JOURNAL OF MULTIVARIATE ANALYSIS 8, 468–478 (1978)

# Inference in Canonical Correlation Analysis

WILLIAM J. GLYNN*

*Harvard University*

AND

ROBB J. MUIRHEAD†

*The University of Michigan*

*Communicated by P. R. Krishnaiah*

The asymptotic behavior, for large sample size, is given for the distribution of the canonical correlation coefficients. The result is used to examine the Bartlett–Lawley test that the residual population canonical correlation coefficients are zero. A marginal likelihood function for the population coefficients is obtained and the maximum marginal likelihood estimates are shown to provide a bias correction.

## 1. INTRODUCTION

Let $r_1, ..., r_p$ be the sample canonical correlation coefficients between variates $y_1, ..., y_p$ and $x_1, ..., x_q$ $(p \leqslant q)$ calculated from a sample of size $N = n + 1$ observations from a $(p + q)$-variate normal distribution. The exact joint density function of $r_1^2, ..., r_p^2$ is (see Constantine [5], James [9])

$$\prod_{i=1}^{p} (1 - \rho_i^2)^{\frac{1}{2}n} \, {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2)$$

$$\times\, k_1 \prod_{i=1}^{p} (r_i^2)^{\frac{1}{2}(q-p-1)} (1 - r_i^2)^{\frac{1}{2}(n-q-p-1)} \prod_{i<j}^{p} (r_i^2 - r_j^2) \qquad (1.1)$$

$$(1 > r_1^2 > r_2^2 > \cdots > r_p^2 > 0),$$

Received November 30, 1976.

AMS 1970 subject classification: Primary 62H10.

Key words and phrases: canonical correlations, asymptotic distributions.

* Research supported in part by the National Science Foundation under Contract MCS75-01493.

† Research supported by the National Science Foundation under Contract MCS76-14876 and MCS78-18583.



0047-259X/78/0083-0468$02.00/0

where $1 \geqslant \rho_1 \geqslant \rho_2 \geqslant \cdots \geqslant \rho_p \geqslant 0$ are the population canonical correlation coefficients, $R = \text{diag}(r_1, ..., r_p)$, $P = \text{diag}(\rho_1, ..., \rho_p)$,

$$k_1 = \Gamma_p(\tfrac{1}{2}n)\pi^{\frac{1}{2}p^2}/[\Gamma_p(\tfrac{1}{2}(n-q))\,\Gamma_p(\tfrac{1}{2}q)\,\Gamma_p(\tfrac{1}{2}p)], \tag{1.2}$$

and ${}_2F_1^{(p)}$ is a hypergeometric function with the matrices $P^2$ and $R^2$ as arguments. The distribution of $r_1^2, ..., r_p^2$ depends only on $\rho_1, ..., \rho_p$ and hence that part of the distribution involving $\rho_1, ..., \rho_p$ can be regarded as a marginal likelihood. From (1.1) we see that the marginal likelihood function is

$$\prod_{i=1}^{p} (1-\rho_i^2)^{\frac{1}{2}n}\, {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2). \tag{1.3}$$

In Section 2 we derive an asymptotic representation for the ${}_2F_1^{(p)}$ function, and hence for the distribution (1.1) and marginal likelihood (1.3), for large sample size $n$. This is done by expressing ${}_2F_1^{(p)}$ as a complicated multiple integral and using a multivariate extension of Laplace's method for integrals to obtain its asymptotic behavior. In Section 3 the asymptotic distribution is used to examine the Bartlett–Lawley test of the null hypothesis that the last $p-k$ population canonical correlation coefficients are zero. Maximum marginal likelihood estimates of certain transformed population coefficients are also obtained and are shown to provide a bias correction.

## 2. Asymptotic Distributions

Before deriving the asymptotic behavior of the ${}_2F_1^{(p)}$ function in (1.1) we first note Hsu's extension [8] of Laplace's method for obtaining the asymptotic behavior of integrals. If the function $f(x) = f(x_1, ..., x_m)$ has an absolute maximum at an interior point $\xi$ of a domain $\mathscr{S}$ in real $m$-dimensional space, then under suitable conditions, as $n \to \infty$

$$\int_{\mathscr{S}} f(x)^n \varphi(x)\,dx \sim (2\pi/n)^{\frac{1}{2}m} f(\xi)^n \varphi(\xi)\,\Delta(\xi)^{-\frac{1}{2}}, \tag{2.1}$$

where $a \sim b$ means that $\lim_{n\to\infty} a/b = 1$ and $\Delta(x) = \det(-\partial^2 \log f/\partial x_i\,\partial x_j)$.

We begin by looking at the ${}_2F_1$ function with one $k \times k$ matrix $T^2$ as argument (see [9]). Without loss of generality $T$ can be assumed diagonal, $T = \text{diag}(t_1, ..., t_k)$, and we will assume that the roots are distinct with $1 > t_1 > t_2 > \cdots > t_k > 0$. The integrals involved in the subsequent development can be found in James [9] and Herz [7].

THEOREM 1. *As* $n \to \infty$

$$ {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; T^2) \sim c_1 \prod_{i=1}^{k} t_i^{\frac{1}{2}(k-q)}(1-t_i)^{-n+\frac{1}{2}q} \prod_{i<j}^{k} (t_i+t_j)^{-\frac{1}{2}}, \tag{2.2} $$

*where*

$$ c_1 = (\tfrac{1}{2}n)^{-\frac{1}{2}k(q-\frac{1}{2}k-\frac{1}{2})}\, \Gamma_k(\tfrac{1}{2}q)\pi^{-\frac{1}{4}k(k+1)}2^{-k}[1+O(n^{-1})]. $$

*Proof.* The idea here is to express ${}_2F_1$ as a multiple integral to which Hsu's result (2.1) can be applied. We can write

$$ \begin{aligned} {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; T^2) = c_2 \int_{O(k)} \int_{D_U} \int_{O(k)} \int_{D_V} \int_{V(k,q)} & \exp\{-\tfrac{1}{2}n \operatorname{tr}(U^2+V^2)\} \\ & \times |UV|^{n-k} \exp\{n \operatorname{tr}([TH_1UH_1'H_2VH_2' \vdots O]Q_1)\} \\ & \times \prod_{i<j}^{k} (u_i^2-u_j^2)(v_i^2-v_j^2)(dQ_1)(dV)(dH_2)(dU)(dH_1), \end{aligned} \tag{2.3} $$

where

$$ c_2 = (\tfrac{1}{2}n)^{nk}2^{-k}\pi^{\frac{1}{2}k(k-2q)}\, \Gamma_q(\tfrac{1}{2}q)/[\Gamma_k(n/2)^2\, \Gamma_{q-k}(\tfrac{1}{2}(q-k))], $$

$O(k)$ is the group of $k \times k$ orthogonal matrices, $(dH_i)$ $(i = 1, 2)$ is the unnormalized measure on $O(k)$, so that the volume of $O(k)$ is $2^k\pi^{(1/2)k^2}/\Gamma_k(\frac{1}{2}k)$, $U = \operatorname{diag}(u_1, ..., u_k)$, $V = \operatorname{diag}(v_1, ..., v_k)$, $D_u = \{(u_1, ..., u_k); u_1 > u_2 > \cdots > u_k > 0\}$, and $V(k, q)$ is the Stiefel manifold consisting of all $q \times k$ matrices $Q_1$ with orthonormal columns. The integral (2.3) is of the form $c_2 \int_{\mathscr{S}} f^n\phi$, where

$$ f = \exp\{-\tfrac{1}{2} \operatorname{tr}(U^2+V^2) + \operatorname{tr}([TH_1UH_1'H_2VH_2' \vdots O]Q_1\}\, |UV| $$

and

$$ \varphi = |UV|^{-k} \prod_{i<j}^{k} (u_i^2-u_j^2)(v_i^2-v_j^2). $$

It can be shown that $f$ achieves its maximum value at the $2^{2k}$ points in $\mathscr{S}$ of the form

$$H_1 = \begin{bmatrix} \pm 1 & & & 0 \\ & \pm 1 & & \\ & & \ddots & \\ 0 & & & \pm 1 \end{bmatrix}, \qquad Q_1 = \begin{bmatrix} I_k \\ \cdots \\ 0 \end{bmatrix},$$

$$H_2 = \begin{bmatrix} \pm 1 & & & 0 \\ & \pm 1 & & \\ & & \ddots & \\ 0 & & & \pm 1 \end{bmatrix},$$

$$U = V = \operatorname{diag}\{(1 - t_1)^{-\frac{1}{2}}, \ldots, (1 - t_k)^{-\frac{1}{2}}\},$$

and the maximum value of $f$ is

$$\hat{f} = e^{-k} \prod_{i=1}^{k} (1 - t_i)^{-1}. \tag{2.4}$$

At these maxima $\varphi$ has the value

$$\hat{\varphi} = \prod_{i=1}^{k} (1 - t_i)^{2-k} \prod_{i<j}^{k} (t_i - t_j)^2 \tag{2.5}$$

and it can be shown that the Hessian is

$$\Delta = 2^{2k} \prod_{i=1}^{k} t_i^{q-k} (1 - t_i)^{4-2k-q} \prod_{i<j}^{k} (t_i - t_j)^4 (t_i + t_j). \tag{2.6}$$

The number of variables $m$ in Hsu's result being integrated is $\frac{1}{2}k(k + 2q + 1)$. Substitution of (2.4), (2.5), and (2.6) in (2.1), together with an obvious simplification of $c_2$, yields the theorem. As a check on some very tedious algebra it can be noted that when $k = 1$, (2.2) agrees with the known asymptotic behavior of the classical hypergeometric function (see Luke [13, Sect. 7.2]).

The asymptotic behavior of the two-matrix ${}_2F_1^{(p)}$ function follows from Theorem 1. Let $R = \operatorname{diag}(r_1, \ldots, r_p)$, where $1 > r_1 > \cdots > r_p > 0$ and let $P$ be a $p \times p$ diagonal matrix of the form

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where $P_1 = \operatorname{diag}(\rho_1, \ldots, \rho_k)$ with $1 > \rho_1 > \cdots > \rho_k > 0$. Then we have

THEOREM 2. *As* $n \to \infty$,

$$
{}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) \sim c_3 \prod_{i=1}^{k} (1 - r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)} (r_i\rho_i)^{\frac{1}{2}(p-q)} \times \prod_{\substack{i=1 \\ i<j}}^{k} \prod_{j=1}^{p} C_{ij}^{-\frac{1}{2}}, \tag{2.7}
$$

*where*

$$
c_3 = (\tfrac{1}{2}n)^{-(1/2)k(p+q-k-1)} \pi^{-(1/2)k(k+1)} \Gamma_k(\tfrac{1}{2}q)\, \Gamma_k(\tfrac{1}{2}p)\, 2^{-k}[1 + O(n^{-1})]
$$

*and*

$$
C_{ij} = (r_i^2 - r_j^2)(\rho_i^2 - \rho_j^2) \qquad i = 1,\dots, k; \quad j = 1,\dots, p.
$$

*Proof.* This follows from (2.1), (2.2), and the fact that

$$
{}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) = \frac{\Gamma_k(\frac{1}{2}p)}{\pi^{\frac{1}{2}pk}} \int_{V(k,p)} {}_2F_1(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P_1H_1R^2H_1'P_1)(dH_1). \tag{2.8}
$$

Using (2.2) in (2.8) it follows that

$$
\begin{aligned}
&{}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2) \\
&\quad \sim c_4 \int_{V(k,p)} \prod_{i=1}^{k} [\delta_i^{\frac{1}{2}(k-q)}(1 - \delta_i)^{-n+\frac{1}{2}q}] \prod_{i<j}^{k} (\delta_i + \delta_j)^{-\frac{1}{2}} (dH_1),
\end{aligned}
$$

where $c_4 = c_1\Gamma_k(\frac{1}{2}p)\, \pi^{-(1/2)pk}$ and $\delta_1 > \cdots > \delta_k$ are the positive square roots of the latent roots of $P_1H_1R^2H_1'P_1$. This integral is of the form $c_4 \int_{V(k,p)} f^n\phi$, where

$$
f = \prod_{i=1}^{k} (1 - \delta_i)^{-1}
$$

and

$$
\varphi = \prod_{i=1}^{k} \delta_i^{\frac{1}{2}(k-q)} (1 - \delta_i)^{\frac{1}{2}q} \prod_{i<j}^{k} (\delta_i + \delta_j)^{-\frac{1}{2}}.
$$

It can be shown that $f$ has $2^k$ maxima which are obtained when $H_1$ has the form

$$
H_1 = \underbrace{\left[\begin{array}{ccc:c} \pm 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \pm 1 & \end{array}\right]}_{k \qquad\qquad p-k} k.
$$

At these values for $H_1$, $\delta_i = r_i\rho_i$ for $i = 1,..., k$, and the maximum value of $f$ is

$$\hat{f} = \prod_{i=1}^{k} (1 - r_i\rho_i)^{-1}.$$

The value of $\varphi$ at these maxima is

$$\hat{\varphi} = \prod_{i=1}^{k} (r_i\rho_i)^{\frac{1}{2}(k-q)}(1 - r_i\rho_i)^{\frac{1}{2}q} \prod_{i<j}^{k} (r_i\rho_i + r_j\rho_j)^{-\frac{1}{2}},$$

and it can be shown that the Hessian is

$$\varDelta = \prod_{i=1}^{k} (r_i\rho_i)^{k-p}(1 - r_i\rho_i)^{1-p} \prod_{i<j}^{k} \frac{(r_i^2 - r_j^2)(\rho_i^2 - \rho_j^2)}{(r_i\rho_i + r_j\rho_j)} \prod_{i=1}^{k} \prod_{j=k+1}^{p} [\rho_i^2(r_i^2 - r_j^2)].$$

The theorem now follows from a straightforward application of (2.1).

Substitution of (2.7) in (1.1) gives an asymptotic representation for the distribution of $r_1^2,..., r_p^2$ under the assumption that the population canonical correlation coefficients satisfy

$$1 > \rho_1 > \cdots > \rho_k > \rho_{k+1} = \cdots = \rho_p = 0. \qquad (2.9)$$

This is summarized in the following

THEOREM 3. *The asymptotic density function of* $r_1^2,..., r_p^2$ *for large* $n$, *when the population coefficients satisfy* (2.9), *is*

$$k_2 \prod_{i=1}^{k} (1 - r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}(r_i^2)^{\frac{1}{2}(q-p)-\frac{1}{2}}(1 - r_i^2)^{\frac{1}{2}(n-p-q-1)} \prod_{i<j}^{k} \left(\frac{r_i^2 - r_j^2}{\rho_i^2 - \rho_j^2}\right)^{\frac{1}{2}}$$

$$\times \prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}}$$

$$\times \prod_{i=k+1}^{p} (r_i^2)^{\frac{1}{2}(q-p-1)}(1 - r_i^2)^{\frac{1}{2}(n-q-p-1)} \prod_{\substack{k+1\\ i<j}}^{p} (r_i^2 - r_j^2), \qquad (2.10)$$

*where*

$$k_2 = k_1(\tfrac{1}{2}n)^{-\frac{1}{2}k(p+q-k-1)}\pi^{-\frac{1}{2}k(k+1)}\Gamma_k(\tfrac{1}{2}q)\, \Gamma_k(\tfrac{1}{2}p)2^{-k}$$
$$\times \prod_{i=1}^{k} (1 - \rho_i^2)^{\frac{1}{2}n}\, \rho_i^{k-\frac{1}{2}(p+q)}[1 + O(n^{-1})]$$

*and* $k_1$ *is given by* (1.2).

An alternative asymptotic result has been given by Chattopadhyay and Pillai [3] and Chattopadhyay, Pillai, and Li [4]; however the asymptotic behavior given by these authors involves a ${}_2F_1$ function with the matrix $P^2R^2$ as argument and appears to be incorrect. From Theorem 3 it is easy to obtain the following:

COROLLARY. *The asymptotic conditional density function of the $p - k$ smallest sample coefficients $r_{k+1}^2, ..., r_p^2$ given the first $k$ coefficient $r_1^2, ..., r_k^2$ is proportional to*

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}} \prod_{i=k+1}^{p} (r_i^2)^{\frac{1}{2}(q-p-1)} (1 - r_i^2)^{\frac{1}{2}(n-q-p-1)} \prod_{\substack{k+1 \\ i<j}}^{p} (r_i^2 - r_j^2). \quad (2.11)$$

From this we see that the largest $k$ sample coefficients $r_1^2, ..., r_k^2$ are asymptotically sufficient for $\rho_1^2, ..., \rho_k^2$. This suggests the use of (2.11) as a basis for testing the null hypothesis that the smallest $p - k$ population coefficients are all zero; this approach will be followed in the next section.

## 3. TESTING AND ESTIMATION

We first investigate the Barlett–Lawley test of the null hypothesis $H_0$: $\rho_k^2 > \rho_{k+1}^2 = \cdots = \rho_p^2 = 0$ against $H$: $\rho_{k+1}^2 \geqslant \cdots \geqslant \rho_p^2 \geqslant 0$ and $\rho_k^2 > 0$ using the distribution (2.11) of $r_{k+1}^2, ..., r_p^2$ given $r_1^2, ..., r_k^2$ which does not depend on the nuisance prameters $\rho_1, ..., \rho_k$. The approach given here is similar to that used by James [10] in another context.

The likelihood ratio statistic is

$$T_k = -\log \prod_{j=k+1}^{p} (1 - r_i^2)$$

and under $H_0$ Bartlett [1, 2] showed that $\{n - \frac{1}{2}(p + q + 1)\}T_k$ has an asymptotic $\chi^2$ distribution with $(p - k)(q - k)$ degrees-of-freedom. Lawley [12] obtained a correction to Bartlett's multiplying factor which makes the moments equal to those of the asymptotic $\chi^2$ distribution, apart from errors of order $n^{-2}$. Fujikoshi [6] has obtained an expansion for the asymptotic distribution of Lawley's statistic. This statistic involves the $k$ largest population coefficients and since these will usually be unknown, Lawley suggested, somewhat tentatively, that they be replaced by the $k$ largest sample coefficients. Here we attempt to provide some information about the accuracy of the approximation when this is done.

The appropriate multiplier of $T_k$ can be obtained by finding its expected value. For notational convenience let $E_c$ denote expectation taken with respect to the conditional distribution (2.11) of $r_{k+1}^2, ..., r_p^2$ given $r_1^2, ..., r_k^2$ and let $E_N$

denote expectation with respect to the null distribution obtained by ignoring the linkage factor

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} (r_i^2 - r_j^2)^{\frac{1}{2}}$$

in (2.11). In order to obtain $E_c(T_k)$ we first find $E_c(e^{-hT_k})$. This can obviously be done by finding

$$E_N\left[\left\{\prod_{i=1}^{k} \prod_{j=k+1}^{p} \left(1 - \frac{r_j^2}{r_i^2}\right)^{\frac{1}{2}}\right\} e^{-hT_k}\right]. \tag{3.1}$$

Writing

$$\prod_{i=1}^{k} \prod_{j=k+1}^{p} \left(1 - \frac{r_j^2}{r_i^2}\right)^{\frac{1}{2}} = 1 - \frac{\alpha}{2} \sum_{j=k+1}^{p} r_j^2 + O_p(n^{-2}),$$

where

$$\alpha = \sum_{i=1}^{k} r_i^{-2},$$

and substituting this in (3.1) it is seen that we need the following:

LEMMA.

$$E_N\left(e^{-hT_k} \sum_{j=k+1}^{p} r_j^2\right) = \frac{(p-k)(q-k)}{n-2k+2h} E_0(h), \tag{3.2}$$

*where* $E_0(h) = E_N(e^{-hT_k})$.

The proof of this follows easily from the fact that

$$\sum_{j=k+1}^{p} r_j^2 = \operatorname{tr}(I - U),$$

where $U$ is a $(p-k) \times (p-k)$ matrix having a multivariate Beta$\{\frac{1}{2}(n-q-k), \frac{1}{2}(q-k)\}$ distribution (see Kshirsagar [11, Chap. 8]). Using the lemma we can then show, from (3.1), that

$$E_c(e^{-hT_k}) = \theta(h)/\theta(0),$$

where $\theta(h) = E_0(h) f(h)$ with $f(h) = 1 - \alpha(p-k)(q-k)/[2(n-2k+2h)]$. Now

$$E_c(T_k) = -\frac{d}{dh}\left\{\frac{\theta(h)}{\theta(0)}\right\}_{h=0}$$

$$= -E_0'(0) - \frac{\alpha(p-k)(q-k)}{(n-2k)^2} + O(n^{-3}).$$

But $-E_0'(0) = E_N(T_k)$ and when $H_0$ is true we know that

$$[n-k-\tfrac{1}{2}(p+q+1)]T_k$$

has an asymptotic $\chi^2$ distribution with $(p-k)(q-k)$ degrees-of-freedom and the means agree to $O(n^{-2})$ so that

$$-E_0'(0) = (p-k)(q-k)/[n-k-\tfrac{1}{2}(p+q+1)] + O(n^{-3}).$$

Hence it follows that

$$E_c(T_k) = (p-k)(q-k)/\{[n-k-\tfrac{1}{2}(p+q+1)+\alpha] + O(n^{-1})\}.$$

Thus the appropriate multiplier of $T_k$ is $n-k-\frac{1}{2}(p+q+1)+\alpha$. Summarizing this, together with Lawley's result [12] we have the following theorem.

THEOREM 4. *The statistic*

$$L_k = \left\{n-k-\tfrac{1}{2}(p+q+1)+\sum_{i=1}^{k} r_i^{-2}\right\}T_k$$

*has an asymptotic* $\chi^2$ *distribution with* $(p-k)(q-k)$ *degrees-of-freedom and* $E_c(L_k) = (p-k)(q-k) + O(n^{-2})$.

We now turn to the problem of estimating the parameters $\xi_1, \ldots, \xi_p$ defined via the familiar transformation

$$\xi_i = \tanh^{-1}\rho_i = \tfrac{1}{2}\log\frac{1+\rho_i}{1-\rho_i}.$$

Let $z_i = \tanh^{-1} r_i$ $(i = 1, \ldots, p)$, the usual maximum likelihood estimate of $\xi_i$, which has a bias term of order $n^{-1}$. We will show that the maximum marginal likelihood estimate of $\xi_i$ provides a bias correction.

From (1.3) and (2.7) with $k = p$ we see that the asymptotic marginal log likelihood function is

$$\log L = \tfrac{1}{2}n \sum_{i=1}^{p} \log(1 - \rho_i^2) + \{\tfrac{1}{2}(p + q - 1) - n\} \sum_{i=1}^{p} \log(1 - r_i\rho_i)$$

$$+ \tfrac{1}{2}(p - q) \sum_{i=1}^{p} \log \rho_i - \tfrac{1}{2} \sum_{i<j} \log(\rho_i^2 - \rho_j^2)$$

from which it follows easily that the maximum marginal likelihood estimate of $\xi_i$ is

$$\hat{\xi}_i = z_i - \frac{1}{2nr_i} \left\{ p + q - 2 + r_i^2 + 2(1 - r_i^2) \sum_{j \neq i} \frac{r_j^2}{r_i^2 - r_j^2} \right\} + O(n^{-2}).$$

Using expressions for the mean and variance of $r_i$ given by Lawley [12] it can readily be verified

$$E(\hat{\xi}_i) = \xi_i + O(n^{-2})$$

and

$$\mathrm{Var}(\hat{\xi}_i) = 1/n + O(n^{-2})$$

so that these estimates stabilize the variance to order $n^{-1}$ and also provide a correction for bias.

## Acknowledgment

The authors would like to thank a referee for his useful comments.

## References

[1] Bartlett, M. S. (1938). Further aspects of the theory of multiple regression. *Proc. Cambridge Philos. Soc.* **34** 33–40.

[2] Bartlett, M. S. (1947). Multivariate Analysis. *J. Roy. Statist. Soc. Suppl.* **9** 176–190.

[3] Chattopadhyay, A. K., and Pillai, K. C. S. (1973). Asymptotic expansions for the distributions of characteristic roots when the parameter matrix has several multiple roots. In *Multivariate Analysis-III* (P. R. Krishnaiah, Ed.), pp. 117–127. Academic Press, New York.

[4] Chattopadhyay, A. K., Pillai, K. C. S., and Li, H. C. (1976). Maximization of an integral of a matrix function and asymptotic expansions of distributions of latent roots of two matrices. *Ann. Statist.* **4** 796–806.

[5] Constantine, A. G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

[6] Fujikoshi, Y. (1977). Asymptotic expansions for the distributions of some multivariate tests. In *Multivariate Analysis-IV* (P. R. Krishnaiah, Ed.), pp. 55–70. North–Holland, New York.
[7] Herz, C. S. (1955). Bessel functions of matrix argument. *Ann. of Math.* **61** 474–523.
[8] Hsu, L. C. (1948). A theorem on the asymptotic behavior of a multiple integral. *Duke Math. J.* **15** 623–632.
[9] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.
[10] James, A. T. (1969). Tests of equality of latent roots of the covariance matrix. In *Multivariate Analysis-II* (P. R. Krishnaiah, Ed.), pp. 205–218. Academic Press, New York.
[11] Kshirsagar, A. M. (1972). *Multivariate Analysis*. Dekker, New York.
[12] Lawley, D. N. (1959). Tests of significance in canonical analysis. *Biometrika* **46** 59–66.
[13] Luke, Y. L. (1969). *The Special Functions and Their Approximations*. Vol. 1. Academic Press, New York.