HIROSHIMA MATH. J.
22 (1992), 469–477

On confidence regions in canonical discriminant analysis

Yasunori FUJIKOSHI

(Received May 20, 1991)

1. Introduction

Consider q p-variate populations $\Pi_j, j = 1,\dots,q$ with means $\boldsymbol{\mu}_j$ and the same covariance matrix Σ, where $\boldsymbol{\mu}_j$ and Σ are unknown. Suppose that there are N_j observations $\boldsymbol{x}_{jk}$ from the j-th population Π_j $(k = 1,\dots,N_j; j = 1,\dots,q; N = \sum N_j)$. Let S and B be the matrices of sums of squares and products due to within populations and between populations, respectively, i.e.

$$S = \sum_{j=1}^{q} \sum_{k=1}^{N_j} (\boldsymbol{x}_{jk} - \bar{\boldsymbol{x}}_j)(\boldsymbol{x}_{jk} - \bar{\boldsymbol{x}}_j)'$$

and

$$B = \sum_{j=1}^{q} N_j(\bar{\boldsymbol{x}}_j - \bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_j - \bar{\boldsymbol{x}})',$$

where $\bar{\boldsymbol{x}}_j = (1/N_j)\sum_{k=1}^{N_j} \boldsymbol{x}_{jk}$ and $\bar{\boldsymbol{x}} = (1/N)\sum_{j=1}^{q}\sum_{k=1}^{N_j} \boldsymbol{x}_{jk}$. The canonical discriminant analysis introduced by Fisher [3] was developed by Rao [5, 6]. The method is used to summarize the differences between populations in terms of only a few transformed variates. Let $y_\alpha = \boldsymbol{c}_\alpha'(\boldsymbol{x} - \bar{\boldsymbol{x}})$, $\alpha = 1,\dots,p$ be the transformed variates, which are called canonical discriminant variates. The coefficient vectors $\boldsymbol{c}_\alpha$'s are defined as the solutions of

$$B\boldsymbol{c}_\alpha = \ell_\alpha S\boldsymbol{c}_\alpha, \quad \boldsymbol{c}_\alpha' S\boldsymbol{c}_\beta = n\delta_{\alpha\beta}, \tag{1.1}$$

where $\ell_1 \geq \cdots \geq \ell_p \geq 0$ and $n = N - q$. Let $\zeta_\alpha = \gamma_\alpha'(\boldsymbol{x} - \bar{\boldsymbol{\mu}})$, $\alpha = 1,\dots,p$ be the corresponding population canonical discriminant variates whose discriminant vectors are defined by

$$\Omega\gamma_\alpha = \lambda_\alpha \Sigma\gamma_\alpha, \quad \gamma_\alpha'\Sigma\gamma_\beta = \delta_{\alpha\beta}, \tag{1.2}$$

where $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, $\Omega = \sum_{j=1}^{q} (N_j/n)(\boldsymbol{\mu}_j - \bar{\boldsymbol{\mu}})(\boldsymbol{\mu}_j - \bar{\boldsymbol{\mu}})'$ and $\bar{\boldsymbol{\mu}} = (1/N)\sum_{i=1}^{q} N_j\boldsymbol{\mu}_j$. We assume that $\operatorname{rank}(\Omega) = m \leq \min(p, q-1)$. Then $\lambda_{m+1} = \cdots = \lambda_p = 0$, and only the first m canonical discriminant variates are meaningful. Suppose we are interesting in the canonical discriminant variates based on the first $s(\leq m)$ discriminant variates. Let $C = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_s]$, and

$$\bar{\boldsymbol{y}}_j = C'(\bar{\boldsymbol{x}}_j - \bar{\boldsymbol{x}}), \ \boldsymbol{\eta}_j = C'(\boldsymbol{\mu}_j - \bar{\boldsymbol{\mu}}). \tag{1.3}$$

It is assumed that all the observations are normal.

For the confidence regions on η_j, it has been used that $N_j(\bar{y}_j - \eta_j)'(\bar{y}_j - \eta_j)$ has an asymptotic χ^2-distribution with s degrees of freedom. We note that the asmptotic distributional result should be corrected under the sampling variability of the canonical discriminant vectors. Krzanowski [4] has noted that such regions are not appropriate as a surrounding for the set of transformed data $y_{jk} = C'(x_{jk} - \bar{x})$, $k = 1,\dots,N_j$. In the connection with the latter regions, we consider the confidence regions for a new observation from Π_j based on

$$w_j = y_j - \bar{y}, \tag{1.4}$$

where $y_j = C'(x_j - \bar{x})$ and x_j is a new observation from Π_j. In the case $q = 2$, w_j is equal to the studentized classification statistic W, whose asymptotic distribution has been obtained by Anderson [1]. In Section 2 we give a fundamental reduction for the distributions of $\sqrt{N_j}(\bar{y}_j - \eta_j)$ and w_j. In Section 3 asymptotic confidence regions for η_j and y_j are given by obtaining asymptotic distributions of $N_j(\bar{y}_j - \eta_j)'(\bar{y}_j - \eta_j)$ and $w_j'w_j$, respectively. In Section 4 we obtain an asymptotic expansion of the distribution of $w_j'w_j$, which gives the confidence regions for y_j with confidence coefficients up to the order N^{-1}.

2. A fundamental reduction

As is well known, S and B are independently distributed as a central Wishart distribution $W_p(\Sigma, n)$ and a noncentral Wishart distribution $W_p(\Sigma, q-1; n\Omega)$, respectively. Let

$$\tilde{S} = \frac{1}{n}\Gamma' S \Gamma,\ \tilde{B} = \frac{1}{n}\Gamma' B \Gamma,\ h_\alpha = \Gamma^{-1} c_\alpha,\ \alpha = 1,\dots,p. \tag{2.1}$$

Then the transformed vectors h_α's are the solutions of

$$\tilde{B} h_\alpha = \ell_\alpha \tilde{S} h_\alpha,\quad h_\alpha' \tilde{S} h_\beta = \delta_{\alpha\beta}. \tag{2.2}$$

Here $\tilde{S} \sim W_p(I_p, n)$. Further, it is well known that we can write $\tilde{B}$ as

$$\tilde{B} = \Lambda + \frac{1}{\sqrt{n}} M + \frac{1}{n} U_1 U_1', \tag{2.3}$$

where $U_1 = [u_1 \cdots u_{q-1}]$, the columns of $U = [U_1\ u_q]$ are independently distributed as $N_p(0, I_p)$, $M = [\sqrt{\lambda_1} u_1 \cdots \sqrt{\lambda_m} u_m\ O] + [\sqrt{\lambda_1} u_1 \cdots \sqrt{\lambda_m} u_m\ O]'$ and $\Lambda = \mathrm{diag}(\lambda_1 \cdots \lambda_p)$.

Since the distributions of $\sqrt{n_j}(\bar{y}_j - \eta_j)$ and w_j depend on $\bar{x}_j$ and $\bar{x}$ as well as $\tilde{S}$ and U, it is important to express these statistics in terms of $\tilde{S}$ and U only. Let

$$\Xi = \sqrt{N/n}\,[\sqrt{N_1/N}(\mu_1 - \bar{\mu}) \cdots \sqrt{N_q/N}(\mu_q - \bar{\mu})].$$

Then $\Omega = \Xi\Xi'$, and $\text{rank}(\Xi) = \text{rank}(\Omega) = m$.

LEMMA 2.1. *There exist a nonsingular $p \times p$ matrix Γ and an orthogonal $q \times q$ matrix $G = [G_1\ g_q]$ such that*

$$\begin{aligned} &\Gamma'\Omega\Gamma = \Lambda,\ \Gamma'\Sigma\Gamma = I_p \text{ and} \\ &\Gamma'\Xi G_1 = \begin{bmatrix} \Lambda_1^{1/2} & O \\ O & O \end{bmatrix}, \end{aligned} \tag{2.4}$$

where $g_q = (\sqrt{N_1/N} \cdots \sqrt{N_q/N})'$ and $\Lambda_1 = \text{diag}(\lambda_1, \ldots, \lambda_m)$.

PROOF. Let $\tilde{G}_1$ be a $q \times \overline{q-1}$ matrix such that $I_q - g_q g_q' = \tilde{G}_1\tilde{G}_1'$ and $\tilde{G}_1'\tilde{G}_1 = I_{q-1}$. It is easily checked that $\Xi\tilde{G}_1\tilde{G}_1' = \Xi$. Let $\Gamma = \Sigma^{-1/2}\Gamma_0$ and $G_1 = \tilde{G}_1 G_0$, where Γ_0 and G_0 are orthogonal $p \times p$ and $\overline{q-1} \times \overline{q-1}$ matrices, respectively. Then $\Gamma'\Sigma\Gamma = I_p$ and

$$\Gamma'\Xi G_1 = \Gamma_0'\Sigma^{-1/2}\Xi\tilde{G}_1 G_0. \tag{2.5}$$

Using Singular-valued Decomposition Theorem (see e.g., Rao [7, p. 42]) and noting that $\Sigma^{-1/2}\Xi\tilde{G}_1(\Sigma^{-1/2}\Xi\tilde{G}_1)' = \Sigma^{-1/2}\Xi\Xi'\Sigma^{-1/2}$, it is seen that there exist Γ_0 and G_0 such that the right-hand side of (2.5) is equal to the desired matrix. This completes the proof.

LEMMA 2.2. *Let $G = [G_1\ g_q]$ be an orthogonal $q \times q$ matrix satisfying* (2.4), *and let U be the random matrix difined in* (2.3). *Then*

(i) $\sqrt{N_j}(\bar{y}_j - \eta_j) = C'\Gamma'^{-1}U_1 g_{j1}$

(ii) $w_j = y_j - \bar{y}_j = C'(x_j - \mu_j) - (1/\sqrt{N_j})C'\Gamma'^{-1}U\tilde{g}_j$,

where $G_1 = [g_{11} \cdots g_{q1}]'$ and *$\tilde{g}_j = (g_{j1}'\ \sqrt{N_j/N})'$.*

PROOF. Let $z_j = \sqrt{N_j}\Gamma'(\bar{x}_j - \mu_j)$ and $Z = [z_1 \cdots z_q]$. Then $z_j \sim N_p(0, I_p)$. We have

$$\bar{x}_j - \bar{x} = \mu_j - \mu + (1/\sqrt{N_j})\Gamma'^{-1}\tilde{z}_j, \tag{2.6}$$

where $\tilde{z}_j = z_j - \sqrt{N_j/N}(\sqrt{N_1/N}\, z_1 + \cdots + \sqrt{N_q/N}\, z_q)$. First we show that the random matrix U in (2.3) may be defined by

$$U = ZG = Z[G_1\ g_q] \tag{2.7}$$

whose columns are independently distributed as $N_p(0, I_p)$. This will be shown by substituting (2.6) into $\tilde{B} = (1/n)\Gamma'B\Gamma$. In fact,

$$\tilde{Z} = [\tilde{z}_1 \cdots \tilde{z}_q] = ZG_1G_1' = U_1G_1', \tag{2.8}$$

and hence

$$\tilde{B} = \Lambda + (1/\sqrt{n})(\Gamma'\Xi\tilde{Z}' + \tilde{Z}\Xi'\Gamma) + (1/n)\tilde{Z}\tilde{Z}'$$
$$= \Lambda + (1/\sqrt{n})M + (1/n)U_1U_1'.$$

Using (2.6) $\sim$ (2.8) we obtain the expressions (i) and (ii) in the following way:

$$\sqrt{N_j}(\bar{y}_j - \eta_j) = C'\Gamma'^{-1}\tilde{z}_j = C'\Gamma'^{-1}U_1 g_{j1}$$

and

$$w_j = C'(x_j - \mu_j) - (1/\sqrt{N_j})C'\Gamma'^{-1}z_j$$
$$= C'(x_j - \mu_j) - (1/\sqrt{N_j})C'\Gamma'^{-1}U\tilde{g}_j.$$

Next we consider perturbation expansions for

$$C = \Gamma[h_1 \cdots h_s] = \Gamma H. \tag{2.9}$$

We make the following assumptions:

A1. *All the first s characteristic roots of* $\Omega\Sigma^{-1}$ *are simple, i.e.,*

$$\lambda_1 > \cdots > \lambda_s > \lambda_{s+1} \geq \cdots \geq \lambda_m > \lambda_{m+1} = \cdots = \lambda_p = 0.$$

A2. $\lim_{N\to\infty} N_j/N = d_j > 0,\ j = 1,\dots,q.$

Let

$$\tilde{S} = I_p + \frac{1}{\sqrt{n}}V. \tag{2.10}$$

Then, h_α's and ℓ_α's are the solutions of

$$\left[\Lambda + \frac{1}{\sqrt{n}}M + \frac{1}{n}U_1U_1'\right]h_\alpha = \ell_\alpha\left(I_p + \frac{1}{\sqrt{n}}V\right)h_\alpha,$$
$$h_\alpha'\left(I_p + \frac{1}{\sqrt{n}}V\right)h_\beta = \delta_{\alpha\beta}. \tag{2.11}$$

Under Assumption A1 it is known (see, e.g., Siotani, Hayakawa and Fujikoshi [8, *p.* 464]) that ℓ_α and h_α, $\alpha = 1,\dots,s$ are expanded as

$$\ell_\alpha = \lambda_\alpha + \frac{1}{\sqrt{n}}\ell_\alpha^{(1)} + \frac{1}{n}\ell_\alpha^{(2)} + \frac{1}{n\sqrt{n}}\ell_\alpha^{(3)} + O_p(n^{-2}),$$
$$h_\alpha = e_\alpha + \frac{1}{\sqrt{n}}h_\alpha^{(1)} + \frac{1}{n}h_\alpha^{(2)} + \frac{1}{n\sqrt{n}}h_\alpha^{(3)} + O_p(n^{-2}) \tag{2.12}$$

where e_α is the $p \times 1$ vector with α-th element one and other zero. The coefficients in (2.12) can be obtained by substituting (2.12) into (2.1) and equating the terms of $n^{-1/2}$ and n^{-1} in the expansions. These imply that

$$C = \Gamma\left\{\begin{bmatrix} I_s \\ O \end{bmatrix} + \frac{1}{\sqrt{n}}H^{(1)} + \frac{1}{n}H^{(2)} + \frac{1}{n\sqrt{n}}H^{(3)} + O_p(n^{-2})\right\}. \tag{2.13}$$

The matrices $H^{(j)} = [h_{i\alpha}^{(j)}]$ are given as follows.

$$h_{\alpha\alpha}^{(1)} = -\frac{1}{2}v_{\alpha\alpha},$$

$$h_{i\alpha}^{(1)},\ i \neq \alpha$$
$$= \lambda_{\alpha i}(m_{i\alpha} - \lambda_\alpha v_{i\alpha}),$$

$$h_{\alpha\alpha}^{(2)} = \frac{3}{8}v_{\alpha\alpha}^2 - \sum_{i\neq\alpha}^p \lambda_{\alpha i}v_{\alpha i}(m_{i\alpha} - \lambda_\alpha v_{i\alpha})$$
$$-\frac{1}{2}\sum_{i\neq\alpha}^p \lambda_{\alpha i}^2(m_{\alpha i} - \lambda_\alpha v_{\alpha i})^2,$$

$$h_{i\alpha}^{(2)},\ i \neq \alpha$$
$$= \lambda_{\alpha i}[(U_1'U_1)_{i\alpha} + \sum_{j\neq\alpha}^p \lambda_{\alpha j}(m_{ij} - \lambda_\alpha v_{ij})(m_{j\alpha} - \lambda_\alpha v_{ij})$$
$$-\frac{1}{2}(m_{i\alpha} - \lambda_\alpha v_{i\alpha})v_{\alpha\alpha} - v_{i\alpha}(m_{\alpha\alpha} - \lambda_\alpha v_{\alpha\alpha})$$
$$-\lambda_{\alpha i}(m_{i\alpha} - \lambda_\alpha v_{i\alpha})(m_{\alpha\alpha} - \lambda_\alpha v_{\alpha\alpha})],$$

where $\lambda_{\alpha i} = (\lambda_\alpha - \lambda_i)^{-1}$, $i \neq \alpha$, $M = [m_{i\alpha}]$, and $h_{i\alpha}^{(3)}$ is a homogeneous polynomial (not depending on n) of degree 3 in the elements of U and V. The coefficients $h_{i\alpha}^{(1)}$ have been given in Anderson [1].

3. Asymptotic confidence regions

In this section we obtain asymptotic confidence regions for η_i and y_j, based on asymptotic distributions of $N_j(\bar{y}_j - \eta_j)'(\bar{y}_j - \eta_j)$ and $w_j'w_j$, respectively.

THEOREM 3.1. *Under Assumptions* A.1 *and* A.2 *it holds that* $\sqrt{N_j}(\bar{y}_j - \eta_j)$ *is asymptotically distributed as* $N_s(\mathbf{0}, (1 - N_j/N)I_s)$.

PROOF. Using Lemma 2.2 and the fact that C converges to $\Gamma[I_s\ O]'$ in probability, we have that the asymptotic distribution of $\sqrt{N_j}(\bar{y}_j - \eta_j)$ is the same as the distribution of $U_{11}g_{i1}$. The distribution of $U_{11}g_{j1}$ is an s-dimensional normal with mean zero and covariance matrix

$$g_{j1}'g_{j1}I_s = (1 - N_j/N)I_s.$$

This completes the proof.

From Theorem 3.1 we have

$$N_j(1 - N_j/N)^{-1}(\bar{y}_j - \eta_j)'(\bar{y}_j - \eta_j) \longrightarrow \chi_s^2. \tag{3.1}$$

This gives confidence regions for η_j as hypersphere centered at $\bar{y}_j$ and having squared radii $\{(1 - N_j/N)/N_j\}\chi_{s,\alpha}^2$, where $\chi_{s,\alpha}^2$ is the upper α point of the χ^2-distribution with s degrees of freedom. It should be noted that the radii are not $\{N_j^{-1}\chi_{s,\alpha}^2\}^{1/2}$, but $[\{(1 - N_j/N)/N_j\}\chi_{s,\alpha}^2]^{1/2}$. Krzanowski [4] has pointed that the traditional confidence regions as hypersphere centered at $\bar{y}_j$ and having squared raii $N_j^{-1}\chi_{s,\alpha}^2$ are not appropriate as a surrounding for the set of transformed data y_{jk}, $k = 1,\ldots, N_j$. This note may be extended to the corrected regions as hypersphere centered at $\bar{y}_j$ and having squared radii $\{(1 - N_j/N)/N_j\}\chi_{s,\alpha}^2$.

Next we consider asymptotic confidence regions for the canonical discriminant value y_j of a new observation x_j from Π_j, which are closely related to a surrounding for the set of transformed data y_{jk}, $k = 1,\ldots, N_j$. From Lemma 2.2. (ii) it is easily seen that the asymptotic distribution of w_j is the same as the distribution of $[I_s\ O]\Gamma'(x_j - \mu_j)$. The latter distribution is an s variate normal with mean zero and covariance matrix $[I_s\ O]\Gamma'\Sigma\Gamma[I_s\ O]' = I_s$. Therefore, w_j is asymptotically distributed as $\mathrm{N}_s(\mathbf{0}, I_s)$. This asymptotic result is also obtained by noting that y_j is independent of C and C converges to $\Gamma[I_s\ O]'$ in probability. This implies that $w_j'w_j$ is asymptotically distributed as χ_s^2, and we obtain confidence regions for y_j,

$$(y_j - \bar{y}_j)'(y_j - \bar{y}_j) \leq \chi_{s,\alpha}^2 \tag{3.2}$$

as hyperspheres centered at $\bar{y}_j$ and having squared radii $\chi_{s,\alpha}^2$.

4. Asymptotic expansion

In order to obtain more accurate confidence coefficients of the confidence regions (3.2), we shall obtain an asymptotic expansion of the distribution of $w_j'w_j$. The conditional distribution of w_j given S and B is

$$\mathrm{N}_s\left[-\frac{1}{\sqrt{N_j}}C'\Gamma'^{-1}Ug_i,\ C'\Sigma C\right].$$

Therefore, the characteristic function of $w_j'w_j$ is given by

$$\psi(t) = \mathrm{E}\{e^{itw_j'w_j}\}$$

$$= \mathrm{E}[|I_s - 2itC'\Sigma C|^{-1/2} \tag{4.1}$$

$$\times \exp\{itN_j^{-1}g_j'U'\Gamma'^{-1}C(I_s - 2itC'\Sigma C)^{-1}C'\Gamma'Ug_j\}].$$

From Lemma 2.3 we can write

$$C'\Sigma C = I_s + \frac{1}{\sqrt{n}}Q^{(1)} + \frac{1}{n}Q^{(2)} + \frac{1}{n\sqrt{n}}Q^{(3)} + O_p(n^{-2}), \tag{4.2}$$

where

$$Q^{(1)} = [I_s\ O]H^{(1)} + H^{(1)'}\begin{bmatrix} I_s \\ O \end{bmatrix},$$

$$Q^{(2)} = [I_s\ O]H^{(2)} + H^{(2)'}\begin{bmatrix} I_s \\ O \end{bmatrix} + H^{(1)'}H^{(1)},$$

and $Q^{(3)}$ is a homogeneous polynomial of degree 3 in the elements of U and V. Substituting (4.2) into (4.1) and using $-\log|I_s - A| = \operatorname{tr} A + \frac{1}{2}\operatorname{tr} A^2 + \cdots$, we have

$$\begin{aligned}\psi(t) = \delta(t)^{s/2}\mathrm{E}\Big[& 1 + \frac{1}{2\sqrt{n}}(\delta(t) - 1)\operatorname{tr} Q^{(1)} \\ & + \frac{1}{n}(\delta(t) - 1)\Big\{\frac{1}{2}\operatorname{tr} Q^{(2)} + \frac{1}{4}(\delta(t) - 1)(\operatorname{tr} Q^{(1)})^2 \\ & + \frac{1}{8}(\delta(t) - 1)^2(\operatorname{tr} Q^{(1)})^2\Big\} \\ & + \frac{1}{2N_j}(\delta(t) - 1)g_j'U'\begin{bmatrix} I_s & O \\ O & O \end{bmatrix}Ug_j\Big] + O(n^{-2}),\end{aligned}$$

where $\delta(t) = (1 - 2it)^{-1}$. Here we used that $\mathrm{E}[\{$homogeneous polynomial of degree 3 in the elements of U and $V\}] = O(n^{-1/2})$. After much simplication, we obtain

$$\begin{aligned}\psi(t) = \delta(t)^{s/2}\Big[& 1 + \frac{1}{n}(\delta(t) - 1)\Big\{s + \textstyle\sum_{\alpha=1}^{s}\sum_{\beta\neq\alpha}^{p}\lambda_{\alpha\beta}\lambda_\alpha \\ & + \frac{1}{4}(\delta(t) - 1)(3s + 2\textstyle\sum\sum_{\alpha\neq\beta}^{s}\lambda_{\alpha\beta}\lambda_\alpha) + \frac{ns}{2N_j}\Big\}\Big] \\ & + O(n^{-2}).\end{aligned} \tag{4.3}$$

This gives the following result.

THEOREM 4.1. *Under Assumptions* A1 *and* A2 *it holds that*

$$P(w_j'w_j \le x) = P(\chi_s^2 \le x) - \frac{1}{n} g_s(x)$$

$$\left[\left\{\frac{1}{2}s + (\textstyle\sum_{\alpha=1}^{s}\sum_{\beta\neq\alpha}^{p} + \sum_{\alpha=1}^{s}\sum_{\beta=s+1}^{p})\lambda_{\alpha\beta}\lambda_\alpha + \frac{ns}{N_j}\right\}\frac{x}{s} + \left(\frac{3}{2}s + \textstyle\sum_{\alpha\neq\beta}^{s}\lambda_{\alpha\beta}\lambda_\alpha\right)\frac{x^2}{s(s+2)}\right] + O(n^{-2}),$$

where $g_s(x)$ is the density function of a χ^2-variate with s degrees of freedom and $\lambda_{\alpha\beta} = (\lambda_\alpha - \lambda_\beta)^{-1}$.

Theorem 4.1 implies that the $\chi^2_{s,\alpha}$ in (3.2) can be expanded as

$$\chi^2_{s,\alpha}\left[1 + \frac{1}{n}\left\{\frac{1}{2} + \frac{1}{s}(\textstyle\sum_{\alpha=1}^{s}\sum_{\beta\neq\alpha}^{p} + \sum_{\alpha=1}^{s}\sum_{\beta=s+1}^{p})\lambda_{\alpha\beta}\lambda_\alpha + \frac{n}{N_j} + \left(\frac{3}{2} + \frac{1}{s}\textstyle\sum_{\alpha\neq\beta}^{s}\lambda_{\alpha\beta}\lambda_\alpha\right)\frac{\chi^2_{s,\alpha}}{s+2}\right\}\right] + O(n^{-2}). \tag{4.4}$$

For a practical use, we need to replace λ_j by its estimate ℓ_j.

In a special case $q = 2$, $\lambda_1 > \lambda_2 = \cdots = \lambda_p = 0$ and

$$P(w_j'w_j \le x) = P(\chi_1^2 \le x) - \frac{1}{n} g_1(x)\left[\left(2p - \frac{3}{2} + \frac{n}{N_j}\right)x + \frac{1}{2}x^2\right] + O(n^{-2}). \tag{4.5}$$

The upper α point of $w_j'w_j$ can be expanded as

$$\chi^2_{1,\alpha}\left[1 + \frac{1}{n}\left\{2p - \frac{1}{2} + \frac{N_2}{N_1} + \frac{1}{2}\chi^2_{1,\alpha}\right\}\right] + O\,(n^{-2}). \tag{4.6}$$

These special results (4.5) and (4.6) can be also obtained from the result of Anderson [1]. In order to see this, first note that the coefficient vector of the canonical or linear discriminant function is

$$c = \frac{1}{D}\left(\frac{1}{n}S\right)^{-1}(\bar{x}_1 - \bar{x}_2),$$

where $D = \{(\bar{x}_1 - \bar{x}_2)'S^{-1}(\bar{x}_1 - \bar{x}_2)\}^{1/2}$. Anderson [1] has shown that

$$\begin{aligned} &P(c'(x - \bar{x}) - c'(\bar{x}_1 - \bar{x}) \le u \mid x \in \Pi_1) \\ &= P\left(\frac{1}{D}(\bar{x}_1 - \bar{x}_2)'\left(\frac{1}{n}S\right)^{-1}(x - \bar{x}_1) \le u \mid x \in \Pi_1\right) \\ &= \Phi(u) + \frac{1}{n}\phi(x)\left[\frac{(p-1)}{\lambda}\left(1 + \frac{N_2}{N_1}\right)\right. \end{aligned} \tag{4.7}$$

$$-\left(p-\frac{1}{4}+\frac{1}{2}\frac{N_2}{N_1}\right)u-\frac{1}{4}u^3\Bigg]+O(n^{-2}),$$

where $\lambda=\{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)'\Sigma^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)\}^{1/2}$, and $\Phi(x)$ and $\phi(x)$ are the distribution and the density functions of N(0, 1), respectively. This implies that

$$P(\{c'(\boldsymbol{x}_1-\bar{\boldsymbol{x}}_1)\}^2\le v\,|\,\boldsymbol{x}_1\in\Pi_1)$$
$$=P(\chi_1^2\le v)-\frac{2}{n}\phi(\sqrt{v})\left\{\left(p-\frac{1}{4}+\frac{N_2}{2N_1}\right)\sqrt{v}+\frac{1}{4}(\sqrt{v})^3\right\}+O(n^{-2})$$

which is coincident with (4.5).

References

[1] T. W. Anderson, The asymptotic distribution of certain characteristic roots and vectors, Proc. 2nd. Berkeley Symp. Math. Statist. Prob., Univ. of California Press, Berkeley, 1951, 103–130.

[2] T. W. Anderson, An asymptotic expansion of the distribution of the studentized classification statistic *W*, Ann. Statist. **1** (1973), 964–972.

[3] R. A. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugen. **7** (1936), 179–184.

[4] W. J. Krzanowski, On confidence regions in canonical variate analysis, Biometrika **76** (1989), 107–116.

[5] C. R. Rao, The utilization of multiple measurements in problems of biological classification (with discussion), J. R. Statist. Soc. B **10** (1948), 159–203.

[6] C. R. Rao, Advanced Statistical Methods in Biometric Research, Wiley, New York, 1952.

[7] C. R. Rao, Linear Statistical Inference and Its Applications, 2nd ed., Wiley, New York, 1973.

[8] M. Siotani, T. Hayakawa, and Y. Fujikoshi, Modern Multivariate Statistical Analysis: A Graduate Course and Handbook, American Science Press, INC., Ohio, 1985.

Department of Mathematics,
Faculty of Science,
Hiroshima University