5	5.3	0.9	11.6	5.3	1.0	24.1	14.3	5.4	40.6	29.6	14.4
		10.5	5.3	0.9	11.4	5.3	0.9	14.7	8.5	1.9	5.3	2.8	0.3
		6.6	2.6	0.2	7.4	3.1	0.3	15.2	8.7	2.1	27.2	17.8	6.2
	10	10.4	4.8	1.1	8.2	4.7	0.5	22.9	14.4	4.8	42.4	31.7	15.2
		10.4	4.8	1.1	8.1	4.6	0.5	16.2	9.7	2.5	7.8	3.5	0.4
		6.0	2.9	0.5	5.7	2.6	0.2	15.5	9.0	2.4	31.3	21.1	7.1

bootstrap version of Hotelling's T^2 test statistic is defined by

$$T_G^* = n(\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}})' S^{*-1} (\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}}),$$

where

$$\bar{\boldsymbol{y}}^* = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{y}_j^*, \qquad S^* = \frac{1}{n-1}\sum_{j=1}^{n} (\boldsymbol{y}_j^* - \bar{\boldsymbol{y}}^*)(\boldsymbol{y}_j^* - \bar{\boldsymbol{y}}^*)'.$$

The percentage point of bootstrap version t_{B} can be calculated as

$$\mathrm{P}(T_G^* \geq t_{\mathrm{B}} \,|\, \chi) = \alpha.$$

Furthermore, Mardia [20] proposed a robust method on Hotelling's T^2 test statistic. Let

$$\delta = 1 + \frac{1}{n}\left\{\frac{b_{2,p} - p(p+2)}{p}\right\}, \qquad b_{2,p} = \frac{1}{n}\sum_{j=1}^{n} (\tilde{\boldsymbol{y}}_j' \tilde{\boldsymbol{y}}_j)^2,$$

then we use u_M, a percentage point of F distribution with δp and $\delta(n-p)$ degrees of freedom, as an approximation of one. For the percentage points of F distribution with non-integer values of degrees of freedom, see, Mardia and Zemroch [21].

Table 4.2: Actual test sizes of Hotelling's T^2 test; several methods.

n	p		Normal Nominal Sizes 10%	5%	1%	Uniform Nominal Sizes 10%	5%	1%	Exponential Nominal Sizes 10%	5%	1%	Log-Normal Nominal Sizes 10%	5%	1%
10	2	α_1	9.81	4.96	0.89	10.63	5.53	1.48	18.62	12.75	6.30	14.02	5.67	0.90
		α_2	9.81	4.96	0.89	10.55	5.38	1.33	12.34	7.35	1.95	0.20	0.05	0.00
		α_3	8.35	3.84	0.63	9.16	4.57	0.99	15.35	10.02	4.01	8.39	3.27	0.37
		α_4	6.35	2.23	0.27	6.56	2.52	0.25	10.35	5.80	1.13	3.89	1.11	0.09
		α_5	11.08	6.03	1.42	11.47	6.51	1.96	20.37	14.88	7.74	18.25	8.58	1.76
	3	α_1	9.82	4.92	1.00	10.69	5.52	1.25	21.16	14.64	6.33	13.98	6.10	1.21
		α_2	9.82	4.92	1.00	10.55	5.42	1.21	14.09	8.41	2.20	0.25	0.08	0.00
		α_3	7.62	3.55	0.59	8.78	4.22	0.78	16.98	10.91	3.93	8.30	3.11	0.41
		α_4	3.36	0.83	0.02	3.55	0.91	0.03	7.10	2.68	0.12	1.43	0.37	0.00
		α_5	11.65	6.38	1.68	12.29	6.96	2.00	23.93	17.46	9.13	19.41	9.89	2.37
	4	α_1	10.02	4.92	1.04	10.82	5.56	1.25	22.44	15.23	5.84	14.13	6.84	1.40
		α_2	10.02	4.92	1.04	10.69	5.48	1.19	15.55	9.15	2.52	0.52	0.21	0.01
		α_3	7.21	3.38	0.56	8.35	3.80	0.76	17.78	10.95	3.48	8.70	3.72	0.65
		α_4	0.71	0.08	0.00	0.76	0.07	0.00	2.15	0.27	0.00	0.36	0.03	0.00
		α_5	12.74	7.11	1.84	12.99	7.85	2.16	26.05	18.97	9.31	19.91	10.96	3.10

Table 4.2 shows the actual sizes for the nominal 10%, 5% and 1% tests in the case of $n = 10$ and $p = 2$, 3 and 4. Each test size α_j is defined by

$$\alpha_1 = \mathrm{P}(T_G \geq p(n-1)u_F/(n-p)), \qquad \alpha_2 = \mathrm{P}(T_G \geq t_{\mathrm{E}}(u)),$$

$$\alpha_3 = \mathrm{P}(T_G \geq \hat{t}_{\mathrm{E}}(u)), \qquad \alpha_4 = \mathrm{P}(T_G \geq t_{\mathrm{B}}),$$

$$\alpha_5 = \mathrm{P}(T_G \geq p(n-1)u_M/(n-p)).$$

Four error distributions considered are the same ones as in the previous simulation. From Table 4.2, we can see that the bootstrap method gives conservative approximations but the approximations are not so good. Especially, for $p = 3$ and 4, the bootstrap approximation is very bad since the determinant of S^* is near 0 occasionally when p is large in comparison with n. On the other hand, the asymptotic expansion with estimators improves the first order approximations for actual test sizes constantly. However, these improvements are not enough, in comparison with the case of normality. Mardia's method is robust in the non-skewness data only.

5. Bartlett corrections

In this section, first we consider the situation where the Bartlett corrections do work even under nonnormality. More precisely, we shall find conditions such that the Bartlett correction in the normal case implies an improved χ^2-approximation, even under nonnormality. Note that T_G has been adjusted by the Bartlett correction in the normal case, namely, under normality, $\mathrm{E}(T_G) = ph + \mathrm{o}(n^{-1})$.

By using the formula in Theorem 2.1, the expectation of T_G can be calculated as

$$\begin{aligned}\mathrm{E}(T_G) &= ph + \frac{1}{n}\sum_{j=0}^{3} b_j(ph+2j) + \mathrm{o}(n^{-1}) \\ &= ph\left(1+\frac{c_1}{n}\right) + \mathrm{o}(n^{-1}).\end{aligned}$$

Noting that $\sum_{j=0}^{3} b_j = 0$, we obtain

$$c_1 = \frac{2}{ph}\sum_{j=1}^{3} jb_j.$$

From (2.1),

$$b_1 + 2b_2 + 3b_3 = 4a_5(\kappa_3^{(1)} + \kappa_3^{(2)}). \tag{5.1}$$

Therefore

$$c_1 = \frac{8}{ph}a_5(\kappa_3^{(1)} + \kappa_3^{(2)}).$$

If $a_5 = 0$, then $\mathrm{E}(T_G) = ph + \mathrm{o}(n^{-1})$. This means that if $a_5 = 0$, then T_G has an improved χ^2 approximation by the Bartlett correction in the normal case. On the other hand,

$$a_5 = \frac{1}{8n}\mathbf{1}_n' X(X'X)^{-1/2}\Omega(X'X)^{-1/2}X'\mathbf{1}_n.$$

So, $a_5 = 0$ is equivalent to

$$\Omega(X'X)^{-1/2}X'\mathbf{1}_n = \mathbf{0}. \tag{5.2}$$

If H is given by a concrete form, then the condition (5.2) may be changed into a simpler form as

$$H(X'X)^{-1}X'\mathbf{1}_n = \mathbf{0}. \tag{5.3}$$

Moreover,

$$a_4 = \frac{1}{12}\mathbf{1}_n' D_\Psi X(X'X)^{-1/2}\Omega(X'X)^{-1/2}X'\mathbf{1}_n.$$

Therefore, under condition (5.2), the coefficient a_4 becomes 0. This result can be summarized as the following Theorem 5.1.

Theorem 5.1. *Suppose that X and H satisfy the condition* (5.2) *(or more concretely* (5.3)*), then the test statistics adjusted by the Bartlett correction in the normal case have an improved χ^2-approximation, i.e., even under nonnormality,* $\mathrm{E}(T_G) = ph + \mathrm{o}(n^{-1})$.

Related to Theorem 5.1, we examine the condition (5.2) in the one-way MANOVA model. In this model, the design matrix X is defined by (3.1). Then the condition (5.2) can be rewritten as $H\mathbf{1}_k = \mathbf{0}$, which means that the rows of H are contrast vectors. Therefore, the test statistics for equality of means in the one-way MANOVA model can be improved by the Bartlett correction in the normal case.

Next, we consider the second moment of T_G, which can be calculated as

$$\begin{aligned}\mathrm{E}(T_G^2) &= ph(ph+2) + \frac{1}{n}\sum_{j=0}^{3} b_j(ph+2j)(ph+2+2j) + \mathrm{o}(n^{-1}) \\ &= ph(ph+2)\left(1+\frac{c_2}{n}\right) + \mathrm{o}(n^{-1}).\end{aligned}$$

Note that $\sum_{j=0}^{3} b_j = 0$, and we obtain

$$\begin{aligned}c_2 &= \frac{4}{ph(ph+2)}\sum_{j=1}^{3} j(ph+j+1)b_j \\ &= \frac{4}{ph(ph+2)}\{(ph+2)(b_1+2b_2+3b_3)+2(b_2+3b_3)\}.\end{aligned}$$

From (2.1),

$$\begin{aligned}b_2+3b_3 &= a_1\kappa_4^{(1)} - \{6a_4 - 4(h+2)a_5\}(\kappa_3^{(1)}+\kappa_3^{(2)}) \\ &\quad + \frac{1}{4}ph(h+p+1)(1+2r_2). \qquad (5.4)\end{aligned}$$

Substituting (5.1) and (5.4) into c_2 yields

$$\begin{aligned}c_2 = \frac{8}{ph(ph+2)}\Big[&a_1\kappa_4^{(1)} - \{6a_4 - 2(ph+2h+6)a_5\}(\kappa_3^{(1)}+\kappa_3^{(2)}) \\ &+ \frac{1}{4}ph(h+p+1)(1+2r_2)\Big].\end{aligned}$$

If the condition (5.2) is satisfied, then the coefficients a_4 and a_5 become 0. Then c_2 has a simpler form as

$$c_2 = \frac{8}{ph(ph+2)}\left\{a_1\kappa_4^{(1)} + \frac{1}{4}ph(h+p+1)(1+2r_2)\right\}.$$

Furthermore, if

$$a_1 = 0, \tag{5.5}$$

then c_2 does not depend on unknown parameters, i.e., $c_2 = 2(h+p+1)\cdot(1+2r_2)/(ph+2)$. So, the n^{-1} term of variance of T_G is independent of unknown parameters $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(1)}$, in fact

$$\mathrm{Var}(T_G) = 2ph\left\{1 + \frac{1}{n}(p+h+1)(1+2r_2)\right\} + \mathrm{o}(n^{-1}).$$

Under these conditions, a modified Bartlett transformation (see, Fujikoshi [11]) can be defined by

$$\tilde{T}_G = (ph+2)\left\{\frac{n}{(p+h+1)(1+2r_2)} + \frac{1}{2}\right\} \log\left\{1 + \frac{(p+h+1)(1+2r_2)}{n(ph+2)} T_G\right\}.$$

Furthermore, the mean and variance of $\tilde{T}_G$ become

$$\mathrm{E}(\tilde{T}_G) = ph + \mathrm{o}(n^{-1}), \qquad \mathrm{Var}(\tilde{T}_G^2) = 2ph + \mathrm{o}(n^{-1}).$$

These results can be summarized as Theorem 5.2.

THEOREM 5.2. *Suppose that X and H satisfy the condition* (5.2) *(or more concretely* (5.3)*) and* (5.5), *then the Lawley-Hotelling and Bartlett-Nanda-Pillai trace criteria can be improved in the variances as well as the means by the modified Bartlett transformation in the normal case, which are defined by*

$$\tilde{T}_{HL} = (ph+2)\left\{\frac{n}{(p+h+1)} + \frac{1}{2}\right\} \log\left\{1 + \frac{(p+h+1)}{n(ph+2)} T_{HL}\right\},$$

$$\tilde{T}_{BNP} = -(ph+2)\left\{\frac{n}{(p+h+1)} - \frac{1}{2}\right\} \log\left\{1 - \frac{(p+h+1)}{n(ph+2)} T_{BNP},\right\},$$

i.e., even under nonnormality,

$$\mathrm{E}(\tilde{T}_{HL}) = ph + \mathrm{o}(n^{-1}), \qquad \mathrm{Var}(\tilde{T}_{HL}) = 2ph + \mathrm{o}(n^{-1}),$$

$$\mathrm{E}(\tilde{T}_{BNP}) = ph + \mathrm{o}(n^{-1}), \qquad \mathrm{Var}(\tilde{T}_{BNP}) = 2ph + \mathrm{o}(n^{-1}).$$

For T_{LR}, it may be noted that under conditions (5.2) and (5.5), $\mathrm{Var}(T_{LR}) = 2ph + \mathrm{o}(n^{-1})$, so it shall not be necessary to consider the transformation such as T_{HL} and T_{BNP}.

Furthermore, we consider the two-way MANOVA test statistics as in § 3.6. In this case, it can be checked that condition (5.2) holds through a simple calculation. More specifically, we consider the case $r = s = 3$ and $n_{i1} = n_{i2} = n_{i3}$ $(i = 1, 2, 3)$. In this case, if $n_{1j} : n_{2j} : n_{3j} = 1 : 2 : 3$ $(j = 1, 2, 3)$, then the condition (5.5) holds. Therefore, when $r = s = 3$ and $n_{1j} : n_{2j} : n_{3j} = 1 : 2 : 3$,

TABLE 5.1: Actual test sizes of test statistics in two-way MANOVA model with interaction; Bartlett corrections.

			Normal			Uniform			Exponential			Log-Normal		
			Nominal Sizes			Nominal Sizes			Nominal Sizes			Nominal Sizes		
	a_1		10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
T_{LR}	−1.00	T_O	10.0	5.0	0.9	9.6	4.9	1.1	9.2	4.7	0.9	7.9	3.7	0.5
		T_G	10.3	5.2	1.0	9.9	5.1	1.1	9.5	4.9	1.0	8.1	3.8	0.6
	0.00	T_O	9.6	4.9	0.9	10.2	4.9	1.0	10.0	5.1	0.9	9.6	4.9	1.2
		T_G	9.8	5.0	0.9	10.4	5.1	1.1	10.2	5.2	0.9	9.8	5.0	1.3
	1.25	T_O	9.6	4.7	0.9	9.6	4.9	0.9	10.7	5.6	1.4	12.0	6.6	1.9
		T_G	9.9	5.0	0.9	9.8	5.1	0.9	11.0	5.8	1.5	12.2	6.8	2.0
T_{HL}	−1.00	T_O	12.5	7.0	1.7	11.9	6.6	1.8	11.4	6.5	1.8	10.1	4.9	1.0
		T_G	10.8	5.7	1.3	10.5	5.7	1.6	10.1	5.4	1.4	8.6	4.2	0.8
		$\tilde{T}_G$	10.4	5.3	1.0	10.0	5.1	1.1	9.5	4.9	1.0	8.2	3.8	0.6
	0.00	T_O	12.0	6.6	1.6	12.6	6.7	1.8	12.2	6.9	1.7	12.0	6.6	1.9
		T_G	10.5	5.6	1.3	11.1	5.7	1.5	10.6	5.8	1.3	10.4	5.6	1.6
		$\tilde{T}_G$	10.1	5.1	1.0	10.4	5.2	1.1	10.3	5.3	1.0	9.8	5.1	1.3
	1.25	T_O	12.4	6.5	1.6	12.0	6.7	1.6	13.1	7.5	2.4	14.4	8.6	2.9
		T_G	10.6	5.6	1.3	10.4	5.8	1.2	11.6	6.5	2.0	12.8	7.6	2.4
		$\tilde{T}_G$	10.0	5.0	1.0	10.0	5.2	1.0	11.1	5.8	1.5	12.4	6.9	2.1
T_{BNP}	−1.00	T_O	11.3	5.6	1.0	7.3	3.4	0.5	7.2	3.1	0.5	5.8	2.3	0.2
		T_G	9.5	4.6	0.7	9.2	4.5	0.8	8.9	4.2	0.6	8.6	4.2	0.8
		$\tilde{T}_G$	10.2	5.2	1.0	9.9	5.1	1.1	9.4	4.8	1.0	8.1	3.8	0.6
	0.00	T_O	10.9	5.4	0.9	7.7	3.4	0.6	7.6	3.5	0.4	7.3	3.4	0.8
		T_G	9.2	4.5	0.7	9.8	4.5	0.8	9.6	4.7	0.7	10.4	5.6	1.6
		$\tilde{T}_G$	9.8	5.0	0.9	10.4	5.0	1.0	10.1	5.2	0.9	9.8	5.0	1.3
	1.25	T_O	11.1	5.3	0.9	7.6	3.2	0.5	8.3	4.1	0.7	9.4	4.7	1.2
		T_G	9.3	4.1	0.6	9.3	4.5	0.7	10.4	5.2	1.1	12.8	7.6	2.4
		$\tilde{T}_G$	9.9	4.9	0.9	9.9	5.1	0.9	11.0	5.8	1.4	12.2	6.7	1.9

the Lawley-Hotelling and Bartlett-Nanda-Pillai trace criteria in the two-way MANOVA model can be improved by the modified Bartlett transformation in the normal case.

Table 5.1 shows the actual test size of the three test statistics in two-way MANOVA model with iteration in the case $r = s = 3$ and $n_{i1} = n_{i2} = n_{i3}$ $(i = 1, 2, 3)$. For each test statistic, let T_O denote the test statistics without the Bartlett correction, T_G the statistics with the Bartlett correction and $\tilde{T}_G$ the transformed test statistics based on a modified Bartlett transformation.

Needless to say, each correction is obtained under normality and we do not deal with $\tilde{T}_G$ for the likelihood ratio statistic. In Table 5.1, values of a_1 are given by

$$a_1 = -1.00, \qquad (n_{1j} = 10, n_{2j} = 10, n_{3j} = 10),$$

$$a_1 = 0.00, \qquad (n_{1j} = 5, n_{2j} = 10, n_{3j} = 15),$$

$$a_1 = 1.25, \qquad (n_{1j} = 5, n_{2j} = 5, n_{3j} = 20).$$

Four error distributions considered are the same ones as in the previous section. From Table 5.1, we note that the Bartlett correction in the normal case can improve the approximation well enough, even if an error vector is not distributed as a normal distribution. Moreover, when $a_1 = 0$, the difference between the actual test size and the nominal ones is the smallest in all the cases, since the n^{-1} terms of mean and variance of T_G are independent of $\kappa_3^{(1)}$, $\kappa_3^{(2)}$ and $\kappa_4^{(1)}$. In this case, $\tilde{T}_G$ gives better size than T_O and T_G. Furthermore, $\tilde{T}_G$ under the condition (5.5) gives the best size in all the cases. Therefore, we can say that test statistics $\tilde{T}_G$, which satisfies the conditions (5.2) and (5.5), is robust for nonnormality. So, when we consider a test statistic which satisfies the condition (5.2), we recommend to get the sample data satisfying the condition (5.5) and transforming the test statistic by the modified Bartlett transformation.

Appendix

A.1. Some basic results on validity

The aim of this section is to prepare some basic theorems in order to assure the validity of asymptotic expansions in Appendices A.2 and A.3, which are given later.

In this section, we may assume, without loss of generality, that $\Sigma = I_p$. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ be a sequence of non-random k-dimensional vectors and $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots$ be a sequence of *i.i.d.* random vectors with $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}) = I_p$. Set

$$Z = (X'X)^{-1/2}X'\mathscr{E}, \qquad V = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_j' - I_p),$$

and $Q' = (X'X)^{-1/2}X' = (\boldsymbol{q}_1, \dots, \boldsymbol{q}_n)$ and $\boldsymbol{q}_j = (q_1^{(j)}, \dots, q_k^{(j)})'$. The assumptions in Section 2 are used, but A1 and B1 are replaced by more general ones as follows:

A1. For some integer $s \geq 3$, $\limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \|\boldsymbol{x}_j\|^s < \infty$.

B1. For some integer $s \geq 3$, $\mathrm{E}(\|\boldsymbol{\varepsilon}\|^{2s}) < \infty$.

We prepare the following two lemmas.

Lemma A.1. *Under the assumptions* A1, A2 *and* A3,

$$\max\{\|\boldsymbol{q}_j\| : j = 1, \dots, n\} = \mathrm{O}(n^{-\delta/2}),$$

$$\frac{1}{n}\sum_{j=1}^{n} |q_{i(1)}^{(j)} \dots q_{i(r)}^{(j)}| = \mathrm{O}(n^{-r/2}) \qquad (0 < r \le s),$$

where $i(1), \dots, i(r)$ *are arbitrary positive integers not larger than* k.

The proof of Lemma A.1 is easy, and is therefore omitted.

Lemma A.2. *Set*

$$L(n) = \left\{ j : 1 \le j \le n, \inf_{T_1 \ne 0} \frac{n\boldsymbol{q}_j' T_1 T_1' \boldsymbol{q}_j}{\mathrm{tr}(T_1 T_1')} > \frac{1}{2} \right\},$$

where T_1 *is a* $k \times p$ *matrix, then under the assumptions* A1, A2 *and* A3,

$$\liminf_{n \to \infty} n^{-\delta} \# L(n) > 0,$$

where $\# L(n)$ *denotes the number of integers in* $L(n)$.

Proof. If $T_1 \ne 0$, then

$$\begin{aligned} 1 &= \sum_{j=1}^{n} \frac{\boldsymbol{q}_j' T_1 T_1' \boldsymbol{q}_j}{\mathrm{tr}(T_1 T_1')} \\ &\le \frac{n - \# L(n)}{2n} + \# L(n) \max\{\boldsymbol{q}_j' \boldsymbol{q}_j : 1 \le j \le n\} \\ &\le \frac{1}{2} + n^{-\delta} K \# L(n), \end{aligned}$$

for some positive constant K.

For the underlying distribution of $\boldsymbol{\varepsilon}$, we make alternative assumptions:
B1′. For some integer $s \ge 3$, $\mathrm{E}(\|\boldsymbol{\varepsilon}\|^s) < \infty$.
B2′. The Cramér's condition holds, that is, for any $b > 0$,

$$\sup_{\|\boldsymbol{t}\| > b} |\mathrm{E}(\exp(i\boldsymbol{t}'\boldsymbol{\varepsilon}))| < 1,$$

where $\boldsymbol{t}$ is a $p \times 1$ vector.

In the proof of the following theorem, Bhattacharya and Ranga Rao [5] is referred to as BR because of its frequent usage.

Theorem A.1. *Under the assumptions* A1, A2, A3 *and* B1′, B2′,

$$\sup_{B \in \mathscr{B}_{c,\alpha}} \left| \mathrm{P}(Z \in B) - \int_B \psi_{s,n}(z) dz \right| = \mathrm{o}(n^{-(s-2)/2}), \qquad (\mathrm{A.1})$$

where $\psi_{s,n}$ *is the asymptotic expansion of the density function of* Z *which is formally derived up to the order* $n^{-(s-2)/2}$, *and*

$$\mathscr{B}_{c,\alpha} = \{B \in \mathscr{B}^{k\times p} : \Phi((\partial B)^{\varepsilon}) \le c\varepsilon^{\alpha} \text{ for all } \varepsilon > 0\}.$$

Here, $\mathscr{B}^{k\times p}$ *denotes the class of all* $k \times p$*-dimensional Borel sets.*

PROOF. Set $\boldsymbol{r}_j = \mathrm{vec}(\boldsymbol{x}_j\boldsymbol{\varepsilon}_j')$ $(j = 1, 2, \dots)$ and $V_n = n^{-1}\sum_{j=1}^{n}\mathrm{Cov}(\boldsymbol{r}_j)$. Then $V_n = n^{-1}I_p \otimes (X'X)$ and

$$\mathrm{vec}(Z) = (I_p \otimes (X'X)^{-1/2})\sum_{j=1}^{n}\boldsymbol{r}_j,$$

which is a standardized sum of independent random vectors. If the result (20.56) of Theorem 20.6 of BR is true in our problem, then it implies Theorem 2.1. The moment conditions (i), (ii) and (iii) in Theorem 20.6 can be easily checked as well as (20.54) in our case. In our problem, each distribution of $\boldsymbol{r}_j$ is degenerate, and so, the uniform Cramér condition (20.55) is not satisfied. The proof of Theorem 20.6 is similar to that of Theorem 20.1 of BR. Therefore we have to check the parts where the nonsingularity of each covariance matrices and the condition (20.55) have been used. The first part where the nonsingularity of $\mathrm{Cov}(\boldsymbol{r}_j)$ is required is (20.21) and (20.22) with Theorem 9.10 of BR. In the non-*i.i.d.* case, instead of Theorem 9.10, Theorem 9.9 of BR is used in order to estimate (20.21). In Theorem 9.9, the nonsingularity is not required. Let $g_j(\boldsymbol{t}_1)$ be the characteristic function of $\boldsymbol{r}_j$, where $\boldsymbol{t}_1 = \mathrm{vec}(T_1)$. Then Lemma A.1 implies that for any positive constant δ, there exists a positive constant d such that

$$\begin{aligned} &\|\boldsymbol{t}_1\| < d \qquad \text{implies} \qquad \sup_j |g_j(\boldsymbol{t}_1) - 1| < \frac{1}{2} - \delta, \\ &\lim_{n\to\infty} \max\{\mathrm{P}(\|\boldsymbol{r}_j\| > \sqrt{n}) : j = 1, \dots, n\} = 0, \\ &\sup_n \left[\max\left\{\int_{\{\|\boldsymbol{r}_j\|>\sqrt{n}\}} \|\boldsymbol{r}_j\|^{s'} : j = 1, \dots, n\right\}\right] < K, \end{aligned} \tag{A.2}$$

for some $K > 0$ and $s' \ge 3$. Replace the right-hand side of (20.21) with

$$c_6(s,k)n^{-(s+k-1)/2}\eta_{s+k+1},$$

and A_n in (20.22) with

$$A_n = c_7(s,k)n^{1/2}(\eta_{s+k+1})^{-1/(s+k-1)},$$

where

$$\eta_r = \frac{1}{n}\sum_{j=1}^{n}\mathrm{E}(\|V_n^{-1/2}\boldsymbol{r}_{j,n}\|^{r}) \qquad (r > 0).$$

Here $\boldsymbol{r}_{j,n}$ is a truncated and centralized version of $\boldsymbol{r}_j$ as in Theorem 20.1. Then (20.23) holds with

$$\rho_s = \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}(\|\boldsymbol{r}_j\|^s).$$

In the rest of the proof of Theorem 20.1, (20.26), (20.35), (20.36), (20.37), (20.39), (20.42) and (20.43) use the covariance matrix D_n of (in our problem) $\boldsymbol{r}_{j,n}$. In the non-*i.i.d.* case we can replace D_n with

$$\frac{1}{n}\sum_{j=1}^{n} \mathrm{Cov}(\boldsymbol{r}_{j,n})$$

which is nonsingular for sufficiently large n by (20.54). Therefore we do not need to require the nonsingularity of each covariance matrix of $\boldsymbol{r}_j$. It remains to check the parts where the uniform Cramér condition (20.55) is used. The unique part where (20.55) is used is the estimation of I_1 in (20.26). Let $g_{j,n}(\boldsymbol{t}_1)$ be the characteristic function of $R_{j,n}$. Then the left-hand side of (20.28) is estimated as

$$\begin{aligned}|D^{\beta-\alpha}\hat{Q}_n'(\boldsymbol{t}_1)| &= \left|D^{\beta-\alpha}\prod_{j=1}^{n} g_{j,n}(n^{-1/2}V_n^{-1/2}\boldsymbol{t}_1)\right| \\ &\le n^{|\beta-\alpha|}\left(\frac{n}{\lambda_n}\right)^{|\beta-\alpha|}\left[\sup_{1\le j\le n}\mathrm{E}(\|n^{-1/2}\boldsymbol{r}_{j,n}\|^{|\beta-\alpha|})\right] \\ &\quad\times\left[\sup_{J_n}\left|\prod_{j\notin J_n} g_{j,n}(n^{-1/2}V_n^{-1/2}\boldsymbol{t}_1)\right|\right],\end{aligned}$$

where $\beta-\alpha$ is a nonnegative integral vector, $|\beta-\alpha| = \sum v_i$, $D^{\beta-\alpha} = (\partial/\partial t_1)^{v_1}\ldots(\partial/\partial t_{kp})^{v_{kp}}$ for $\beta-\alpha = (v_1,\ldots,v_{kp})$, $\boldsymbol{t}_1 = (t_1,\ldots,t_{kp})'$ and $J_n = \{j_1,\ldots,j_{|\beta-\alpha|} : 1\le j_1\le\cdots\le j_{|\beta-\alpha|}\le n\}$. From Lemma 14.1 of BR and A3, we see that if $s' = |\beta-\alpha| > s$, then

$$\mathrm{E}(\|n^{-1/2}\boldsymbol{r}_{j,n}\|^{|\beta-\alpha|}) \le 2^{s'}n^{(s'-s)/2}n^{-s/2}\|\boldsymbol{x}_j\|^s\mathrm{E}(\|\boldsymbol{\varepsilon}\|^s) = \mathrm{o}(n^{(s'-s)/2}),$$

and if $s'\le s$, then

$$\mathrm{E}(\|n^{-1/2}\boldsymbol{r}_{j,n}\|^{|\beta-\alpha|}) \le 2^{s'}n^{-s/2}\|\boldsymbol{x}_j\|^s\mathrm{E}(\|\boldsymbol{\varepsilon}\|^s) = \mathrm{o}(1).$$

Furthermore,

$$\begin{aligned}&|g_{j,n}(n^{-1/2}V_n^{-1/2}\boldsymbol{t}_1)| \\ &\quad\le |\mathrm{E}[\exp\{i\boldsymbol{t}_1'(I_p\otimes(X'X)^{-1/2})(\boldsymbol{\varepsilon}_j\otimes\boldsymbol{x}_j)\}]| + 2\mathrm{P}(\|\boldsymbol{\varepsilon}\| > n^{1/2}M_n^{-1}) \\ &\quad= |\mathrm{E}(\exp(i\boldsymbol{q}_j'T_1\boldsymbol{\varepsilon}))| + 2\mathrm{P}(\|\boldsymbol{\varepsilon}\| > n^{1/2}M_n^{-1}).\end{aligned}$$

Therefore

$$\begin{aligned}
\sup_{\operatorname{tr}(T_1T_1')>b}&\left|\prod_{j\notin J(m,n)}\mathrm{E}(\exp(in^{1/2}\boldsymbol{q}_j'T_1\boldsymbol{\varepsilon}_j))\right| \\
&\le \prod_{j\in J(m,n)^c\cap L(n)}\sup_{\operatorname{tr}(T_1T_1')>b}|\mathrm{E}(\exp(in^{1/2}\boldsymbol{q}_j'T_1\boldsymbol{\varepsilon}_j))| \\
&\le \prod_{j\in J(m,n)\cap L(n)}\sup_{\tau_j'\tau_j>b/2}|\mathrm{E}(\exp(i\boldsymbol{\tau}_j'\boldsymbol{\varepsilon}_j))| \\
&\le \left(\sup_{\tau'\tau>b/2}|\mathrm{E}(\exp(i\boldsymbol{\tau}'\boldsymbol{\varepsilon}))|\right)^{\#L(n)-m},
\end{aligned}$$

where $\boldsymbol{\tau}_j = n^{1/2}T_1'\boldsymbol{q}_j$. By the same argument as in Theorem 20.1, from Lemma 2.2 and (A.2) we see that $I_1 = \mathrm{o}(n^{-(s-2)/2})$.

In order to expand the joint distribution of (Z, V) up to the order $\mathrm{o}(n^{-(s-2)/2})$, we use the assumptions B1 and B2.

Corollary A.2. *Under the assumptions* A1, A2, A3 *and* B1, B2,

$$\sup_{B\in\mathscr{B}_{c,\alpha}}\left|\mathrm{P}((Z,V)\in B)-\int_B\psi_{s,n}(z,v)dzdv\right| = \mathrm{o}(n^{-(s-2)/2}), \tag{A.3}$$

where $\psi_{s,n}$ is an asymptotic expansion of the joint density function of (Z, V) formally derived up to the order $n^{-(s-2)/2}$, and

$$\mathscr{B}_{c,\alpha} = \{B\in\mathscr{B}^{kp+p(p+1)/2} : \Phi((\partial B)^{\varepsilon}) \le c\varepsilon^{\alpha} \ \text{for all}\ \ \varepsilon>0\}.$$

Here, $\mathscr{B}^{kp+p(p+1)/2}$ denotes the class of all $kp+p(p+1)/2$-dimensional Borel sets, considering (Z, V) as a point of $kp+p(p+1)/2$-dimensional Euclidean space.

A.2. Edgeworth expansion of *t*-statistic

In this section, using the assumption $\Sigma = I_p$ as in the previous section, we derive an asymptotic expansion for the distribution function of U.

In order to get a valid expansion of U up to the order n^{-1}, we need some assumptions for the design matrix X and the distribution of $\boldsymbol{\varepsilon}$. For the design matrix X, we assume A1 with $s = 4$, A2 and A3, given in Appendix A.1. Moreover, for the distribution of $\boldsymbol{\varepsilon}$, we also assume B1 with $s = 4$ and B2.

Since S_e is rewritten as $S_e = \mathscr{E}'\mathscr{E} - Z'Z = nI_p + \sqrt{n}V - Z'Z$, we can expand $(S_e/n)^{-1/2}$ as

$$\left(\frac{1}{n}S_e\right)^{-1/2} = I_p - \frac{1}{2\sqrt{n}}V + \frac{1}{n}\left(\frac{3}{8}V^2 + \frac{1}{2}Z'Z\right) + \mathrm{O}_p(n^{-3/2}).$$

Therefore

$$U = Z - \frac{1}{2\sqrt{n}} ZV + \frac{1}{n} Z\left(\frac{3}{8} V^2 + \frac{1}{2} Z'Z\right) + \mathrm{O}_p(n^{-3/2}). \tag{A.4}$$

Using (A.4), under the assumptions A1, A2, A3 and B1 the characteristic function $C_U(T_1)$ of U can be expanded as

$$\begin{aligned} C_U(T_1) &= \mathrm{E}[\exp\{i \,\mathrm{tr}(T_1'U)\}] \\ &= \mathrm{E}\Bigg[\exp\{i \,\mathrm{tr}(T_1'Z)\}\Bigg\{1 - \frac{i}{2\sqrt{n}} \mathrm{tr}(T_1'ZV) \\ &\quad + \frac{1}{n}\left\{\frac{3}{8} i \,\mathrm{tr}(T_1'ZV^2) + \frac{i}{2} \mathrm{tr}(T_1'ZZ'Z) + \frac{i^2}{8} (\mathrm{tr}(T_1'ZV))^2\right\}\Bigg\}\Bigg] + \mathrm{o}(n^{-1}) \\ &= C_U^{(0)}(T_1) + \frac{1}{\sqrt{n}} C_U^{(1)}(T_1) + \frac{1}{n} C_U^{(2)}(T_1) + \mathrm{o}(n^{-1}), \end{aligned}$$

where $T_1 = [t_{a'a}^{(1)}]$ is a $k \times p$ matrix. Each term in the expansion of $C_U(T_1)$ can be evaluated by using the joint characteristic function of Z and V, which can be expressed as

$$\begin{aligned} \Psi(T_1, T_2) &= \mathrm{E}[\exp\{i \,\mathrm{tr}(T_1'Z + n^{-1/2} T_2 V)\}] \\ &= \prod_{\alpha=1}^{n} \mathrm{E}[\exp\{i \,\mathrm{tr}(T_1' \boldsymbol{q}_\alpha \boldsymbol{\varepsilon}' + n^{-1/2} T_2(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' - I_p))\}] \\ &= \prod_{\alpha=1}^{n} h_\alpha(T_1, T_2) = \exp\left[\sum_{\alpha=1}^{n} \log\{h_\alpha(T_1, T_2)\}\right], \end{aligned}$$

where T_2 is defined in §2. Then the following identities hold:

$$\begin{aligned} &\mathrm{E}[\exp\{i \,\mathrm{tr}(T_1'Z)\}\{i \,\mathrm{tr}(T_1'ZV)\}] \\ &\qquad = -i \sum_{a'=1}^{k} \sum_{ab}^{p} t_{a'a}^{(1)} \frac{\partial^2}{\partial t_{a'b}^{(1)} \partial t_{ab}^{(2)}} \Psi(T_1, T_2)\Bigg|_{T_2=0}, \\ &\mathrm{E}[\exp\{i \,\mathrm{tr}(T_1'Z)\}\{i \,\mathrm{tr}(T_1'ZV^2)\}] \\ &\qquad = \sum_{a'=1}^{k} \sum_{abc}^{p} t_{a'c}^{(1)} \frac{\partial^3}{\partial t_{a'a}^{(1)} \partial t_{ab}^{(2)} \partial t_{bc}^{(2)}} \Psi(T_1, T_2)\Bigg|_{T_2=0}, \end{aligned}$$

$$\mathrm{E}[\exp\{i\ \mathrm{tr}(T_1'Z)\}\{i\ \mathrm{tr}(T_1'ZZ'Z)\}]$$

$$= \sum_{a'b'}^{k} \sum_{ab}^{p} t_{a'b}^{(1)} \frac{\partial^3}{\partial t_{a'a}^{(1)} \partial t_{b'a}^{(1)} \partial t_{b'b}^{(1)}} \Psi(T_1, T_2)\bigg|_{T_2=0},$$

$$\mathrm{E}[\exp\{i\ \mathrm{tr}(T_1'Z)\}\{i\ \mathrm{tr}(T_1'ZV)\}^2]$$

$$= \sum_{a'b'}^{k} \sum_{abcd}^{p} t_{a'c}^{(1)} t_{b'd}^{(1)} \frac{\partial^4}{\partial t_{a'a}^{(1)} \partial t_{b'b}^{(1)} \partial t_{ac}^{(2)} t_{bd}^{(2)}} \Psi(T_1, T_2)\bigg|_{T_2=0}.$$

Note that

$$\frac{\partial^2}{\partial t_{a'b}^{(1)} \partial t_{ab}^{(2)}} \Psi(T_1, T_2)\bigg|_{T_2=0}$$

$$= \sum_{\alpha=1}^{n} \left[\frac{\partial^2}{\partial t_{a'b}^{(1)} \partial t_{ab}^{(2)}} \log\{h_\alpha(T_1, T_2)\} + \frac{\partial}{\partial t_{a'b}^{(1)}} \log\{h_\alpha(T_1, T_2)\} \frac{\partial}{\partial t_{ab}^{(2)}} \log\{h_\alpha(T_1, T_2)\} \right] \Psi(T_1, T_2)\bigg|_{T_2=0},$$

and

$$\Psi(T_1, 0) = \exp\left\{ \frac{i^2}{2} \mathrm{tr}(T_1'T_1) + \frac{i^3}{6\sqrt{n}} \sum_{a'b'c'}^{k} \sum_{abc}^{p} t_{a'a}^{(1)} t_{b'b}^{(1)} t_{c'c}^{(1)} \bar{\chi}_{a'b'c'} \kappa_{abc} + \frac{i^4}{24n} \sum_{a'b'c'd'}^{k} \sum_{abcd}^{p} t_{a'a}^{(1)} t_{b'b}^{(1)} t_{c'c}^{(1)} t_{d'd}^{(1)} \bar{\chi}_{a'b'c'd'} \kappa_{abcd} + \mathrm{o}(n^{-1}) \right\},$$

where the coefficient $\bar{\chi}_{a_1 \dots a_j}$ is defined by

$$\bar{\chi}_{a_1 \dots a_j} = \frac{1}{n} \sum_{i=1}^{n} \prod_{l=1}^{j} \sqrt{n} q_{a_l}^{(j)},$$

e.g.,

$$\bar{\chi}_{abc} = \sqrt{n} \sum_{j=1}^{n} q_a^{(j)} q_b^{(j)} q_c^{(j)}, \qquad \bar{\chi}_{abcd} = n \sum_{j=1}^{n} q_a^{(j)} q_b^{(j)} q_c^{(j)} q_d^{(j)}.$$

Therefore we can get an expansion of $C_U(T_1)$, whose formal inversion yields a valid expansion of the distribution function of U as in the following Theorem A.2.

Theorem A.2. *Suppose that X and $\mathscr{E}$ satisfy* A1, A2, A3, B1 *and* B2 *with $s = 4$. Set $\boldsymbol{u} = \mathrm{vec}(U)$, then the distribution function of U can be expanded as*

$$\mathrm{P}(\mathrm{vec}(U) \le \boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_{pk}} \phi_{pk}(\boldsymbol{u}) \left[1 + \frac{1}{\sqrt{n}} R_1(\boldsymbol{u}) + \frac{1}{n} R_2(\boldsymbol{u})\right] d\boldsymbol{u} + \mathrm{o}(n^{-1}),$$

where

$$R_1(\boldsymbol{u}) = -\frac{1}{2} \sum_{a'=1}^{k} \sum_{ab}^{p} \bar{\chi}_{a'} \kappa_{abb} H_{a'a}(\boldsymbol{u}) + \frac{1}{6} \sum_{a'b'c'}^{k} \sum_{abc}^{p} (\bar{\chi}_{a'b'c'} - 3\bar{\chi}_{a'} \delta_{b'c'}) \kappa_{abc} H_{a'a,b'b,c'c}(\boldsymbol{u}), \tag{A.5}$$

$$\begin{aligned} R_2(\boldsymbol{u}) = {} & \frac{1}{8} \sum_{a'b'}^{k} \sum_{abcd}^{p} \bar{\chi}_{a'} \bar{\chi}_{b'} (\kappa_{acc}\kappa_{bdd} + 3\kappa_{abc}\kappa_{cdd} + 4\kappa_{acd}\kappa_{bcd}) H_{a'a,b'b}(\boldsymbol{u}) \\ & + \frac{1}{2}(p + k + 1) \sum_{a'}^{k} \sum_{a}^{p} H_{a'a,a'a}(\boldsymbol{u}) \\ & + \frac{1}{24} \sum_{a'b'c'd'}^{k} \left[\sum_{abcd}^{p} (\bar{\chi}_{a'b'c'd'} - 3\delta_{a'b'}\delta_{c'd'}) \kappa_{abcd} \right. \\ & - 2 \sum_{abcde}^{p} \{ \bar{\chi}_{a'b'c'} \bar{\chi}_{d'} (\kappa_{abc}\kappa_{dee} + 3\kappa_{ade}\kappa_{bce}) \\ & \left. - 3\bar{\chi}_{a'}\bar{\chi}_{b'}\delta_{c'd'} (\kappa_{aee}\kappa_{bcd} + \kappa_{abe}\kappa_{cde} + 2\kappa_{ace}\kappa_{bde}) \} \right] H_{a'a,b'b,c'c,d'd}(\boldsymbol{u}) \\ & + \frac{1}{4} \sum_{a'b'}^{k} \sum_{ab}^{p} H_{a'a,a'b,b'a,b'b}(\boldsymbol{u}) \\ & + \frac{1}{72} \sum_{a'b'c'd'e'f'}^{k} \sum_{abcbdef}^{p} (\bar{\chi}_{a'b'c'} \bar{\chi}_{d'e'f'} - 6\bar{\chi}_{a'b'c'} \bar{\chi}_{d'} \delta_{e'f'} \\ & + 9\bar{\chi}_{a'} \delta_{b'c'} \bar{\chi}_{d'} \delta_{e'f'}) \kappa_{abc} \kappa_{def} H_{a'a,b'b,c'c,d'd,e'e,f'f}(\boldsymbol{u}). \end{aligned} \tag{A.6}$$

Here $\phi_{pk}(\boldsymbol{u})$ is the probability density function of $\mathrm{N}_{pk}(\mathbf{0}, I_{pk})$ given by $\phi_{pk}(\boldsymbol{u}) = (2\pi)^{-pk/2} \exp(-\boldsymbol{u}'\boldsymbol{u}/2)$, and $H_{a_1'a_1,\ldots,a_j'a_j}(\boldsymbol{u})$ is the Hermite polynomial defined by

$$H_{a_1'a_1,\ldots,a_j'a_j}(\boldsymbol{u}) = (-1)^j \frac{\partial^j}{\partial u_{a_1'a_1} \ldots \partial u_{a_j'a_j}} \phi_{pk}(\boldsymbol{u}),$$

where $u_{a'a}$ is the (a', a)th element of U. Furthermore, we call the integrand for the distribution function of U the pseudo density function of U.

Using Corollary A.2, we can prove the validity of Theorem A.2 in the same manner as in Bhattacharya and Ghosh [3].

When $k = 1$, the test statistic becomes the usual multivariate t-statistic defined by

$$\boldsymbol{u} = \sqrt{n} S^{-1/2}(\bar{\boldsymbol{y}} - \mathrm{E}(\boldsymbol{y})),$$

where $\bar{\boldsymbol{y}} = n^{-1} \sum_{j=1}^{n} \boldsymbol{y}_j$ and $S = n^{-1} \sum_{j=1}^{n} (\boldsymbol{y}_j - \bar{\boldsymbol{y}})(\boldsymbol{y}_j - \bar{\boldsymbol{y}})'$. Its pseudo density can be written as

$$\begin{aligned}
\phi_p(\boldsymbol{u}) \Bigg[1 &- \frac{1}{6\sqrt{n}} \left\{ 3 \sum_{ab}^{p} \kappa_{abb} H_a(\boldsymbol{u}) + 2 \sum_{abc}^{p} \kappa_{abc} H_{a,b,c}(\boldsymbol{u}) \right\} \\
&+ \frac{1}{24n} \Bigg\{ 3 \sum_{abcd}^{p} (\kappa_{acc}\kappa_{bdd} + 3\kappa_{abc}\kappa_{cdd} + 4\kappa_{acd}\kappa_{bcd}) H_{a,b}(\boldsymbol{u}) \\
&\quad + 12(p+2) \sum_{a}^{p} H_{a,a}(\boldsymbol{u}) + 4 \sum_{abcde}^{p} (\kappa_{aee}\kappa_{bcd} + 3\kappa_{abe}\kappa_{cde}) H_{a,b,c,d}(\boldsymbol{u}) \\
&\quad - 2 \sum_{abcd}^{p} \kappa_{abcd} H_{a,b,c,d}(\boldsymbol{u}) + 6 \sum_{ab}^{p} H_{a,a,b,b}(\boldsymbol{u}) \\
&\quad + \frac{3}{4} \sum_{abcdef}^{p} \kappa_{abc}\kappa_{def} H_{a,b,c,d,e,f}(\boldsymbol{u}) \Bigg\} \Bigg].
\end{aligned}$$

The corresponding results in the case where $S = (n-1)^{-1} \sum_{j=1}^{n} (\boldsymbol{y}_j - \bar{\boldsymbol{y}})(\boldsymbol{y}_j - \bar{\boldsymbol{y}})'$ was derived by Fujikoshi [10].

The moment condition B1 will be replaced with B1$'$ with $s = 4$ in Appendix A.1 as in Hall [14], Bhattacharya and Ghosh [4] and Babu and Bai [2].

A.3. Outline of computation on Theorem 2.1

In this section, we explain our method for finding an asymptotic expansion of the null distribution of T_G up to the order n^{-1}. Without loss of generality, we may replace $\mathscr{E}$ by $\mathscr{E}\Sigma^{-1/2}$, which has $\mathrm{E}[\mathrm{vec}(\mathscr{E})] = \boldsymbol{0}$ and $\mathrm{Cov}[\mathrm{vec}(\mathscr{E})] = I_{np}$, in the expressions of T_G, since T_G is invariant under the transformation from Y to $Y\Sigma^{-1/2}$.

Suppose that X and the distribution of $\boldsymbol{\varepsilon}$ satisfy A1 with $s = 4$, A2, A3, B1 with $s = 4$ and B2. Note that T_G is a smooth function of U. From the results of Chandra and Ghosh [6] and Corollary A.2, it can be shown that T_G has a valid expansion up to the order n^{-1} under the assumptions A1, A2, A3, B1 and B2. In the following we will derive an asymptotic expansion of the characteristic function of T_G up to the order n^{-1}, which can be inverted formally. From (1.1), we can write the characteristic function of T_G as

$$C_{T_G}(t) = C_0(t) + \frac{1}{n} C_1(t) + \mathrm{o}(n^{-1}), \tag{A.7}$$

where

$$C_0(t) = \mathrm{E}[\exp\{it\ \mathrm{tr}(U'\Omega U)\}],$$
$$C_1(t) = it\mathrm{E}[\{(r_1 - k)\ \mathrm{tr}(U'\Omega U) + r_2(\mathrm{tr}(U'\Omega U))^2\} \exp\{it\ \mathrm{tr}(U'\Omega U)\}].$$

For an evaluation of each term in (A.7), we will use the pseudo density function of U in Theorem A.2.

For the computation of $C_0(t)$, using the pseudo density of U, we have

$$C_0(t) = \int_{\mathscr{R}^{pk}} \exp\{it\ \mathrm{tr}(U'\Omega U)\} \times \phi_{pk}(\boldsymbol{u})\left\{1 + \frac{1}{\sqrt{n}} R_1(\boldsymbol{u}) + \frac{1}{n} R_2(\boldsymbol{u})\right\} d\boldsymbol{u} + \mathrm{o}(n^{-1}),$$

where $R_1(\boldsymbol{u})$ and $R_2(\boldsymbol{u})$ are defined by (A.5) and (A.6) respectively. Set $\varphi = (1-2it)^{-1}$ and $\Gamma = I_k + (\varphi - 1)\Omega$. Then

$$\exp\{it\ \mathrm{tr}(U'\Omega U)\} \exp\left\{-\frac{1}{2}\ \mathrm{tr}(U'U)\right\} = \exp\left\{-\frac{1}{2}\ \mathrm{tr}(U'\Gamma^{-1}U)\right\}.$$

Using the transformation from U to $U^* = \Gamma^{-1/2}U$ and the equation $\boldsymbol{u} = \mathrm{vec}(\Gamma^{1/2}U^*) = (I_p \otimes \Gamma^{1/2})\boldsymbol{u}^*$, where $\boldsymbol{u}^* = \mathrm{vec}(U^*)$, $C_0(t)$ is expressed as the expectation on U^*, that is

$$C_0(t) = \mathrm{E}_{U^*}\left[1 + \frac{1}{\sqrt{n}} R_1((I_p \otimes \Gamma^{1/2})\boldsymbol{u}^*) + \frac{1}{n} R_2((I_p \otimes \Gamma^{1/2})\boldsymbol{u}^*)\right] + \mathrm{o}(n^{-1}). \tag{A.8}$$

This expectation is taken with respect to U^* whose columns are independently distributed as $\mathrm{N}_p(\boldsymbol{0}, I_p)$. Set $W = \Gamma^{1/2}U^*$, then it is seen that $\boldsymbol{w} = \mathrm{vec}(W) \sim \mathrm{N}_{pk}(\boldsymbol{0}, I_p \otimes \Gamma)$. Therefore the expansion (A.8) can be rewritten as

$$C_0(t) = \varphi^{ph/2}\mathrm{E}_W\left[1 + \frac{1}{\sqrt{n}} R_1(\boldsymbol{w}) + \frac{1}{n} R_2(\boldsymbol{w})\right] + \mathrm{o}(n^{-1}). \tag{A.9}$$

Applying a similar method to $C_1(t)$ yields

$$C_1(t) = \frac{1}{2}(1-\varphi^{-1})\varphi^{ph/2}\mathrm{E}_W[\{(r_1-k)\ \mathrm{tr}(W'\Omega W) + r_2\{\mathrm{tr}(W'\Omega W)\}^2\}] + \mathrm{o}(1). \tag{A.10}$$

Through the calculations of (A.9) and (A.10), we use the following identities which are expectations of the Hermite polynomials, and the relations among the elements of Ω. As for the former, let the (a,b)th elements of Ω and Γ be denoted by ω_{ab} and γ_{ab} respectively. Note that $\gamma_{ab} = \delta_{ab} + (\varphi-1)\omega_{ab}$ and

$$\begin{aligned}
&\mathrm{E}_W[H_{a'a}(\boldsymbol{w})] = 0, \qquad \mathrm{E}_W[H_{a'a,b'b}(\boldsymbol{w})] = (\varphi-1)\omega_{a'b'}\delta_{ab},\\
&\mathrm{E}_W[H_{a'a,b'b,c'c}(\boldsymbol{w})] = 0,\\
&\mathrm{E}_W[H_{a'a,b'b,c'c,d'd}(\boldsymbol{w})] = (\varphi-1)^2\sum_{[3]}\omega_{a'b'}\omega_{c'd'}\delta_{ab}\delta_{cd},\\
&\mathrm{E}_W[H_{a'a,b'b,c'c,d'd,e'e,f'f}(\boldsymbol{w})] = (\varphi-1)^3\sum_{[15]}\omega_{a'b'}\omega_{c'd'}\omega_{e'f'}\delta_{ab}\delta_{cd}\delta_{ef},
\end{aligned} \tag{A.11}$$

where $\sum_{[j]}$ means the sum of all j possible combinations of the set a_i and a_i', for example

$$\sum_{[3]}\omega_{a'b'}\omega_{c'd'}\delta_{ab}\delta_{cd} = \omega_{a'b'}\omega_{c'd'}\delta_{ab}\delta_{cd} + \omega_{a'c'}\omega_{b'd'}\delta_{ac}\delta_{bd} + \omega_{a'd'}\omega_{b'c'}\delta_{ad}\delta_{bc}.$$

As for the latter, using the property that Ω is an idempotent matrix, we have

$$\begin{aligned}
&\sum_{c=1}^{k}\omega_{ac}\omega_{bc} = \omega_{ab},\\
&\mathrm{tr}(\Omega) = \sum_{a=1}^{k}\omega_{aa} = h, \qquad \mathrm{tr}(\Omega^2) = \sum_{ab}^{k}\omega_{ab}^2 = h,\\
&\mathrm{tr}(\Omega^3) = \sum_{abc}^{k}\omega_{ab}\omega_{bc}\omega_{ac} = h, \qquad \mathrm{tr}(\Omega^4) = \sum_{abcd}^{k}\omega_{ab}\omega_{bc}\omega_{cd}\omega_{ad} = h.
\end{aligned} \tag{A.12}$$

Substituting (A.11) and (A.12) into both of (A.9) and (A.10) yields

$$C_T(t) = \varphi^{ph/2}\left[1 + \frac{1}{n}\sum_{j=0}^{3} b_j\varphi^j + \mathrm{o}(n^{-1})\right], \tag{A.13}$$

where the coefficients b_j's are defined by (2.1). Finally, by inverting (A.13), we have Theorem 2.1.

Acknowledgement

The authors would like to thank the referee for several valuable comments.

References

[1] T. W. Anderson, An introduction to multivariate analysis (2nd ed.), John Wiley & Sons, 1984.

[2] G. J. Babu and Z. D. Bai, Edgeworth expansions of a function of sample means under minimal moment conditions and partial Cramér's condition, Sankhyā Ser. **A 55** (1993), 244–258.

[3] R. N. Bhattacharya and J. K. Ghosh, On the validity of formal Edgeworth expansion, Ann. Statist. **6** (1978), 434–451.

[4] R. N. Bhattacharya and J. K. Ghosh, On moment conditions for valid formal Edgeworth expansions, J. Multivariate Anal. **27** (1988), 68–79.

[5] R. N. Bhattacharya and R. Ranga Rao, Normal approximation and asymptotic expansions, John Wiley & Sons, New York, 1976.

[6] T. K. Chandra and J. K. Ghosh, Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables, Sankhyā Ser. **A 41** (1979), 22–47.

[7] G. R. Chase and W. G. Bulgren, A Monte Carlo investigation of the robustness of T^2, J. Amer. Statist. Assoc. **66** (1971), 499–502.

[8] B. S. Everitt, A Monte Carlo investigation of the robustness of Hotelling's one- and two-sample T^2 tests, J. Amer. Statist. Assoc. **74** (1979), 48–51.

[9] Y. Fujikoshi, Two-way ANOVA models with unbalanced data, Discrete Math. **116** (1993), 315–334.

[10] Y. Fujikoshi, An asymptotic expansion for the distribution of Hotelling's T^2-statistic under nonnormality, J. Multivariate Anal. **61** (1997), 187–193.

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

[12] Y. Fujikoshi, Asymptotic expansions for the distributions of multivariate basic statistics and one-way MANOVA tests under nonnormality, J. Statist. Plann. Inference. (2001), to appear.

[13] Y. Fujikoshi, M. Ohmae and H. Yanagihara, Asymptotic approximations of the null distribution of the one-way ANOVA test statistic under nonnormality, J. Japan Statist. Soc. **29** (1999), 147–161.

[14] P. Hall, Edgeworth expansion for student's t statistic under minimal moment condition, Ann. Probab. **15** (1987), 920–931.

[15] P. J. Huber, Robust regression: asymptotics, conjectures and Monte Carlo, Ann. Statist. **1** (1973), 799–821.

[16] T. Isogai, On measures of multivariate skewness and kurtosis, Math. Japon. **28** (1983), 251–261.

[17] Y. Kano, An asymptotic expansion of the distribution of Hotellings T^2-statistic under general condition, Amer. J. Math. Management Sci. **15** (1995), 317–341.

[18] E. L. Kaplan, Tensor notation and the sampling cumulants of k-statistics, Biometrika **39** (1952), 319–323.

[19] K. V. Mardia, Measures of multivariate skewness and kurtosis with applications, Biometrika **57** (1970), 519–530.

[20] K. V. Mardia, Assessment of multinormality and the robustness of Hotelling's T^2 test, J. Roy. Statist. Soc. Ser. **C** Appl. Statist. **24** (1975), 163–171.

[21] K. V. Mardia and P. J. Zemroch, Tables of the *F*- and related distributions with algorithms, Academic Press, Inc., London-New York, 1978.

[22] M. B. Qumsiyeh, Edgeworth expansion in regression models, J. Multivariate Anal. **35** (1990), 86–101.

[23] M. Siotani, T. Hayakawa and Y. Fujikoshi, Modern multivariate statistical analysis: a graduate course and handbook, American Sciences Press, Columbus, Ohio, 1985.

[24] H. Yanagihara, Asymptotic expansion of the null distribution of test statistic for linear hypothesis in nonnormal linear model, TR No. 99-16, Statistical Research Group, Hiroshima University, 1999.

Hirofumi Wakaki and Yasunori Fujikoshi
Department of Mathematics
Graduate School of Science
Hiroshima University
Higashi-Hiroshima 739-8526, *Japan*
Hirofumi Wakaki: wakaki@math.sci.hiroshima-u.ac.jp
Yasunori Fujikoshi: fuji@math.sci.hiroshima-u.ac.jp

Hirokazu Yanagihara
Department of Statistical Methodology
Institute of Statistical Mathematics
Tokyo 106-8569, *Japan*
Hirokazu Yanagihara: hyanagi@ism.ac.jp