

Edgeworth expansion of Wilks' lambda statistic

Hirofumi Wakaki

Department of Mathematics, Hiroshima University, Kagamiyama 1-3-1, Higashi-Hiroshima 739-8526, Japan

Received 19 May 2005

Abstract

An asymptotic expansion of the null distribution of the Wilks' lambda statistic is derived when some of the parameters are large. Cornish–Fisher expansions of the upper percent points are also obtained. A monotone transformation which reduces the third and the fourth order cumulants is also derived. 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.

© 2006 Published by Elsevier Inc.

AMS 1991 subject classification: primary 62H10secondary 62E20

Keywords: MANOVA; High dimension; Likelihood ratio; Monotone transformation; Edgeworth expansion

1. Introduction

Consider one-way MANOVA in which the equality of $q + 1$ mean vectors is tested. Let V_b and V_w be the between and the within matrices of sum of squares and products, respectively. Assume normal populations with common covariance matrix Σ . Then the likelihood ratio test is based on the statistic

$$\Lambda = \det\{V_w(V_b + V_w)^{-1}\}, \quad (1.1)$$

and

$$V_w \sim W_p(n, \Sigma), V_b \sim W_p(q, \Sigma)$$

under the null hypothesis, where n depends on the sample sizes and is assumed to be larger than p .

E-mail address: wakaki@math.sci.hiroshima-u.ac.jp.

The criterion (1.1) was introduced by Wilks [9] and is called Wilks' lambda-criterion. We denote the null distribution of Λ as $\Lambda_p(q, n)$. It is known that Λ is distributed as the product of some independent beta random variables under the null hypothesis. So the exact null distribution function is complicated and hard to use. When n is large relative to p and q , an asymptotic expansion based on the formula of Box [2] can be used. Let

$$M = n - \frac{1}{2}(p + q + 1), \quad \gamma = \frac{1}{48}pq(p^2 + q^2 - 5) \quad (1.2)$$

and $f = pq$. Then

$$\Pr\{-M \log \Lambda \leq x\} = G_f(x) + \frac{\gamma}{M^2}\{G_{f+4}(x) - G_f(x)\} + O(M^{-4}). \quad (1.3)$$

Distributional results for Λ and other test statistics can be found in Muirhead [5, Chapter 10], Anderson [1, Chapter 8] and Siotani et al. [6, Chapter 6]. Numerical experiments show that the approximation based on (1.3) is poor when p is large relative to n .

Tonda and Fujikoshi [7] derived an asymptotic expansion formula when q is fixed, $n \rightarrow \infty$, $p \rightarrow \infty$ with $p/n \rightarrow c \in (0, 1)$ by using the fact that $\Lambda_p(q, n) = \Lambda_q(p, n - p + q)$. Wakaki et al. [8] derived asymptotic expansion formulas for three test statistics, Λ , Lawley–Hotelling's trace and Bartlett–Nanda–Pillai trace in the same setup with Tonda and Fujikoshi [7]. Since the leading term of the expansion is normal distribution we refer there formula as "normal type", while the formulas like (1.3) are referred as "chi-square type". Their numerical experiments show that when p is large the normal type approximations perform better than the chi-square type. However, when q is large both normal type and chi-square approximations has poor performance. In this paper we derive an asymptotic expansion of the null distribution of Λ when q is also large.

2. Edgeworth expansion for $\log \Lambda$

From

$$\Lambda_q(p, n - p + q) = \prod_{j=1}^q \text{beta} \left(\frac{n - p + j}{2}, \frac{p}{2} \right),$$

the h th order moment of Λ is given by

$$E(\Lambda^h) = \frac{\Gamma_q \left[\frac{n-p+q}{2} + h \right] \Gamma_q \left[\frac{n+q}{2} \right]}{\Gamma_q \left[\frac{n+q}{2} + h \right] \Gamma_q \left[\frac{n-p+q}{2} \right]}, \quad (2.1)$$

where

$$\Gamma_q(a) = \pi^{q(q-1)/4} \prod_{j=1}^q \Gamma \left[a - \frac{j-1}{2} \right], \quad \mathcal{R}(a) > \frac{q-1}{2}.$$

Hence the cumulants generating function of $-\log \Lambda$ can be expanded as

$$\begin{aligned} \log E[\exp(-it \log \Lambda)] &= \sum_{j=1}^q \log \left\{ \frac{\Gamma \left[\frac{n-p+j}{2} - it \right] \Gamma \left[\frac{n+j}{2} \right]}{\Gamma \left[\frac{n+j}{2} - it \right] \Gamma \left[\frac{n-p+j}{2} \right]} \right\} \\ &= \sum_{s=1}^{\infty} \frac{(-1)^s}{s!} \left\{ \psi_q^{(s-1)} \left(\frac{n-p+q}{2} \right) - \psi_q^{(s-1)} \left(\frac{n+q}{2} \right) \right\}, \quad (2.2) \end{aligned}$$

where

$$\psi_q^{(s)}(a) = \sum_{j=1}^q \psi^{(s)}\left(a - \frac{j-1}{2}\right) \quad (s = 0, 1, \dots; a > 0),$$

and $\psi^{(s)}(a)$ is the polygamma function defined as

$$\psi^{(s)}(a) = \left(\frac{d}{da}\right)^s \log \Gamma(a) = \begin{cases} -C + \sum_{k=0}^{\infty} \left(\frac{1}{1+k} - \frac{1}{k+a}\right) & (s = 0) \\ \sum_{k=0}^{\infty} \frac{(-1)^{s+1} s!}{(k+a)^{s+1}} & (s = 1, 2, \dots) \end{cases}. \quad (2.3)$$

Here C is the Euler's constant.

The expansion (2.2) and the series representation in (2.3) give the s th order cumulant $\kappa^{(s)}$ of $-\log \Lambda$ as

$$\begin{aligned} \kappa^{(s)} &= (-1)^s \left\{ \psi_q^{(s-1)}\left(\frac{n-p+q}{2}\right) - \psi_q^{(s-1)}\left(\frac{n+q}{2}\right) \right\} \\ &= \sum_{j=1}^q \sum_{k=0}^{\infty} f_s(j+2k), \end{aligned} \quad (2.4)$$

where

$$f_s(x) = (s-1)! \left\{ \left(\frac{2}{n-p+x}\right)^s - \left(\frac{2}{n+x}\right)^s \right\}.$$

Since f_s is a decreasing and convex function,

$$\begin{aligned} &\int_j^{j+1} \left\{ \int_k^{k+1} f_s(x+2y) dy \right\} dx \\ &< f_s(j+2k) < \int_{j-1/2}^{j+1/2} \left\{ \int_{k-1/2}^{k+1/2} f_s(x+2y) dy \right\} dx. \end{aligned}$$

Therefore, we obtain upper and lower bounds for $\kappa^{(s)}$ as

$$\begin{aligned} b^{(s)}(n-p+1, p, q) &< \kappa^{(s)} < b^{(s)}\left(n-p-\frac{1}{2}, p, q\right), \\ b^{(2)}(l, p, q) &= 2 \log \left\{ 1 + \frac{pq}{(l+p+q)l} \right\}, \\ b^{(s)}(l, p, q) &= \frac{2^{s-1}(s-3)!}{l^{s-2}} \\ &\quad \times \left\{ 1 - \left(\frac{l}{l+q}\right)^{s-2} - \left(\frac{l}{l+p}\right)^{s-2} + \left(\frac{l}{l+p+q}\right)^{s-2} \right\} \end{aligned} \quad (2.5)$$

for $s \geq 3$. We note that $b^{(s)}(l, p, q) > 0$ if l, p and q are positive.

Table 1

The orders of the standardized cumulants

		$\tilde{\kappa}^{(s)}$	ε
(i)	$l \sim p \sim q \gg 1, p \sim l \gg q \sim 1,$ $p \gg l \gg q \sim 1, p \gg l \sim q \gg 1$	$\sim \varepsilon^{-(s-2)}$	$l^{-1/2}$
(ii)	$l \gg p \gg q \gg 1, l \gg p \sim q \gg 1$	$\sim \varepsilon^{-(s-2)}$	$(pq)^{-1/2}$
(iii)	$l \gg p \gg q \sim 1$	$\sim \varepsilon^{-(s-2)}$	$p^{-1/2}$
(iv)	$p \sim l \gg q \gg 1, p \gg l \gg q \gg 1$	$\sim \varepsilon^{-(s-2)}$	$(lq)^{-1/2}$
(v)	$p \gg q \gg l \gg 1, p \sim q \gg l \gg 1$	$\sim (\log \frac{q}{l})^{-1} \varepsilon^{-(s-2)}$	$(l^2 \log \frac{q}{l})^{-1/2}$
(vi)	$p \gg q \gg l \sim 1, p \sim q \gg l \sim 1$	$\sim (\log q)^{-1} \varepsilon^{-(s-2)}$	$(\log q)^{-1/2}$
(vii)	$p \gg q \sim l \sim 1, l \gg p \sim q \sim 1$	~ 1	

Assume at least one of $l = n - p$, p , q tends to infinity. We use the following notations:

$$a \sim b \Leftrightarrow \frac{a}{b} \text{ and } \frac{b}{a} \text{ are bounded,}$$

$$a \gg b \Leftrightarrow \frac{b}{a} \rightarrow 0.$$

Let the standardized cumulant $\tilde{\kappa}^{(s)}$ be defined as

$$\tilde{\kappa}^{(s)} = \kappa^{(s)} (\kappa^{(2)})^{-s/2} \quad (s \geq 3). \quad (2.6)$$

Then using (2.5) the order of the standardized cumulants can be calculated in several cases. The results are summarized in Table 1.

Note that the orders in the cases that $q \gg p$ are obtained by using the fact that $\Lambda_p(q, n) = \Lambda_q(p, n - p + q)$. Table 1 shows that the standardized cumulants can be expressed as $\tilde{\kappa}^{(s)} = \varepsilon^{s-2} \tau^{(s)}$ with $\varepsilon \rightarrow 0$ and $\tau^{(s)} = O(1)$ in all cases except (vii). Only chi-square type approximation can be used in the second case of (vii). Neglecting the cumulants $\tilde{\kappa}^{(s)}$ ($s \geq 5$), we obtain the Edgeworth expansion of the Wilks' distribution as in the following theorem.

Theorem 1. In each case of (i)–(vi) in Table 1, the null distribution of $T = \frac{-\log \Lambda - \kappa^{(1)}}{(\kappa^{(2)})^{1/2}}$ can be expanded as

$$\Pr \{T \leq z\} = \Phi(z) - \phi(z) \left\{ \frac{1}{6} \tilde{\kappa}^{(3)} h_2(z) + \frac{1}{24} \tilde{\kappa}^{(4)} h_3(z) + \frac{1}{72} (\tilde{\kappa}^{(3)})^2 h_5(z) \right\} + O(\tilde{\kappa}^{(5)}), \quad (2.7)$$

where $\kappa^{(1)}$, $\kappa^{(2)}$ and $\tilde{\kappa}^{(s)}$ ($s \geq 3$) are given by (2.4) and (2.6), and $h_s(z)$'s are the Hermite polynomials given by

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

Note that the order of the remainder term in each case of (i) to (iv) is $O(\varepsilon^3)$, while the order in each case of (v) and (vi) is $o(\varepsilon^3)$.

We note that if p or q are even number, $\kappa^{(s)}$'s are simplified as

$$\kappa^{(s)} = \begin{cases} (s-1)! \sum_{j=0}^{q/2-1} \sum_{k=1}^p \left(\frac{2}{n-p+2j+k} \right)^s & (q : \text{even}), \\ (s-1)! \sum_{j=1}^q \sum_{k=0}^{p/2-1} \left(\frac{2}{n-p+j+2k} \right)^s & (p : \text{even}). \end{cases}$$

By using the above theorem, we obtain the Cornish–Fisher expansion of the upper $100\alpha\%$ point of the distribution as in the following theory.

Corollary 2. Let z_α be the upper $100(1-\alpha)\%$ point of the standard normal distribution, and let

$$z_{CF}(z) = z + \left\{ (z^2 - 1) \frac{\tilde{\kappa}^{(3)}}{6} - z(2z^2 - 5) \left(\frac{\tilde{\kappa}^{(3)}}{6} \right)^2 + z(z^2 - 3) \frac{\tilde{\kappa}^{(4)}}{24} \right\}. \quad (2.8)$$

Then

$$\Pr\{T \leq z_{CF}(z_\alpha)\} = 1 - \alpha + O(\tilde{\kappa}^{(5)}).$$

3. Monotone transformation

In this section we focus to the case that the order of the remainder term in (2.7) is $O(\varepsilon^3)$ for simplicity, although the derived transformation can be also applied to the cases (v) and (vi) in Table 1.

Expanding the inverse transformation of $z_{CF}(z)$, we obtain a cubic transformation which makes the statistic more close to normal distribution. Let

$$g(t) = c\{t + \varepsilon a(t^2 - 1) + \varepsilon^2 b t^3\},$$

$$a = -\frac{1}{6}\tau^{(3)}, \quad b = \frac{1}{9}(\tau^{(3)})^2 - \frac{1}{24}\tau^{(4)}, \quad c = \frac{1 + \varepsilon^2 \frac{1}{8}\tau^{(4)}}{1 + \varepsilon^2 \frac{7}{36}(\tau^{(3)})^2}. \quad (3.1)$$

Then

$$\Pr\{g(T) \leq z_\alpha\} = 1 - \alpha + O(\varepsilon^3).$$

However, g is not monotone and $g(T)$ leads to the different test with Λ . Hall [3] suggested the following monotone transformation. Let $g_u(t)$ be defined as

$$g_u(t) = t + \varepsilon u t^2 + \frac{1}{3}\varepsilon^2 u^2 t^3. \quad (3.2)$$

Then g_u is monotone, and since $g_a(t) - a\varepsilon = g(t) + O(\varepsilon^2)$,

$$\Pr\{g_a(T) - a\varepsilon \leq z_\alpha\} = 1 - \alpha + O(\varepsilon^2).$$

One advantage of Hall's transformation is that the inverse transformation has the simple form:

$$g_u^{-1}(t) = \frac{3t}{1 + (1 + 3\varepsilon u t)^{1/3} + (1 + 3\varepsilon u t)^{2/3}}. \quad (3.3)$$

We need a monotone transformation of which the Taylor expansion agrees with g in (3.1) up to the order $O(\varepsilon^3)$ for our purpose. The composite function of g_u^{-1} and g_v can be expanded as

$$g_u^{-1}\{g_v(t)\} = t + \varepsilon(v - u)t^2 + \frac{1}{3}\varepsilon^2(u - v)(5u - v)t^3 + O(\varepsilon^3).$$

By equating the coefficients of t^2 and t^3 with those of $g(t)/c$, we obtain

$$\varepsilon u = \frac{11\tilde{\kappa}^{(3)}}{24} - \frac{3\tilde{\kappa}^{(4)}}{16\tilde{\kappa}^{(3)}}, \quad \varepsilon v = \frac{7\tilde{\kappa}^{(3)}}{24} - \frac{3\tilde{\kappa}^{(4)}}{16\tilde{\kappa}^{(3)}}. \quad (3.4)$$

Then we obtain the following theorem.

Theorem 3. Let u, v be defined by (3.4), c be defined by (3.1), and g_u be defined by (3.2). Then

$$\Pr \left\{ c g_u^{-1}\{g_v(T)\} - \varepsilon a \leq z_\alpha \right\} = 1 - \alpha + O(\varepsilon^3). \quad (3.5)$$

By inverting the transformation, the monotone version of the Cornish–Fisher expansion of the $100(1 - \alpha)\%$ point is given by

$$z_{\text{mono}}(z_\alpha) = g_v^{-1} \left\{ g_u \left(\frac{z_\alpha + \varepsilon a}{c} \right) \right\}. \quad (3.6)$$

4. Numerical experiment

We study a comparison of the accuracy of the asymptotic expansion formula in Wakaki et al. [8] for the likelihood ratio test and chi-square type approximation formula (1.3) with our new formulas. The comparison was done for the approximated upper 5% points given by the chi-square type Cornish–Fisher expansion, the Cornish–Fisher expansion given in Wakaki et al. [8], z_{CF} in (2.8), and the monotone version (3.6), these formulas are referred in the following table as CHI, WFU, POLY, and MONO.

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).

The number of iteration was 10,000,000. We used a pseudo-random number generator named *Mersenne Twister* which provides a period of $2^{19937} - 1$ and 623-dimensional equidistribution, and is sufficient for our purpose (see Matsumoto and Nishimura [4]).

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

Our new method gives considerably correct approximations while the normal type approximation based on perturbation by Wakaki et al. [8] gets worse when q becomes large. The chi-square type approximation performs good when p is smaller than or equal to the half of n . But chi-square type approximations are poor when p is greater than the half. In the selected parameters we could not find any difference between the usual Cornish–Fisher expansion and the monotone version.

Table 2

The actual error probabilities of the first kind

<i>n</i>	<i>p</i>	<i>q</i>	POLY	MONO	WFO	CHI
40	10	2	4.992	4.999	5.104	5.003
40	20	2	5.000	5.005	5.186	5.164
40	30	2	4.986	4.995	5.428	7.303
80	10	2	4.998	5.004	5.034	5.004
80	40	2	5.001	5.002	5.070	5.215
80	70	2	4.983	4.991	5.411	21.717
40	10	4	5.009	5.012	5.340	5.019
40	20	4	4.985	4.986	5.608	5.187
40	30	4	5.000	5.003	6.555	7.832
80	10	4	4.998	5.000	5.089	5.000
80	40	4	5.006	5.007	5.245	5.278
80	70	4	4.989	4.992	6.507	28.071
40	10	8	5.001	5.002	6.240	5.020
40	20	8	5.005	5.006	7.562	5.299
40	30	8	4.995	4.996	11.578	8.536
80	10	8	5.015	5.016	5.333	5.017
80	40	8	4.998	4.999	5.996	5.354
80	70	8	5.002	5.003	11.464	35.171

Acknowledgments

The author would like to thank the referees and the managing editor for carefully reading the paper and for their help in improving the paper. This study was supported in part by a Grant-in-Aid for Scientific Research (B) (no. 1530092).

References

- [1] T.W. Anderson, *An Introduction to Multivariate Analysis*, second ed., Wiley, New York, 1984.
- [2] G.E.P. Box, A general distribution theory for a class of likelihood criteria, *Biometrika* 36 (1949) 317–346.
- [3] P. Hall, On the Removal of Skewness by Transformation, *J. R. Statist. Soc. B* 54 (1992) 221–228.
- [4] M. Matsumoto, T. Nishimura, Mersenne Twister: a 623-dimensionally equidistributed uniform pseudorandom number generator, *ACM Trans. Modeling Comput. Simul.* 8 (1998) 3–30.
- [5] R.J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.
- [6] M. Siotani, T. Hayakawa, Y. Fujikoshi, *Modern multivariate statistical analysis: a graduate course and handbook*, American Sciences Press, Columbus, Ohio, 1985.
- [7] T. Tonda, 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.
- [8] H. Wakaki, Y. Fujikoshi, V. Ulyanov, Asymptotic expansions of the distributions of MANOVA test statistics when the dimension is large, *Hiroshima Statistical Group Technical Report* 10, 2002, pp. 1–10, submitted for publication.
- [9] S.S. Wilks, Certain generalizations in the analysis of variance, *Biometrika* 24 (1932) 471–494.