

Asymptotic expansions of the distributions of MANOVA test statistics when the dimension is large

Hirofumi WAKAKI, Yasunori FUJIKOSHI and Vladimir V. ULYANOV

(Received October 1, 2013)

(Revised February 26, 2014)

ABSTRACT. Asymptotic expansions of the null distribution of the MANOVA test statistics including the likelihood ratio, Lawley-Hotelling and Bartlett-Nanda-Pillai tests are obtained when both the sample size and the dimension tend to infinity with assuming the ratio of the dimension and the sample size tends to a positive constant smaller than one. Cornish-Fisher expansions of the upper percent points are also obtained. In order to study the accuracy of the approximation formulas, some numerical experiments are done, with comparing to the classical expansions when only the sample size tends to infinity.

1. Introduction

This paper is concerned with the multivariate linear model:

$$X = Q\mathbf{B} + \mathcal{E}, \quad (1)$$

where X is the $N \times p$ observation matrix, Q is the $N \times k$ design matrix, \mathbf{B} is the $k \times p$ matrix of regression coefficients, and, \mathcal{E} is the $N \times p$ error matrix distributed as $N_{N \times p}(O, I_N \otimes \Sigma)$. Consider testing the hypothesis

$$H_0 : C\mathbf{B} = O, \quad (2)$$

where C is a $q \times k$ known matrix of rank q . Under a certain group of transformations, the testing problem is invariant, and invariant tests depend on the non-zero eigenvalues of $S_h S_e^{-1}$ where

$$S_h = \hat{\mathbf{B}}' C' (C(Q'Q)^{-1}C')^{-1} C \hat{\mathbf{B}} \quad \text{and} \quad S_e = (X - Q\hat{\mathbf{B}})'(X - Q\hat{\mathbf{B}})$$

with $\hat{\mathbf{B}} = (Q'Q)^{-1}Q'X$ (see [4, Theorem 10.2.1]). Among famous invariant tests, we consider the three statistics

$$(i) \ A = \frac{|S_e|}{|S_e + S_h|}, \quad (ii) \ T_0^2 = \text{tr } S_h S_e^{-1}, \quad \text{and} \quad (iii) \ V^{(s)} = \text{tr } S_h (S_e + S_h)^{-1},$$

The third author is partly supported by RFBR grant, No. 14-01-00500.

2010 *Mathematics Subject Classification.* Primary 62H10; Secondary 62E20.

Key words and phrases. MANOVA, high dimension, likelihood ratio, Lawley-Hotelling, Bartlett-Nanda-Pillai, Edgeworth expansion.

which are the likelihood ratio test statistics, Lawley-Hotelling's generalized T^2 statistics, and Bartlett-Nanda-Pillai's test statistics, respectively.

Since the exact distributions of these statistics are complicated and not easy to treat, we need some method to approximate the distributions. One of the method is to use the asymptotic expansions of the distribution functions when the sample size N is large (see [1], [4], or [5]). However, it is known that the approximations are bad when the dimension p is large relative to N .

Tonda and Fujikoshi [6] derived an asymptotic expansion of the null distribution function for the LR test under the framework:

$$q : \text{fixed}, n \rightarrow \infty, p \rightarrow \infty, \frac{p}{n} \rightarrow c \in (0, 1), \quad (3)$$

where $n = N - k$ is the degree of freedom of the Wishart distribution of S_e . Their derivation depends on property of Wilks's lambda distribution:

$$A_p(q, n) = A_q(p, n - p + q) = \prod_{j=1}^q \text{Beta}((n - p + q - j + 1)/2, p/2).$$

In their paper, the problems of deriving the asymptotic expansions of the null distributions for (ii), (iii), and the non-null distributions for the three statistics were left for future work. In this paper, we derive the asymptotic expansions of the null distributions for the three statistics under a similar framework (6) as (3) in a unified way. The results was first presented in [7] as a preprint.

2. Preliminaries

In this section we prepare some lemmas.

By transforming the variables and parameters, the model (1) can be represented as

$$X^* = \begin{pmatrix} X_1^* \\ X_2^* \\ X_3^* \end{pmatrix}$$

where X_1^* , X_2^* and X_3^* are independent random matrices, and

$$X_1^* \sim N_{q \times p}(M_1, I_q \otimes \Sigma), \quad X_2^* \sim N_{(k-q) \times p}(M_2, I_{k-q} \otimes \Sigma),$$

$$X_3^* \sim N_{n \times p}(O, I_n \otimes \Sigma),$$

where $n = N - k$. The null hypothesis (2) is equivalent with $H : M_1 = O$. The basic statistics S_e and S_h can be represented as

$$S_e = (X_3^*)' X_3^*, \quad \text{and} \quad S_h = (X_1^*)' X_1^*.$$

Here A' denotes the transpose of A for arbitrary matrix A . See [4] for details.

Let

$$Z = X_1^* \Sigma^{-1/2}, \quad A = \Sigma^{-1/2} S_e \Sigma^{-1/2}, \quad \text{and} \quad M = M_1 \Sigma^{-1/2}. \quad (4)$$

Then Z and A are independent, $Z \sim N_{q \times p}(M, I_q \otimes I_p)$, $A \sim W_p(n, I_p)$ and the joint distribution of the non-zero eigenvalues of $S_e^{-1} S_h$ is equal to the joint distribution of the eigenvalues (l_1, \dots, l_q) of $ZA^{-1}Z'$.

The following lemma enables us to derive the valid asymptotic expansions of the null distributions of the three test statistics.

LEMMA 1. *Let*

$$B = ZZ' \quad \text{and} \quad W = B^{1/2}(ZA^{-1}Z')^{-1}B^{1/2}.$$

Then B and W are independently distributed as the noncentral and central Wishart distributions, $W_q(p, I_q, \Omega)$ and $W_q(m, I_q)$, respectively, where Ω is the noncentrality matrix given by

$$\Omega = MM' = M_1 \Sigma^{-1} M_1', \quad (5)$$

and $m = n - p + q$.

PROOF. It is obvious about the distribution of B . For given Z , the conditional distribution of $(ZA^{-1}Z')^{-1}$ is $W_q(n - p + q, (ZZ')^{-1})$ (see [4, Theorem 3.2.11]). Therefore the conditional distribution of W is $W_q(m, I_q)$, which proves the independence. \square

The non-zero eigenvalues of $ZA^{-1}Z'$ are equal to the ones of BW^{-1} . Both B and W can be represented as sums of independent random matrices of size $q \times q$. Therefore, the distribution of the smooth and symmetric function of l_1, \dots, l_q , which is the smooth function of BW^{-1} , has the valid asymptotic expansions (see [2]). In particular, the three statistic can be represented as

- (i) $\frac{|S_e|}{|S_e + S_h|} = \frac{|W|}{|W + B|},$
- (ii) $\text{tr}(S_h S_e^{-1}) = \text{tr}(BW^{-1}),$
- (iii) $\text{tr}(S_h(S_e + S_h)^{-1}) = \text{tr}(B(W + B)^{-1}).$

3. Null distributions

In this section we derive the asymptotic expansions of the test statistics under the null hypothesis. By virtue of lemma 1, we can obtain the asymptotic expansions following the usual methods: perturbation expansion of statistics,

expanding the characteristic functions, and inverting the resultant characteristic functions.

3.1. Perturbation expansion. In what follows we assume that

$$q : \text{fixed}, p \rightarrow \infty, m \rightarrow \infty, \text{ and both } \frac{p}{m} \text{ and } \frac{m}{p} \text{ are bounded} \quad (6)$$

instead of (3), since the convergence of the ratio p/n is not necessary.

Let U and V be defined by

$$\frac{1}{p}B = I_q + \frac{1}{\sqrt{p}}U, \quad \frac{1}{m}W = I_q + \frac{1}{\sqrt{m}}V. \quad (7)$$

Then U and V are asymptotically normal. Let

$$D = \sqrt{p} \left(\frac{m}{p} BW^{-1} - I_q \right). \quad (8)$$

Then $D = O_p(1)$ and the three statistics can be expanded in terms of D as follows:

$$\begin{aligned} \log \frac{|W|}{|W+B|} &= -\log |I_q + BW^{-1}| \\ &= -q \log(1+r) - \frac{r_2}{\sqrt{p}} \operatorname{tr}(D) + \frac{r_2^2}{2p} \operatorname{tr}(D^2) \\ &\quad - \frac{r_2^3}{3p\sqrt{p}} \operatorname{tr}(D^3) + O_p\left(\frac{1}{p^2}\right), \\ \operatorname{tr}(BW^{-1}) &= r \operatorname{tr}\left(I_q + \frac{1}{\sqrt{p}}D\right) = rq + \frac{r}{\sqrt{p}} \operatorname{tr}(D), \\ \operatorname{tr}[B(W+B)^{-1}] &= \operatorname{tr}[BW^{-1}(I_q + BW^{-1})^{-1}] \\ &= r_2q + r_2(1-r_2) \left\{ \frac{1}{\sqrt{p}} \operatorname{tr}(D) - \frac{r_2}{p} \operatorname{tr}(D^2) + \frac{r_2^2}{p\sqrt{p}} \operatorname{tr}(D^3) \right\} \\ &\quad + O_p\left(\frac{1}{p^2}\right), \end{aligned}$$

where

$$r = \frac{p}{m} \quad \text{and} \quad r_2 = \frac{r}{1+r}.$$

Let

$$\begin{aligned} T_{LR} &= -\sqrt{p} \left(1 + \frac{m}{p}\right) \left\{ \log \frac{|S_e|}{|S_e + S_h|} + q \log \left(1 + \frac{p}{m}\right) \right\}, \\ T_H &= \sqrt{p} \left\{ \frac{m}{p} \operatorname{tr}(S_h S_e^{-1}) - q \right\}, \\ T_{BNP} &= \sqrt{p} \left(1 + \frac{p}{m}\right) \left\{ \left(1 + \frac{m}{p}\right) \operatorname{tr}[S_h (S_e + S_h)^{-1}] - q \right\}. \end{aligned}$$

Then the expansion of T_G ($G = LR, H$, or BNP) is given by

$$T_G = \operatorname{tr}(D) + \frac{1}{\sqrt{p}} c_1 \operatorname{tr}(D^2) + \frac{1}{p} c_2 \operatorname{tr}(D^3) + O_p\left(\frac{1}{p\sqrt{p}}\right) \quad (9)$$

with $(c_1, c_2) = \left(-\frac{1}{2} \left(\frac{p}{m+p}\right), \frac{1}{3} \left(\frac{p}{m+p}\right)^2\right)$, $(0, 0)$ and $\left(\frac{-p}{m+p}, \left(\frac{p}{m+p}\right)^2\right)$ for $G = LR, H$ and BNP , respectively. Using (7) and (8), (9) can be expanded as

$$\begin{aligned} T_G &= \operatorname{tr}(U - \sqrt{r}V) + \frac{1}{\sqrt{p}} \{c_1 \operatorname{tr}[(U - \sqrt{r}V)^2] - \sqrt{r} \operatorname{tr}[(U - \sqrt{r}V)V]\} \\ &\quad + \frac{1}{p} \{c_2 \operatorname{tr}[(U - \sqrt{r}V)^3] - 2c_1 \sqrt{r} \operatorname{tr}[(U - \sqrt{r}V)^2 V] + r \operatorname{tr}[(U - \sqrt{r}V)V^2]\} \\ &\quad + O_p\left(\frac{1}{p\sqrt{p}}\right). \end{aligned}$$

3.2. The characteristic function. The characteristic function $C_T(t)$ of T_G is given by

$$C_T(t) = \mathbb{E}[\exp\{itT_G\}] = \mathbb{E}[\exp\{it(\operatorname{tr} U - \sqrt{r} \operatorname{tr} V)\}g(U, V)]$$

where

$$\begin{aligned} g(U, V) &= 1 + \frac{it}{\sqrt{p}} \{c_1 \operatorname{tr}[(U - \sqrt{r}V)^2] - \sqrt{r} \operatorname{tr}[(U - \sqrt{r}V)V]\} \\ &\quad + \frac{it}{p} \{c_2 \operatorname{tr}[(U - \sqrt{r}V)^3] - 2c_1 \sqrt{r} \operatorname{tr}[(U - \sqrt{r}V)^2 V] \\ &\quad \quad + r \operatorname{tr}[(U - \sqrt{r}V)V^2]\} \\ &\quad + \frac{(it)^2}{2p} \{c_1 \operatorname{tr}[(U - \sqrt{r}V)^2] - \sqrt{r} \operatorname{tr}[(U - \sqrt{r}V)V]\}^2 + O_p\left(\frac{1}{p\sqrt{p}}\right). \end{aligned}$$

Let Y be a $q \times m$ random matrix distributed as $N_{q \times m}(O, I_q \otimes I_m)$ and Z be the $q \times p$ random matrix given by (4). Then

$$U = \frac{1}{\sqrt{p}}ZZ' - \sqrt{p}I_q \quad \text{and} \quad V = \frac{1}{\sqrt{m}}YY' - \sqrt{m}I_q, \quad (10)$$

and the characteristic function is represented as

$$\begin{aligned} C_T(t) &= (2\pi)^{-q(p+m)/2} \iint \text{etr} \left\{ -\frac{1}{2}ZZ' + it \left(\frac{1}{\sqrt{p}}ZZ' - \sqrt{p}I_q \right) \right\} \\ &\quad \times \text{etr} \left\{ -\frac{1}{2}YY' - it\sqrt{r} \left(\frac{1}{\sqrt{m}}YY' - \sqrt{m}I_q \right) \right\} \\ &\quad \times g \left(\frac{1}{\sqrt{p}}ZZ' - \sqrt{p}I_q, \frac{1}{\sqrt{m}}YY' - \sqrt{m}I_q \right) (dZ)(dY). \end{aligned}$$

Let

$$a = 2it \quad \text{and} \quad b = -2it\sqrt{r}.$$

Considering transformations

$$Z = \left(1 - \frac{a}{\sqrt{p}}\right)^{-1/2} \tilde{Z} \quad \text{and} \quad Y = \left(1 - \frac{b}{\sqrt{m}}\right)^{-1/2} \tilde{Y}, \quad (11)$$

we obtain

$$\begin{aligned} C_T(t) &= (2\pi)^{-q(p+m)/2} \left(1 - \frac{a}{\sqrt{p}}\right)^{-qp/2} \left(1 - \frac{b}{\sqrt{m}}\right)^{-qm/2} \\ &\quad \times \iint \text{etr} \left\{ -\frac{1}{2}\tilde{Z}\tilde{Z}' - \frac{1}{2}\tilde{Y}\tilde{Y}' \right\} \\ &\quad \times g \left(\frac{1}{\sqrt{p}} \left\{1 - \frac{a}{\sqrt{p}}\right\}^{-1} \tilde{Z}\tilde{Z}' - \sqrt{p}I_q, \frac{1}{\sqrt{m}} \left\{1 - \frac{b}{\sqrt{m}}\right\}^{-1} \tilde{Y}\tilde{Y}' - \sqrt{m}I_q \right) \\ &\quad \times (d\tilde{Z})(d\tilde{Y}). \end{aligned}$$

Since

$$\begin{aligned} \frac{1}{\sqrt{p}} \left(1 - \frac{a}{\sqrt{p}}\right)^{-1} \tilde{Z}\tilde{Z}' - \sqrt{p}I_q &= \left(1 - \frac{a}{\sqrt{p}}\right)^{-1} \left\{ \left(\frac{1}{\sqrt{p}} \tilde{Z}\tilde{Z}' - \sqrt{p}I_q \right) + aI_q \right\}, \\ \frac{1}{\sqrt{m}} \left(1 - \frac{b}{\sqrt{m}}\right)^{-1} \tilde{Y}\tilde{Y}' - \sqrt{m}I_q &= \left(1 - \frac{b}{\sqrt{m}}\right)^{-1} \left\{ \left(\frac{1}{\sqrt{m}} \tilde{Y}\tilde{Y}' - \sqrt{m}I_q \right) + bI_q \right\}, \end{aligned}$$

we obtain

$$C_T(t) = \left(1 - \frac{a}{\sqrt{p}}\right)^{-qp/2} \left(1 - \frac{b}{\sqrt{m}}\right)^{-qm/2} E[g(\tilde{U}, \tilde{V})] \quad (12)$$

where

$$\tilde{U} = \left(1 - \frac{a}{\sqrt{p}}\right)^{-1} (U + aI_q), \quad \tilde{V} = \left(1 - \frac{b}{\sqrt{m}}\right)^{-1} (V + bI_q).$$

Here U and V has the same distributions given by (10).

The Jacobian of the transformation (11) is expanded as

$$\begin{aligned} & \left(1 - \frac{a}{\sqrt{p}}\right)^{-pq/2} \left(1 - \frac{b}{\sqrt{m}}\right)^{-mq/2} \\ &= \exp[(it)^2 q(1+r)] \\ & \times \left\{ 1 + \frac{(it)^3}{\sqrt{p}} \frac{4}{3} q(1-r^2) + \frac{(it)^4}{p} 2q(1+r^3) + \frac{(it)^6}{p} \frac{8}{9} q^2(1-r^2)^2 + \dots \right\}. \end{aligned}$$

Basic formulas of the normalized Wishart statistics U and V are given as follows:

$$E[U] = E[V] = O,$$

$$E[(\text{tr}(U))^2] = E[(\text{tr}(V))^2] = 2q, \quad E[\text{tr}(U^2)] = E[\text{tr}(V^2)] = q(q+1)$$

$$E[U^3] \approx E[V^3] \approx O,$$

$$E[(\text{tr}(U^2))^2] \approx E[(\text{tr}(V^2))^2] \approx q(q+1)(q^2 + q + 4),$$

$$E[(\text{tr}(UV))^2] = 2q(q+1),$$

where \approx means the difference between the left and right sides is $O\left(\frac{1}{\sqrt{p}}\right)$. By using the above formulas, the expectation in (12) can be expanded as

$$\begin{aligned} E[g(\tilde{U}, \tilde{V})] &= 1 + \frac{1}{\sqrt{p}} \{ ita_1 + (it)^3 a_3 \} \\ &+ \frac{1}{p} \{ (it)^2 a_2 + (it)^4 a_4 + (it)^6 a_6 \} + O\left(\frac{1}{p\sqrt{p}}\right), \end{aligned}$$

where

$$a_1 = \{c_1(1+r) + r\}q(q+1),$$

$$a_3 = 4\{c_1(1+r) + r\}(1+r)q,$$

$$\begin{aligned}
a_2 &= 6c_2(1+r)^2q(q+1) + \frac{1}{2}c_1^2(1+r)^2q(q+1)(q^2+q+4) \\
&\quad + c_1(1+r)q(q+1)\{4+r(q^2+q+12)\} + \frac{1}{2}q(q+1)r\{6+r(q^2+q+8)\}, \\
a_4 &= 8c_2(1+r)^3q + 4c_1^2(1+r)^3q(q^2+q+4) \\
&\quad + 8c_1q(1+r)^2\{2+r(q^2+q+4)\} + 4(1+r)r\{3+r(q^2+q+2)\}, \\
a_6 &= \frac{1}{2}a_3^2.
\end{aligned}$$

Multiplying by the Jacobian, the characteristic function is expanded as

$$\begin{aligned}
C_T(t) &= \exp[q(1+r)(it)^2] \\
&\quad \times \left[1 + \frac{1}{\sqrt{p}} \{itb_1 + (it)^3b_3\} + \frac{1}{p} \{(it)^2b_2 + (it)^4b_4 + (it)^6b_6\} \right] + O\left(\frac{1}{p\sqrt{p}}\right)
\end{aligned}$$

where

$$\begin{aligned}
b_1 &= a_1, & b_3 &= a_3 + \frac{4}{3}(1-r^2)q, \\
b_2 &= a_2, & b_4 &= \frac{4}{3}a_1(1-r^2)q + a_4 + 2(1+r^3)q, & b_6 &= \frac{1}{2}b_3^2.
\end{aligned}$$

3.3. Expansions of the null distributions. Inverting the characteristic functions we obtain the asymptotic expansion of the distribution function of T_G as in the following theorem.

THEOREM 1.

$$\begin{aligned}
\Pr\left\{\frac{1}{\sigma}T_G \leq z\right\} &= \Phi(z) - \phi(z) \left[\frac{1}{\sqrt{p}} \left\{ \frac{1}{\sigma}b_1 + \frac{1}{\sigma^3}b_3h_2(z) \right\} \right. \\
&\quad \left. + \frac{1}{p} \left\{ \frac{1}{\sigma^2}b_2h_1(z) + \frac{1}{\sigma^4}b_4h_3(z) + \frac{1}{\sigma^6}b_6h_5(z) \right\} \right] + O\left(\frac{1}{p\sqrt{p}}\right),
\end{aligned}$$

where

$$\sigma = \sqrt{2q(1+r)},$$

b_j 's are given above, and $h_j(z)$'s are the Hermite polynomials given by

$$\begin{aligned}
h_1(z) &= z, & h_2(z) &= z^2 - 1, & h_3(z) &= z^3 - 3z, \\
h_4(z) &= z^4 - 6z^2 + 3, & h_5(z) &= z^5 - 10z^3 + 15z.
\end{aligned}$$

By using the above theorem, we obtain the Cornish-Fisher expansion of the upper percent point of the distribution as in the following corollary.

COROLLARY 1. *Let z_α be the upper $100\alpha\%$ point of the standard normal distribution, and let*

$$z_{CF}(\alpha) = z_\alpha + \frac{1}{\sqrt{p}} \left\{ \frac{b_1}{\sigma} + (z_\alpha^2 - 1) \frac{b_3}{\sigma^3} \right\} + \frac{1}{p} \left\{ -\frac{1}{2} z_\alpha \left(\frac{b_1}{\sigma} \right)^2 + z_\alpha \frac{b_2}{\sigma^2} - z_\alpha (z_\alpha^2 - 3) \frac{b_1 b_3}{\sigma^4} - z_\alpha (2z_\alpha^2 - 5) \left(\frac{b_3}{\sigma^3} \right)^2 + z_\alpha (z_\alpha^2 - 3) \frac{b_4}{\sigma^4} \right\}.$$

Then

$$\Pr \left\{ \frac{1}{\sigma} T_G \leq z_{CF}(\alpha) \right\} = 1 - \alpha + O \left(\frac{1}{p\sqrt{p}} \right).$$

4. Numerical experiment

We study a comparison of the accuracy of asymptotic expansions in Theorem 1 with the other approximations based on asymptotic expansions when the sample size tend to infinity, while the dimension is a constant. For the latter asymptotic expansions, we used the approximations up to terms of the second order with respect to n^{-1} . The formulas of these asymptotic expansions can be found in [4, chapter 10].

Our comparison was done for the approximated upper 5 percent points given by Corollary 1, and the Cornish-Fisher expansions based on the large sample asymptotic expansions. The values of p , n and q were chosen as follows:

$$q; 2, 4, 8,$$

$$(p, n); (10, 40), (20, 40), (30, 40), (10, 80), (40, 80), (70, 80).$$

Table 1 shows the estimated upper 5 percent points based on Monte Carlo simulation, and the approximated upper 5 percent points base on our (New) method and the (Old) method.

Table 2 shows the actual error probabilities of the first kind by using the approximated percent points given by Table 1.

Figures 1, 2, and 3 show the actual error probabilities of the first kind of the two methods for the likelihood ratio, Lawley-Hotelling's test and Bartlett-Nanda-Pillai's test, respectively. In these figures the plotting symbols “□” and “X” correspond to our new method and the classical method, respectively.

Table 1. Cornish-Fisher expansions

q	n	p	Likelihood Ratio			Lawley-Hotelling			Bartlett-Nanda-Pillai		
			Simu	New	Old	Simu	New	Old	Simu	New	Old
2	40	10	0.960	0.890	0.888	1.280	1.16	1.14	0.74	0.70	0.53
2	40	20	1.91	1.90	1.861	3.44	3.329	3.31	1.20	1.199	-0.55
2	40	30	3.36	3.607	3.27	9.98	10.80	8.655	1.59	1.638	-5.539
2	80	10	0.457	0.42	0.416	0.52	0.471	0.471	0.40	0.37	0.349
2	80	40	1.720	1.76	1.71	2.84	2.888	3.877	1.14	1.15	-1.337
2	80	70	4.763	5.05	4.024	21.841	24.39	20.777	1.79	1.823	-16.988
4	40	10	1.627	1.531	1.530	2.212	2.005	1.85	1.261	1.200	0.20
4	40	20	3.288	3.340	3.258	5.963	5.73	6.875	2.130	2.150	-7.95
4	40	30	5.98	6.266	5.65	18.32	17.345	24.59	2.96	3.026	-38.47
4	80	10	0.82	0.73	0.73	0.950	0.829	0.81	0.72	0.65	0.52
4	80	40	3.216	3.219	3.12	5.279	5.222	11.323	2.144	2.15	-12.31
4	80	70	8.560	9.061	7.118	40.61	39.37	84.792	3.439	3.508	-107.53
8	40	10	2.913	2.63	2.649	4.190	3.495	0.765	2.205	2.03	-2.730
8	40	20	5.971	5.783	5.71	11.28	9.77	5.485	3.83	3.762	-44.962
8	40	30	10.507	10.575	9.81	34.11	26.48	48.20	5.330	5.411	-193.69
8	80	10	1.467	1.295	1.30	1.741	1.50	1.157	1.262	1.13	0.532
8	80	40	5.929	5.89	5.729	10.034	9.530	28.59	3.98	3.969	-65.05
8	80	70	15.520	15.91	12.88	73.602	60.347	294.221	6.520	6.62	-519.40

Table 2. Actual error probabilities of the first kind when the nominal level is 0.05.

q	n	p	Likelihood Ratio		Lawley-Hotelling		Bartlett-Nanda-Pillai	
			New	Old	New	Old	New	Old
2	40	10	0.089	0.089	0.092	0.103	0.092	0.453
2	40	20	0.053	0.065	0.061	0.064	0.050	1.000
2	40	30	0.017	0.065	0.032	0.089	0.019	1.000
2	80	10	0.109	0.109	0.112	0.111	0.110	0.155
2	80	40	0.042	0.056	0.043	0.003	0.039	1.000
2	80	70	0.014	0.327	0.032	0.064	0.013	1.000
4	40	10	0.107	0.107	0.112	0.219	0.111	1.000
4	40	20	0.041	0.057	0.075	0.013	0.036	1.000
4	40	30	0.024	0.109	0.068	0.007	0.022	1.000
4	80	10	0.147	0.146	0.149	0.166	0.154	0.586
4	80	40	0.048	0.084	0.055	0.000	0.047	1.000
4	80	70	0.011	0.644	0.061	0.000	0.011	1.000
8	40	10	0.184	0.172	0.222	1.000	0.187	1.000
8	40	20	0.094	0.119	0.184	0.986	0.086	1.000
8	40	30	0.045	0.178	0.233	0.002	0.017	1.000
8	80	10	0.235	0.223	0.239	0.814	0.247	1.000
8	80	40	0.063	0.136	0.110	0.000	0.059	1.000
8	80	70	0.017	0.890	0.197	0.000	0.010	1.000

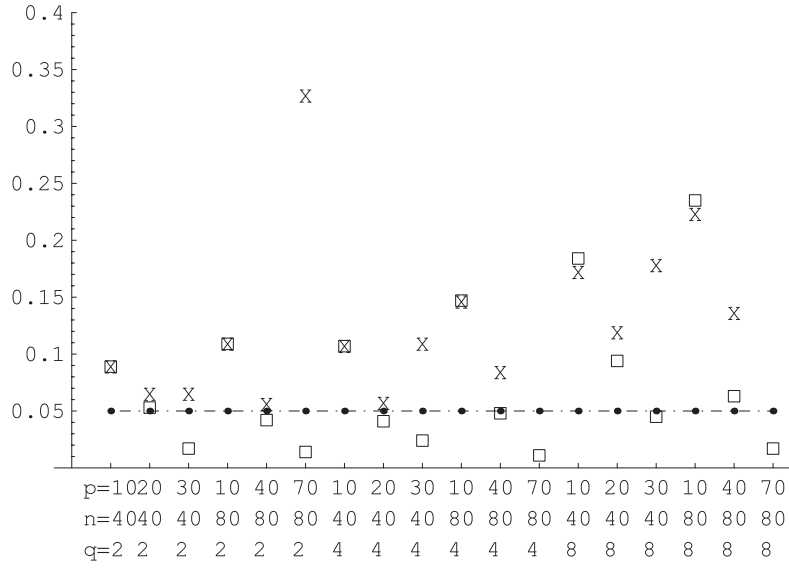


Fig. 1. Actual error probabilities of the first kind for LR test

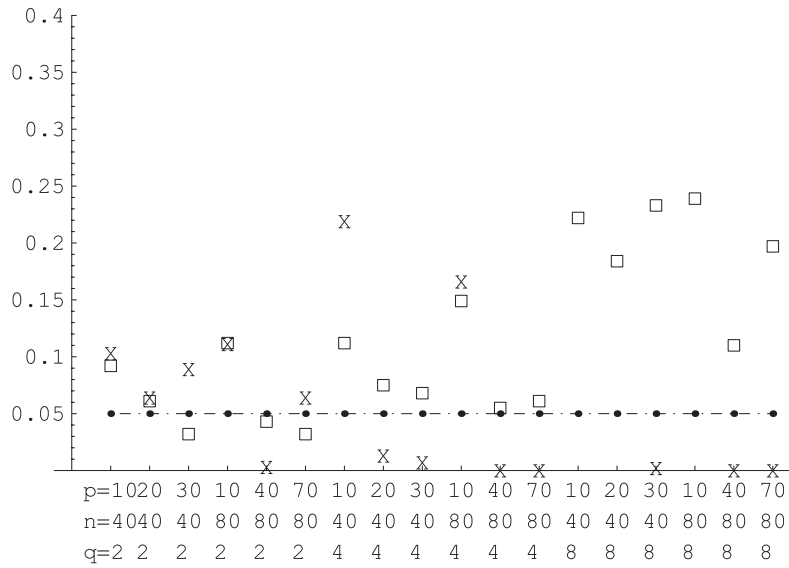


Fig. 2. Actual error probabilities of the first kind for Lawley-Hotelling's test

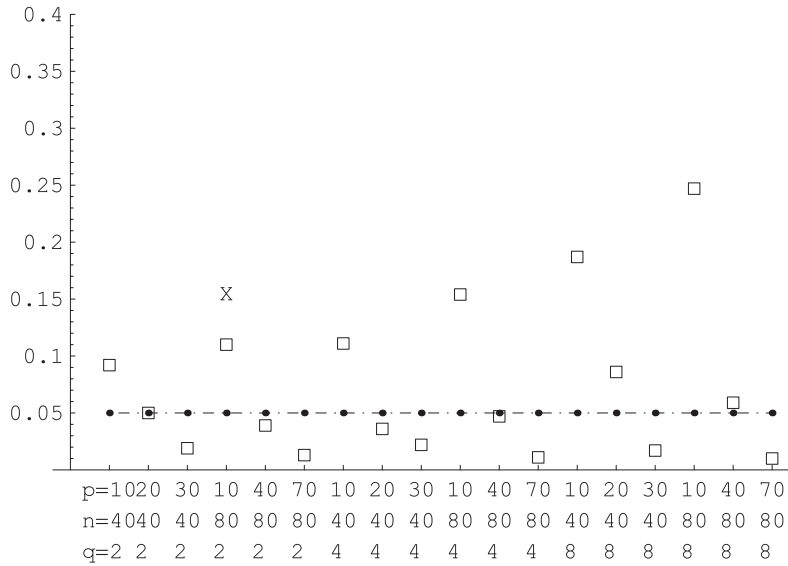


Fig. 3. Actual error probabilities of the first kind for Bartlett-Nanda-Pillai's test

It can be seen that our new method performs better than the old one in most of the cases. Three figures reveal same tendency of the accuracy of our method:

- If the ratio p/n is small, Cornish-Fisher expansion of the percent point is under estimate. (Actual error probabilities are larger than the nominal level.)
- If the ratio p/n is equal to $\frac{1}{2}$, our method performs very good.
- If the ratio p/n is large, Cornish-Fisher expansion gives over estimate approximation.

It is seen that when q becomes large, the accuracy becomes bad as we can expect. This suggests that asymptotic expansion formulas are needed when q , as well as p and n , tends to infinity. Wakaki [8] derived an Edgeworth expansion for Wilks' lambda distribution in such case. Ulyanov, Wakaki and Fujikoshi [9] obtained a Berry–Esseen type bound for Wilks' lambda distribution. The problems of deriving asymptotic expansion formulas and their error bounds for the other test statistics in the case that p , q and n become large are left for future. The detailed discussion of asymptotic expansions for MANOVA tests for large sample and high-dimensional frameworks could be found in [10, Chapter 6].

References

- [1] T. W. Anderson, *An introduction to multivariate analysis* (2nd ed.), John Wiley & Sons, New York, 1984.
- [2] R. N. Bhattacharya and J. K. Ghosh, On the validity of formal Edgeworth expansion, *Ann. Statist.* **6** (1978), 434–451.
- [3] R. N. Bhattacharya and R. Ranga Rao, *Normal approximation and asymptotic expansions*, John Wiley & Sons, New York, 1976.
- [4] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, 1982.
- [5] M. Siotani, T. Hayakawa and Y. Fujikoshi, *Modern multivariate statistical analysis: a graduate course and handbook*, American Sciences Press, Columbus, Ohio, 1985.
- [6] T. Tonda and Y. Fujikoshi, Asymptotic expansion of the null distribution of LR statistic for multivariate linear hypothesis when the dimension is large, *Comm. Statist.-Theory Meth.*, **33** (2004), 1205–1220.
- [7] H. Wakaki, Y. Fujikoshi and V. V. Ulyanov, Asymptotic expansions of the distributions of MANOVA test statistics when the dimension is large. Hiroshima Statistical Research Group Technical Report, **TR02-10** (2002), 1–10.
- [8] H. Wakaki, Edgeworth expansion of Wilks' lambda statistic. *Journal of Multivariate Analysis*, **97** (2006), 1958–1964.
- [9] V. V. Ulyanov, H. Wakaki and Y. Fujikoshi, Berry–Esseen bound for high dimensional asymptotic approximation of Wilks' Lambda distribution. *Statistics & Probability Letters*, **76**, 1191–1200.
- [10] Y. Fujikoshi, V. V. Ulyanov, and R. Shimizu, *Multivariate statistics: High-dimensional and large-sample approximations*, John Wiley and Sons, New York, 2010.

Hirofumi Wakaki
Department of Mathematics
Graduate School of Science
Hiroshima University
Higashi-Hiroshima 739-8526, Japan
E-mail: wakaki@math.sci.hiroshima-u.ac.jp

Yasunori Fujikoshi
Department of Mathematics
Graduate School of Science
Hiroshima University
Higashi-Hiroshima 739-8526, Japan

Vladimir V. Ulyanov
Department of Mathematical Statistics
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
Moscow, 119991, Russia
and National Research University Higher School of Economics (HSE)
Moscow, 101000, Russia