The Annals of Statistics
1980, Vol. 8, No. 5, 958–976

ASYMPTOTIC REPRESENTATIONS OF THE DENSITIES OF CANONICAL CORRELATIONS AND LATENT ROOTS IN MANOVA WHEN THE POPULATION PARAMETERS HAVE ARBITRARY MULTIPLICITY[1]

BY WILLIAM J. GLYNN

Harvard University

Asymptotic representations of the joint densities of the canonical correlation coefficients, calculated from a sample from a multivariate normal population, and of the latent roots of $B(B+W)^{-1}$, where B is $W_p(n_1, \Sigma, \Omega)$ and W is $W_p(n_2, \Sigma)$, are obtained by deriving asymptotic representations of the hypergeometric functions in the joint densities. The results hold in the first case for large sample size and arbitrary values of the population canonical correlations and in the second case for large n_2 and $\Omega = n_2\Theta$, where the latent roots of the noncentrality matrix Ω are arbitrary.

1. Introduction and summary. Noncentral distributions of matrix variates and latent roots based on normal samples involve hypergeometric functions of matrix arguments. A survey of these distributions and definitions of hypergeometric functions can be found in James (1964) and Constantine (1963). The exact results are so complicated, however, that they are of little value for inferential or numerical purposes. A number of people have found approximations to some of these distributions, under various limiting conditions, by finding asymptotic representations of the hypergeometric functions involved in the exact densities (see Muirhead (1978) for a review). The general approach has been to apply a multivariate extension of Laplace's method, due to Hsu (1948), to an integral representation of the hypergeometric function. Usually it has been necessary to place some restrictions on the multiplicities of the population parameters. James (1969) uses an invariance argument to obtain an asymptotic representation, for large sample size, of the ${}_0F_0$ hypergeometric function which occurs in the joint density of the latent roots of a $p \times p$ sample covariance matrix. His results apply when the q smallest population roots are equal and the $p-q$ largest roots are distinct. Chattopadhyay and Pillai (1973) present a technique to obtain asymptotic representations from integral representations of hypergeometric functions with no restrictions on the population parameters. They assert that the method used by James is inappropriate in the MANOVA and canonical correlation cases when there are several multiple population roots because the invariance property does not apply. Instead they introduce a parametrization of a subset of the orthogonal group and argue that it is sufficient to restrict the region of integration to this subset.

Received June 1979.

[1]This paper is based on the author's Ph.D. dissertation from Yale University, 1977. Research supported in part by the National Science Foundation under Contract MCS77-27119.

AMS 1970 *subject classifications.* Primary 62E20, 62H20; secondary 62H10, 62H15.

Key words and phrases. Asymptotic representation, canonical correlations, hypergeometric functions, latent roots, MANOVA.

Asymptotic representations for the MANOVA and canonical correlation cases are presented in this paper. Section 2 contains an extension of Hsu's lemma which is used to derive the representations and to prove that the representations are uniform on certain subsets of the parameter and variable spaces. Section 3 contains a factorization of the invariant measure on the orthogonal group. This factorization is used to show that James' invariance argument does apply in the MANOVA and canonical correlation cases and leads to the parametrization suggested by Chattopadhyay and Pillai. Section 4 contains some results on the maximization of matrix functions.

The following notation is used throughout the paper:

(i) p is the number of population latent roots;

(ii) k is the number of nonzero population latent roots;

(iii) $\delta_1 > \delta_2 > \cdots > \delta_m \geqslant 0$ are the distinct population latent roots with multiplicities $q_1, \cdots, q_m$, respectively;

(iv) l is the number of nonzero δ_i;

(v) $\nu = q_1 + \cdots + q_l$;

(vi) $Q(1), \cdots, Q(m)$ are sets of integers defined by

$$Q(1) = \{1, 2, \cdots, q_1\}$$
$$Q(i) = \{q_1 + \cdots + q_{i-1} + 1, \cdots, q_1 + \cdots + q_i\} \qquad 2 \leqslant i \leqslant m;$$

(vii) $O(p)$ is the group of $p \times p$ orthogonal matrices;

(viii) $V(k, p)$ is the Stiefel manifold of $p \times k$ matrices with orthonormal columns; and

(ix) $q(i) = q_i$.

"Latent roots" refer to the roots of $B(B + W)^{-1}$ or to the canonical correlation coefficients.

Constantine (1963) computes the joint density of the squares, $r_1^2 \geqslant r_2^2, \cdots, \geqslant r_p^2$, of the sample canonical correlation coefficients between variates $x_1, \cdots, x_p$ and $y_1, \cdots, y_q (p \leqslant q)$ calculated from a sample of size $n + 1$ from a $(p + q)$-variate normal distribution. The result is

$$(1.1) \qquad C_n \Pi_{i=1}^p \{(1 - \rho_i^2)^{\frac{1}{2}n} (1 - r_i^2)^{\frac{1}{2}(n-q-p-1)} (r_i^2)^{\frac{1}{2}(q-p-1)}\} \Pi_{i<j}^p (r_i^2 - r_j^2)$$
$$\times {}_2F_1^{(p)}(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2)$$

where

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

$\Gamma_p(a)$ is a multivariate gamma function, $1 \geqslant \rho_1 \geqslant \rho_2 \geqslant \cdots \geqslant \rho_p \geqslant 0$ are the population canonical correlation coefficients, $P = \text{diag}(\rho_1, \cdots, \rho_p)$, and $R = \text{diag}(r_1, \cdots, r_p)$. Glynn and Muirhead (1978) derive an asymptotic representation for large n of the ${}_2F_1^{(p)}$ function which occurs in (1.1) assuming that the nonzero population coefficients are distinct. Their result, which only involves elementary functions, is obtained by a two step application of Hsu's lemma. We assume in this paper that

$$\rho_i = \delta_j \quad \text{for} \quad i \in Q(j). \tag{1.2}$$

Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai, and Li (1976) give an asymptotic expansion of the joint density of the r_i^2 when the ρ_i satisfy (1.2). However, their result involves a ${}_2F_1^{(p)}$ function of one matrix argument. An asymptotic representation of ${}_2F_1^{(p)}$ in (1.1) for large n, when the ρ_i satisfy (1.2), is presented in Section 5. This result is used to derive an asymptotic representation of the joint density of the r_i^2.

Suppose that B and W are independent $p \times p$ random matrices such that W has the Wishart distribution $W_p(n_2, \Sigma)$ and B has the noncentral Wishart distribution $W_p(n_1, \Sigma, \Omega)(n_1 \geqslant p)$. Let $L = \text{diag}(l_1, \cdots, l_p)$, where $1 \geqslant l_1 \geqslant l_2 \geqslant \cdots \geqslant l_p \geqslant 0$ are the latent roots of $B(B + W)^{-1}$. Constantine (1963) shows that the joint density of the l_i is

$$K\Pi_{i=1}^p\left\{ l_i^{\frac{1}{2}\{n(1)-p-1\}}(1 - l_i)^{\frac{1}{2}\{n(2)-p-1\}}\right\}\Pi_{i<j}^p(l_i - l_j) \tag{1.3}$$
$$\times \text{etr}\left(-\tfrac{1}{2}\Omega\right){}_1F_1^{(p)}\left\{\tfrac{1}{2}(n_1 + n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}\Omega, L\right\}$$

where

$$K = \pi^{\frac{1}{2}p^2}\Gamma_p\left\{\tfrac{1}{2}(n_1 + n_2)\right\}\left\{\Gamma_p\left(\tfrac{1}{2}n_1\right)\Gamma_p\left(\tfrac{1}{2}n_2\right)\Gamma_p\left(\tfrac{1}{2}p\right)\right\}^{-1},$$

and $n_i = n(i)$. Assume, without loss of generality, that $\Omega = \text{diag}(\omega_1, \omega_2, \cdots, \omega_p)$ with $\omega_1 \geqslant \omega_2 \geqslant \cdots \geqslant \omega_p \geqslant 0$. In a multivariate analysis of variance situation W and B are respectively the "within groups" and "between groups" matrices of sums of squares and sums of products. Asymptotic expansions of the density of L have been developed for large n_2 (large error df) and large Ω (some or all of the ω_i large). Constantine and Muirhead (1976) give an asymptotic expansion of the ${}_1F_1^{(p)}$ function in (1.3) when some of the ω_i are large. Chattopadhyay and Pillai (1973) and Chattopadhyay, Pillai, and Li (1976) find an asymptotic expansion of the ${}_1F_1^{(p)}$ function for large n_2. However, their result involves a ${}_1F_1^{(p)}$ function of one matrix argument. Let $\Theta = \text{diag}(\theta_1, \cdots, \theta_p)$, $\theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_p \geqslant 0$. An asymptotic representation for large n_2 of the ${}_1F_1^{(p)}$ function in (1.3), assuming

$$\Omega = n_2\Theta \tag{1.4}$$

and

$$\theta_i = \delta_j \quad \text{for} \quad i \in Q(j), \tag{1.5}$$

is presented in Section 6. This result is used to derive an asymptotic representation of the joint density of the l_i.

It is usually of interest in a typical analysis of variance, at least as a first step, to test the hypothesis that $\Omega = 0$. If this is rejected then one is often interested in testing a sequence of hypotheses of the form

$$H_k\colon \omega_{k+1} = \cdots = \omega_p = 0. \tag{1.6}$$

Bartlett (1947) proposed the test statistic

$$-\left\{n_2 + \tfrac{1}{2}(n_1 - p - 1)\right\}\ln\Pi_{i=k+1}^{p}(1 - l_i)$$

which, when H_k is true, is approximately distributed as χ^2 on $(p - k)(n_1 - k)$ df. The asymptotic representation of (1.3) is used in Section 6 to study Bartlett's test. The approach is the same as that used in other contexts by James (1969) and Glynn and Muirhead (1978).

2. An extension of Hsu's lemma. Let $\{f_\gamma: \gamma \in \Gamma\}$ and $g(x)$ be real valued functions defined on a subset D of real m-dimensional Euclidean space, R^m. We will write $f_\gamma(x)$ or $f(x; \gamma)$ for the value of f_γ at x, where $x' = (x_1, x_2, \cdots, x_m)$ is the transpose of x. Theorem 2.1 gives asymptotic representations for a family of integrals of the form $\int_D g(x)\{f(x; \gamma)\}^n dx$. If Γ consists of a single point then the theorem reduces to Corollary 2.1, which is essentially Hsu's result.

The notation '$a \sim b$ for large n' means $\lim(a/b) = 1$ as $n \to \infty$. If N is a subset of D and $a \in R^m$ then $a + N$ is the set $\{a + x : x \in N\}$. The infimums and supremums in the statement of the theorem are taken over all $\gamma \in \Gamma$.

THEOREM 2.1. *Let $\{f_\gamma : \gamma \in \Gamma\}$ and g be real valued functions defined on a subset D of R^m. Define the functions h_γ and p_γ by $h(x; \gamma) = f(x + \xi(\gamma); \gamma)$ and $p(x; \gamma) = g(x + \xi(\gamma))$ for all $x \in -\xi(\gamma) + D$ where $\xi(\gamma)$ is defined in condition* (i). *Suppose that there exists a neighborhood N of 0 such that the following conditions are satisfied.*

(i) *For each γ $f(x; \gamma)$ has an absolute maximum value at an interior point $\xi(\gamma)$ of D such that $0 < \inf\{f(\xi(\gamma); \gamma)\}$ and $0 < \inf\{|g(\xi(\gamma))|\}$;*

(ii) *there exist constants $s > 0$ and $G > 0$ such that $g(x)\{f(x; \gamma)\}^s$ is absolutely integrable on D and*

$$\int_D |g(x)\{f(x; \gamma)\}^s| dx \leqslant G \text{ for all } \gamma \in \Gamma;$$

(iii) *$\xi(\gamma) + N \subset D$ and $h(x; \gamma) > 0$ for all $\gamma \in \Gamma$ and $x \in N$;*

(iv) *all partial derivatives*

$$\frac{\partial h}{\partial x_i}(x; \gamma) \text{ and } \frac{\partial^2 h}{\partial x_i \partial x_j}(x; \gamma)$$

exist and are continuous functions of x on N;

(v) *there exists an $A < 1$ such that*

$$|h(x; \gamma)/h(0; \gamma)| < A \text{ for all } x \in -\xi(\gamma) + D - N \text{ and } \gamma \in \Gamma;$$

(vi) *the family of functions $\{p_\gamma : \gamma \in \Gamma\}$ is continuous on N and equicontinuous at* 0;

(vii) *the family of functions $\{w_{ij}(\cdot\,; \gamma) : \gamma \in \Gamma\}$ is equicontinuous at $x = 0$ where $W(x; \gamma)$ is the $m \times m$ symmetric matrix defined for all $\gamma \in \Gamma$ and $x \in N$ by $W(x; \gamma) = [w_{ij}(x; \gamma)]$ and*

$$w_{ij}(x; \gamma) = -\frac{\partial^2 \ln h}{\partial x_i \partial x_j}(x; \gamma); \text{ and}$$

(viii) $0 < \inf \lambda_m(\gamma)$ *and* $\sup \lambda_1(\gamma) < \infty$ *where* $\lambda_1(\gamma) \geqslant \cdots \geqslant \lambda_m(\gamma)$ *are the latent roots of* $W(0; \gamma)$.

Then for large n

$$\int_D g(x)\{f(x;\gamma)\}^n dx \sim (2\pi/n)^{\frac{1}{2}m}\{f(\xi(\gamma);\gamma)\}^n g(\xi(\gamma))\{\Delta(\xi(\gamma))\}^{-\frac{1}{2}} \tag{2.1}$$

uniformly in γ, *where* Δ *denotes the Hessian of* $-\ln f_\gamma$, i.e.,

$$\Delta(\xi(\gamma)) = \det\left[-\frac{\partial^2 \ln f}{\partial x_i \partial x_j}(\xi(\gamma);\gamma)\right].$$

PROOF. The proof for fixed γ is the same as the proof of Hsu's lemma. Write the integral in (2.1) as

$$\{f(\xi(\gamma);\gamma)\}^n \int_D g(x)\exp\left[n\{\ln f(x;\gamma) - \ln f(\xi(\gamma);\gamma)\}\right]dx$$

Choose a neighborhood $N(\gamma)$ of 0 such that $\ln f(x;\gamma) - \ln f(\xi(\gamma);\gamma)$ is approximately equal to $-\frac{1}{2}(x - \xi(\gamma))' W(0;\gamma)(x - \xi(\gamma))$ in $\xi(\gamma) + N(\gamma)$. Then choose n sufficiently large to make the integral over $-\xi(\gamma) + D - N(\gamma)$ negligible (see Hsu (1948) for details). The conditions of the theorem are sufficient to insure that $N(\gamma)$ and n can be chosen independently of γ.

COROLLARY 2.1. *Let f and g be real valued functions defined on a subset D of* R^m *such that*

(i) *f has an absolute maximum value at an interior point* ξ *of D and* $f(\xi) > 0$;

(ii) *there exists a constant* $s > 0$ *such that* gf^s *is absolutely integrable on* D;

(iii) *all partial derivatives*

$$\frac{\partial f}{\partial x_i} \text{ and } \frac{\partial^2 f}{\partial x_i \partial x_j}$$

exist and are continuous in a neighborhood N of ξ;

(iv) *there exists a constant* $A < 1$ *such that*

$$|f(x)/f(\xi)| < A \text{ for all } x \in D - N;$$

(v) *g is continuous in a neighborhood of* ξ *and* $g(\xi) \neq 0$.

Then for large n

$$\int_D g(x)\{f(x)\}^n dx \sim (2\pi/n)^{\frac{1}{2}m}\{f(\xi)\}^n g(\xi)\{\Delta(\xi)\}^{-\frac{1}{2}}.$$

3. A factorization of the invariant measure on the orthogonal group.

The results in this section follow from the work of James (1954). Let $G(q_1, \cdots, q_l)$, $q_1 + \cdots + q_l = k$, be real manifold defined as follows. A "point" p of $G(q_1, \cdots, q_l)$ consists of l orthogonal subspaces of R^k of dimensions $q_1, q_2, \cdots, q_l$. The manifold $G(q_1, \cdots, q_l)$ consists of all such points. Considered as a group of transformations on $G(q_1, \cdots, q_l)$, $O(k)$ is transitive. If p_0 is any point in $G(q_1, \cdots, q_l)$ then M, the isotropy subgroup at p_0, is defined by

$$M = \{\operatorname{diag}(H_1, \cdots, H_l) : H_i \in O(q_i)\} = O(q_1) \times \cdots \times O(q_l). \tag{3.1}$$

The manifold $G(q_1, \cdots, q_l)$ is identified with the coset space

$$O(k)/O(q_1) \times \cdots \times O(q_l).$$

We need to calculate the invariant measure on $G(q_1, \cdots, q_l)$. The result is summarized in the following lemmas.

LEMMA 3.1. *Let* $H = [h_1, h_2, \cdots, h_k] \in O(k)$. *The (unnormalized) invariant measure on* $G(q_1, \cdots, q_l)$ *is given by the following differential form*

$$\Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}h_j' dh_i$$

where Λ *denotes the exterior product and* $Q(\alpha)$ *is the set defined by* (vi).

PROOF. The result follows directly from the construction of the invariant measure on the Grassmann manifold given by James (1954).

LEMMA 3.2. *The normalized invariant measure on* $G(q_1, \cdots, q_l)$ *is*

$$\Gamma_k\left(\tfrac{1}{2}k\right)\pi^{-\frac{1}{2}k^2}\Pi_{i=1}^l\left\{\Gamma_{q(i)}\left\{\tfrac{1}{2}q(i)\right\}\pi^{-\frac{1}{2}q(i)^2}\right\}^{-1}\Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}h_j' dh_i.$$

PROOF. Let $A = [a_1, \cdots, a_k] \in O(k)$. The columns of A span l orthogonal subspaces in R^k which can be regarded as a point $p \in G(q_1, \cdots, q_l)$, i.e., the first q_1 columns span a subspace of dimension $q_1, \cdots,$ and the last q_l columns span a subspace of dimension q_l. The orthogonal matrix A is uniquely determined by p and the orientation of the q_i-frames in p. Introduce a reference matrix $H \in O(k)$ in p where the elements of H are analytic functions of p for almost all p. Then $A = HG$ for some $G = \text{diag}(G_1, \cdots, G_l) \in M$. If we write $G = [g_1, g_2, \cdots, g_k]$ then

$$\Lambda^l_{\alpha=1}\Lambda_{j>i;\, i,j\in Q(\alpha)}a_j' da_i = \Lambda^l_{\alpha=1}\Lambda_{j>i;\, i,j\in Q(\alpha)}g_j' dg_i + {}^*dH \tag{3.2}$$

and

$$\Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}a_j' da_i = \Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}h_j' dh_i \tag{3.3}$$

where ${}^*dH$ denotes linear differential forms in the elements of H. The right-hand side of (3.3) is a differential form of maximum degree defined on $G(q_1, \cdots, q_l)$ and H is a function defined on $G(q_1, \cdots, q_l)$. Therefore, the exterior product of (3.2) and (3.3) is

$$\Lambda^k_{j>i}a_j' da_i = \left\{\Lambda^l_{\alpha=1}\Lambda_{j>i;\, i,j\in Q(\alpha)}g_j' dg_i\right\}\Lambda\left\{\Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}h_j' dh_i\right\},$$

or equivalently

$$(dA) = B_1\Lambda^l_{i=1}(dG_i)\Lambda\left\{\Lambda^l_{\alpha<\beta}\Lambda_{i\in Q(\alpha)}\Lambda_{j\in Q(\beta)}h_j' dh_i\right\} \tag{3.4}$$

where

$$B_1 = \Gamma_k\left(\tfrac{1}{2}k\right)\pi^{-\frac{1}{2}k^2}\Pi_{i=1}^l\left\{\Gamma_{q(i)}\left\{\tfrac{1}{2}q(i)\right\}\pi^{-\frac{1}{2}q(i)^2}\right\}^{-1}$$

and (dA) and (dG_i) are the normalized invariant measures on $O(k)$ and $O(q_i)$, respectively. The lemma follows by integrating (dA) over

$$O(k) = G(q_1, \cdots, q_l) \times O(q_1) \times \cdots \times O(q_l).$$

LEMMA 3.3. *Let (dH) be the normalized invariant measure on $G(q_1, \cdots, q_l)$. Then*

$$(dA) = \{\Lambda_{i=1}^{l}(dG_i)\}\Lambda(dH).$$

PROOF. The lemma is a restatement of (3.4).

LEMMA 3.4. *If f is a function defined on $O(k)$ such that $f(A) = f(AG)$ for all $A \in O(k)$ and $G \in M$, where M is the isotropy subgroup defined by* (3.1), *then*

$$\int_{O(k)} f(A)(dA) = \int_{G(q_1, \cdots, q_l)} f(H)(dH).$$

PROOF. Write $A = HG$ with $H \in G(q_1, \cdots, q_l)$ and $G \in M$. Then the result follows from the fact that

$$\int_{O(k)} f(A)(dA) = \int_{G(q_1, \cdots, q_l)} \int_{O(q_1)} \cdots \int_{O(q_l)} f(HG)(dH)\Lambda\{\Lambda_{i=1}^{l}(dG_i)\}$$

and $f(HG) = f(H)$.

4. Maximization of matrix functions. This section contains some results on the maximization of certain matrix functions. The result that we need in the derivation of the asymptotic representation of ${}_2F_1^{(p)}$ is given in Corollary 4.1.

LEMMA 4.1. *Suppose that*

(i) *$U = \operatorname{diag}(u_1, \cdots, u_p)$ where $u_i = \eta_\alpha$ for $i \in Q(\alpha)$ and $\eta_1 > \eta_2 > \cdots > \eta_m \geqslant 0$;*

(ii) *l = number of nonzero η_i's;*

(iii) *$B(\gamma) = [b_{ij}(\gamma)]$ is a $p \times p$ matrix defined for all $\gamma \in \Gamma$; and*

(iv) *$\tau_1, \cdots, \tau_p$ are constants such that $\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_p$ and*

$$\Sigma_{\beta\leqslant\alpha}\Sigma_{i\in Q(\beta)} b_{ii}(\gamma) \leqslant \Sigma_{\beta\leqslant\alpha}\Sigma_{i\in Q(\beta)}\tau_i \qquad 1 \leqslant \alpha \leqslant l. \tag{4.1}$$

Then

$$\operatorname{tr}(UB(\gamma)) \leqslant \Sigma_{i=1}^{p} u_i\tau_i$$

with equality if and only if equality holds in (4.1) *for all* $\alpha (1 \leqslant \alpha \leqslant l)$.

PROOF. Let

$$S_\alpha(\gamma) = \Sigma_{\beta\leqslant\alpha}\Sigma_{i\in Q(\beta)} b_{ii}(\gamma) \qquad 1 \leqslant \alpha \leqslant l$$

and

$$\Delta_\alpha = \Sigma_{\beta\leqslant\alpha}\Sigma_{i\in Q(\beta)}\tau_i \qquad 1 \leqslant \alpha \leqslant l.$$

Then

$$\operatorname{tr}(UB(\gamma)) = \Sigma_{i=1}^{l-1} S_i(\gamma)(\eta_i - \eta_{i+1}) + S_l(\gamma)\eta_l$$

and

$$\Sigma_{i=1}^{p} u_i\tau_i = \Sigma_{i=1}^{l-1}\Delta_i(\eta_i - \eta_{i+1}) + \Delta_l\eta_l.$$

Therefore

$$\Sigma_{i=1}^{p} u_i\tau_i - \operatorname{tr}(UB(\gamma)) = \Sigma_{i=1}^{l-1}(\Delta_i - S_i(\gamma))(\eta_i - \eta_{i+1}) + (\Delta_l - S_l(\gamma))\eta_l.$$

By conditions (i) and (ii) $\eta_i - \eta_{i+1} > 0 (1 \leqslant i \leqslant l-1)$ and $\eta_l > 0$. By condition (iv) $\Delta_i - S_i(\gamma) \geqslant 0 (1 \leqslant i \leqslant l)$. Therefore,

$$\Sigma_{i=1}^{p} u_i\tau_i - \operatorname{tr}(UB(\gamma)) \geqslant 0$$

with equality if and only if $\Delta_i = S_i(\gamma)$, $(1 \leqslant i \leqslant l)$.

LEMMA 4.2. *Let*

$$f(H_1, \cdots, H_t) = \operatorname{tr}[U_1H_1U_2H_2 \cdots U_tH_t]$$

where $H_i \in O(p)$, $(1 \leqslant i \leqslant t)$, $U_j = \operatorname{diag}(u_{j1}, \cdots, u_{jp})$ *with* $u_{j1} > u_{j2} > \cdots > u_{jp} > 0$, $(1 \leqslant j \leqslant t-1)$, *and* $U_t = \operatorname{diag}(u_{t1}, \cdots, u_{tp})$ *with* $u_{tj} = \eta_\alpha$ *for* $j \in Q(\alpha)$, $(1 \leqslant \alpha \leqslant m)$, *and* $\eta_1 > \cdots > \eta_m \geqslant 0$. *Then*

$$f \leqslant \Sigma_{i=1}^{p} u_{1i}u_{2i} \cdots u_{ti} \tag{4.2}$$

with equality if and only if

$$H_t = \operatorname{diag}(M_1, \cdots, M_l, G_t), \qquad H_{t-1} = \operatorname{diag}(\pm M_1, \cdots, \pm M_l, G_{t-1}), \tag{4.3}$$

and

$$H_i = \operatorname{diag}(\pm 1, \cdots, \pm 1, G_i) \qquad 1 \leqslant i \leqslant t-2$$

where $l =$ *number of nonzero* η_i's, $\nu = q_1 + \cdots + q_l$, $G_i \in O(p-\nu)$, $M_i \in O(q_i)$, I_ν *is the* $\nu \times \nu$ *identity matrix, and the* H_i's *satisfy*

$$H_1H_2 \cdots H_t = \operatorname{diag}(I_\nu, G)$$

for some $G \in O(p-\nu)$. *If* $l = m$ *then* $p = \nu$ *and the* G_i's *do not appear.*

PROOF. The proof is based on the fact that maximizing f is equivalent to maximizing a class of functions generated by f. The structure of this class is made explicit by Lemma 4.1. The following notation will be used in the proof:

(a) $\Gamma = \{(H_1, \cdots, H_t) : H_i \in O(p)\}$ and γ is a point in Γ;

(b) Γ_0 is the subset of Γ consisting of all $(H_1, \cdots, H_t)$ which satisfy (4.3);

(c) $B_i(\gamma) = H_iU_{i+1}H_{i+1} \cdots U_tH_tH_1 \cdots H_{i-1}$, $(1 \leqslant i \leqslant t)$;

(d) $\tau_{i1} = \sup b_{i1}(\gamma)$ and $\tau_{ij} = \sup(b_{i1}(\gamma) + \cdots + b_{ij}(\gamma)) - \tau_{i,j-1}$, $(1 \leqslant i \leqslant t, 2 \leqslant j \leqslant p)$, where $b_{ij}(\gamma)$ is the jth diagonal element of $B_i(\gamma)$ and the supremum is over all $\gamma \in \Gamma$;

(e) $H(r, s)$ is the $r \times s$ matrix formed by the first r rows and s columns of the $p \times p$ matrix H (in particular $H(p, p) = H$);

(f) $C_i(\gamma, k_1, \cdots, k_i) = U_{i+1}H_{i+1} \cdots U_{t-1}H_{t-1}U_tH_t(p, k_1)H_1(k_1, k_2) \cdots H_i(k_i, p)$, $(1 \leqslant i \leqslant t-1)$, and $C_t(\gamma, k_1, \cdots, k_t) = H_1(k_1, k_2) \cdots H_{t-1}(k_{t-1}, k_t)H_t(k_t, k_1)$;

(g) $u_{t+1,j} = 1$, $(1 \leqslant j \leqslant p)$;

(h) $P_i(k_1, \cdots, k_i)$, $(1 \leqslant i \leqslant t)$ denotes the proposition

$$\operatorname{tr}\{C_i(\gamma, k_1, \cdots, k_i)\} \leqslant \Sigma_{j=1}^{\min(k_1, \cdots, k_i)} u_{i+1,j} \cdots u_{t+1,j} \textit{ for all } \gamma \in \Gamma; \textit{ and}$$

(i) $'a \to b'$ means $'a$ implies b'.

We will write $f(\gamma)$ for the value of f at $\gamma \in \Gamma$. Note that $f(\gamma) = \mathrm{tr}(U_1 B_1(\gamma))$ and by Lemma 4.1 $f(\lambda) \leqslant \Sigma_{j=1}^{p} u_{1j}\tau_{1j}$. The proof consists of showing that

$$(1.) \qquad \tau_{ij} = u_{i+1,j}u_{i+2,j} \cdots u_{t+1,j} \qquad 1 \leqslant j \leqslant p, 1 \leqslant i \leqslant t$$

and

$$(2.) \qquad f(\gamma) = \Sigma_{j=1}^{p} u_{1j}\tau_{1j} \textit{ if and only if } \gamma \in \Gamma_0.$$

If $\gamma \in \Gamma_0$ then $b_{ij}(\gamma) = u_{i+1,j} \cdots u_{t+1,j}$. Therefore, to prove (1.) it is sufficient to prove that

$$(4.4) \qquad \Sigma_{k=1}^{j} b_{ik}(\gamma) \leqslant \Sigma_{k=1}^{j} u_{i+1,k} \cdots u_{t+1,k}$$

for all $\gamma \in \Gamma$, i, $(1 \leqslant i \leqslant t)$, and j, $(1 \leqslant j \leqslant p)$. This is equivalent to $P_i(p, \cdots, p, j)$ for all j, $(1 \leqslant j \leqslant p)$, and i, $(1 \leqslant i \leqslant t)$, since

$$\Sigma_{k=1}^{j} b_{ik}(\gamma) = \mathrm{tr}\{C_i(\gamma, p, \cdots, p, j)\}.$$

By Lemma 4.1

$$P_j(k_1, \cdots, k_j) \textit{ for all } k_j(1 \leqslant k_j \leqslant p) \rightarrow P_{j-1}(k_1, \cdots, k_{j-1}) \qquad 2 \leqslant j \leqslant t.$$

Therefore,

$$P_t(k_1, \cdots, k_t) \textit{ for all } k_1, \cdots, k_t \qquad (1 \leqslant k_1, \cdots, k_t \leqslant p) \rightarrow (4.4).$$

Note that $P_t(k_1, \cdots, k_t)$ is equivalent to

$$(4.5) \qquad \mathrm{tr}(E) \leqslant \min(k_1, \cdots, k_t)$$

where $E = H_1(k_1, k_2) \cdots H_t(k_t, k_1)$. Since E is the product of submatrices of orthogonal matrices, the elements of E are all less than or equal to 1 in magnitude. Therefore (4.5) holds which implies that (1.) is true.

If $\gamma \in \Gamma_0$ then by (1.) $f(\gamma) = u_{11}\tau_{11} + \cdots + u_{1p}\tau_{1p}$. Suppose conversely, that $f(\gamma) = u_{11}\tau_{11} + \cdots + u_{1p}\tau_{1p}$. Note that $\mathrm{tr}(B_i(\gamma)) = \mathrm{tr}(U_{i+1}B_{i+1}(\gamma))$, $(1 \leqslant i \leqslant t - 1)$. It follows from this fact, (1.) and Lemma 4.1 that

$$\mathrm{tr}(U_i B_i(\gamma)) = \Sigma_{j=1}^{p} u_{ij}\tau_{ij} \rightarrow \mathrm{tr}(U_{i+1}B_{i+1}(\gamma)) = \Sigma_{j=1}^{p} u_{i+1,j}\tau_{i+1,j} \qquad 1 \leqslant i \leqslant t-1.$$

In particular, since $f(\gamma) = \mathrm{tr}(U_1 B_1(\gamma))$, $f(\gamma) = \Sigma_{j=1}^{p} u_{1j}\tau_{1j}$ implies

$$(4.6) \qquad \mathrm{tr}(U_i B_i(\gamma)) = \Sigma_{j=1}^{p} u_{ij}\tau_{ij} \qquad 1 \leqslant i \leqslant t.$$

The result (4.3) is obtained by considering (4.6) for $i = t, t-1, \cdots, 1$. If (4.6) holds for $i = t$ then by Lemma 4.1 $\Sigma_{j \in Q(\alpha)} b_{tj}(\gamma) = q_\alpha$, $(1 \leqslant \alpha \leqslant l)$. This implies, since $B_t(\gamma) = H_t H_1 \cdots H_{t-1} \in O(p)$ that

$$(4.7) \qquad H_t H_1 \cdots H_{t-1} = \mathrm{diag}(I_\nu, G)$$

for some $G \in O(p - \nu)$. We get, by substituting (4.7) into the definition of $B_{t-1}(\gamma)$,

$$(4.8) \qquad B_{t-1}(\gamma) = H_{t-1}\mathrm{diag}(u_{t1}, \cdots, u_{t\nu}, 0, \cdots, 0)H'_{t-1}.$$

Again by Lemma 4.1, if (4.6) holds for $i = t - 1$ then $b_{t-1,j}(\gamma) = u_{tj}$, $(1 \leqslant j \leqslant p)$.

This together with (4.8) implies that $H_{t-1} = \operatorname{diag}(M_1, \cdots, M_t, G_{t-1})$. The lemma follows by continuing this procedure for $i = t-2, t-3, \cdots, 1$.

COROLLARY 4.1. *Let*

$$f(G, H, Q, F) = \operatorname{tr}\{[UGVHP_1H'Q'R : O]F\}$$

where

(i) $U = \operatorname{diag}(u_1, \cdots, u_k)$ *with* $u_1 > \cdots > u_k > 0$;
(ii) $V = \operatorname{diag}(v_1, \cdots, v_k)$ *with* $v_1 > \cdots > v_k > 0$;
(iii) $P_1 = \operatorname{diag}(\rho_1, \cdots, \rho_k)$ *where the* ρ_i*'s satisfy* (1.2);
(iv) $R = \operatorname{diag}(r_1, \cdots, r_p)$ *with* $r_1 > r_2 > \cdots > r_p > 0$;
(v) $G, H \in O(k)$, $G = (g_{ij})$, $H = (h_{ij})$;
(vi) $Q \in V(k, p)$, $Q = (q_{ij})$, $p \geqslant k$; *and*
(vii) $F \in V(k, q)$, $F = (f_{ij})$, $q \geqslant p$.

Then

$$f \leqslant \textstyle\sum_{i=1}^k u_i v_i r_i \rho_i$$

with equality if and only if $G = \operatorname{diag}(\pm 1, \cdots, \pm 1)$, $H = \operatorname{diag}(H_1, \cdots, H_l)$,

$$Q = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ & 0 & \end{bmatrix} \text{ and } F = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ & 0 & \end{bmatrix}$$

where $H_i \in O(q_i)$ *and* $g_{ii}q_{ii}f_{ii} = 1$, $(1 \leqslant i \leqslant k)$.

PROOF. Define $U_1 = \operatorname{diag}(r_1, \cdots, r_q)$, $U_2 = \operatorname{diag}(u_1, \cdots, u_q)$, $U_3 = \operatorname{diag}(v_1, \cdots, v_q)$, and $U_4 = \operatorname{diag}(P_1, O)$ where u_i, v_i, $(k+1 \leqslant i \leqslant q)$, and r_i, $(p+1 \leqslant i \leqslant q)$, are chosen so that $u_k > u_{k+1} > \cdots > u_q > 0$, $v_k > v_{k+1} > \cdots > v_q > 0$, and $r_p > \cdots > r_q > 0$. Also choose $G_0, H_0 \in O(q-k)$, and $F_0, Q_0 \in V(q-k, q)$ such that $H_1, H_2, H_3, H_4 \in O(q)$ where

$$H_1 = [F : F_0], \qquad H_2 = \operatorname{diag}(G, G_0), \qquad H_3 = \operatorname{diag}(H, H_0),$$

and

$$H_4 = H_3' \begin{bmatrix} Q' : O \\ \cdots \\ Q_0' \end{bmatrix}. \tag{4.9}$$

Then $f = \operatorname{tr}[U_1H_1U_2H_2U_3H_3U_4H_4]$ and the result follows from Lemma 4.2.

5. An asymptotic representation—canonical correlation case. We first obtain an asymptotic representation for large n of the ${}_2F_1^{(p)}$ function which occurs in (1.1). The result is summarized in Theorem 5.1. The following identities are used in the proof of the theorem to derive the integral representation of ${}_2F_1^{(p)}$:

$$\begin{aligned} {}_{u+1}F_v(a_1, \cdots, a_u, a; b_1, \cdots, b_v; S) &= \{\Gamma_p(a)\}^{-1} \textstyle\int_{T>0} \operatorname{etr}(-T) \det T^{a-\frac{1}{2}(p+1)} \\ &\quad \times {}_uF_v(a_1, \cdots, a_u; b_1, \cdots, b_v; ST)(dT) \end{aligned} \tag{5.1}$$

$$(5.2)\qquad {}_uF_v(a_1, \cdots, a_u; b_1, \cdots, b_v; S, T) = \int_{O(p)} {}_uF_v(a_1, \cdots, a_u; b_1, \cdots, b_v; SH'TH)(dH)$$

and Bessel's integral

$$(5.3)\qquad \int_{V(k,p)} \operatorname{etr}(XH_1)(dH_1) = {}_0F_1\left(\tfrac{1}{2}p; \tfrac{1}{4}XX'\right)$$

where S and T are $p \times p$ symmetric matrices, X is a $k \times p$ matrix, $(k \leqslant p)$, $H \in O(p)$, $H_1 \in V(k,p)$ and (dH) and (dH_1) are the normalized invariant measures on $O(p)$ and $V(k,p)$ respectively.

THEOREM 5.1. *If the population correlation coefficients satisfy* (1.2), *then for large n*

$$\begin{aligned}{}_2F_1^{(p)}\left(\tfrac{1}{2}n, \tfrac{1}{2}n; \tfrac{1}{2}q; P^2, R^2\right) \sim\ & C_1 \Pi_{i=1}^k \left\{(1 - r_i\rho_i)^{-n+\frac{1}{2}(p+q-1)}(\rho_i r_i)^{\frac{1}{2}(p-q)}\right\} \\ (5.4)\qquad & \times \Pi_{\alpha=1}^l \Pi_{i<j;\, i,j \in Q(\alpha)} \left\{(1 - r_i\rho_i)(1 - r_j\rho_j)(r_i + r_j)\rho_i\right\}^{-\frac{1}{2}} \\ & \times \Pi_{\alpha<\beta}^l \Pi_{i \in Q(\alpha)} \Pi_{j \in Q(\beta)} \left\{\left(\rho_i^2 - \rho_j^2\right)\left(r_i^2 - r_j^2\right)\right\}^{-\frac{1}{2}} \\ & \times \Pi_{i=1}^k \Pi_{j=k+1}^p \left\{\left(r_i^2 - r_j^2\right)\rho_i^2\right\}^{-\frac{1}{2}}\end{aligned}$$

where

$$\begin{aligned}C_1 = {}& e^{-nk}\left(\tfrac{1}{2}n\right)^{nk - \frac{1}{2}k(p+q)} \Gamma_k\left(\tfrac{1}{2}p\right)\Gamma_k\left(\tfrac{1}{2}q\right)\left\{\Gamma_k\left(\tfrac{1}{2}n\right)\right\}^{-2} \\ & \times \Pi_{i=1}^l \left[\pi^{\frac{1}{2}q(i)^2}(2\pi/n)^{-\frac{1}{4}q(i)(q(i)-1)}\Gamma_{q(i)}\left\{\tfrac{1}{2}q(i)\right\}^{-1}\right]\end{aligned}$$

$q(i) = q_i$, and k and l are defined by (ii) *and* (iv). *Furthermore,* (5.4) *holds uniformly on any set of r_i's and ρ_i's such that the r_i's are strictly bounded away from* 1, 0, *and one another and the δ_j's (distinct values of the ρ_i's) are similarly bounded.*

PROOF. We can express ${}_2F_1^{(p)}$, using (5.1)–(5.3), as

$$B_n \int_\Lambda h(x)\{f(x)\}^n dx$$

where

$$\begin{aligned} f &= \operatorname{etr}\left(-\tfrac{1}{2}U^2 - \tfrac{1}{2}V^2 + [UGVHP_1H'Q'R : O]F\right)\det(UV), \\ h &= \Pi_{i<j}^k \left\{\left(v_i^2 - v_j^2\right)\left(u_i^2 - u_j^2\right)\right\}\det(UV)^{-k}, \\ dx &= (dU)(dG)(dV)(dH)(dQ)(dF), \\ B_n &= \left(\tfrac{1}{2}n\right)^{nk}\pi^{k^2}2^{2k}\left\{\Gamma_k\left(\tfrac{1}{2}n\right)\Gamma_k\left(\tfrac{1}{2}k\right)\right\}^{-2}, \\ \Lambda &= D(U) \times O(k) \times D(V) \times O(k) \times V(k,p) \times V(k,q), \\ D(U) &= \{\operatorname{diag}(u_1, \cdots, u_k) : u_1 > u_2 > \cdots > u_k > 0\}\end{aligned}$$

$H, G \in O(k)$, $Q \in V(k,p)$, $F \in V(k,q)$, $U \in D(U)$, $V \in D(V)$ and $P_1 = \operatorname{diag}(\rho_1, \cdots, \rho_k)$. The integrand is invariant under transformations of the form $H \to HW$, $W \in M$, where M is the isotropy subgroup defined by (3.1). Therefore, by Lemma 3.4, we can replace Λ by

$$\Upsilon = D(U) \times O(k) \times D(V) \times G(q_1, \cdots, q_l) \times V(k,p) \times V(k,q)$$

where the matrix H is to be interpreted as a point in the manifold $G(q_1, \cdots, q_l)$ and (dH) is the invariant measure on $G(q_1, \cdots, q_l)$ given by Lemma 3.2.

It follows from Corollary 4.1 that the location of the maximum of f, for fixed U and V, is independent of U and V. Consequently, f can be maximized in two steps as follows

$$\max_{U, V, H, G, Q,} f = \max_{U, V} \{\max_{H, G, Q, F,} f\}$$

It follows from this that the maximum value of f is

$$e^{-k}\Pi_{i=1}^{k}(1 - r_i\rho_i)^{-1}$$

and the maximum is obtained at the 2^{2k} points of Υ defined by

$$U = V = \operatorname{diag}\left((1 - \rho_1 r_1)^{-\frac{1}{2}}, \cdots, (1 - \rho_k r_k)^{-\frac{1}{2}}\right)$$

and H, Q, F, and G are defined as in Corollary 4.1. Note that the values of H define a unique point in $G(q_1, \cdots, q_l)$.

The integral over Υ can be replaced by 2^{2k} times the integral over Π where

$$\Pi = D(U) \times O^+(k) \times D(V) \times G(q_1, \cdots, q_l) \times V^+(k,p) \times V^+(k,q),$$

and $O^+(k)$, $V^+(k,p)$, and $V^+(k,q)$ are subsets of $O(k)$, $V(k,p)$, and $V(k,q)$, respectively, consisting of matrices with positive diagonal elements which, in the case of orthogonal matrices, are proper. This is a consequence of the fact that the behavior of f is identical in neighborhoods of each of the points at which it obtains its maximum and it is strictly bounded away from its maximum value outside of these neighborhoods.

Since G is proper in Π, it can be parametrized as

$$G = \exp(S), \qquad S = (s_{ij})$$

where S is $k \times k$ skew symmetric. Anderson (1965) computes the Jacobian of this transformation as

$$J_1 = J(G \to S) = \Gamma_k\left(\tfrac{1}{2}k\right)2^{-k}\pi^{-\frac{1}{2}k^2}\left(1 + O\left(s_{ij}^2\right)\right)$$

where $O(s_{ij}^2)$ denotes terms which are at least quadratic in the elements of S. James (1969) proves that a parametrization of Q and F can be obtained by writing

$$[Q : -] = \exp(W) = \exp\begin{pmatrix} W_{11} & W_{12} \\ -W_{12}' & O \end{pmatrix}, \qquad W = (w_{ij})$$

and

$$[F : -] = \exp(Z) = \exp\begin{pmatrix} Z_{11} & -Z_{21}' \\ Z_{21} & O \end{pmatrix}, \qquad Z = (z_{ij})$$

where $[Q : -]$ is a $p \times p$ orthogonal matrix whose first k columns are Q, $[F : -]$ is a $q \times q$ orthogonal matrix whose first k columns are F, Z_{11} and W_{11} are $k \times k$

skew symmetric, W_{12} is $k \times (p-k)$, and Z_{21} is $(q-k) \times k$. The Jacobians of these transformations are (James (1969))

$$J_2 = J(Q \to (W_{11}, W_{12})) = \Gamma_k(\tfrac{1}{2}p)2^{-k}\pi^{-\frac{1}{2}kp}(1 + O(w_{ij}^2))$$

and

$$J_3 = J(F \to (Z_{11}, Z_{21})) = \Gamma_k(\tfrac{1}{2}q)2^{-k}\pi^{-\frac{1}{2}kq}(1 + O(z_{ij}^2)).$$

Similarly, $H \in G(q_1, \cdots, q_l)$ can be parametrized by writing

$$H = \exp(T), \qquad T = (t_{ij})$$

where T is a $k \times k$ skew symmetric matrix such that $t_{ij} = 0$ if $i, j \in Q(\alpha)$ for some α, $(1 \leqslant \alpha \leqslant l)$. This is the parametrization used by Chattopadhyay and Pillai (1973). The Jacobian of this transformation is

$$J_4 = J(G \to T) = \Gamma_k(\tfrac{1}{2}k)\pi^{-\frac{1}{2}k^2}\Pi_{i=1}^{l}\left\{\Gamma_{q(i)}\{\tfrac{1}{2}q(i)\}\pi^{-\frac{1}{2}q(i)^2}\right\}^{-1}(1 + O(t_{ij}^2)).$$

By making this change of variables we can express ${}_2F_1^{(p)}$ as

$$B_n \textstyle\int_{\Xi} g f^n \tag{5.5}$$

where

$$g = \Pi_{i=1}^{k}(u_i v_i)^{-k}\Pi_{i<j}^{k}\{(u_i^2 - u_j^2)(v_i^2 - v_j^2)\}J_1 J_2 J_3 J_4,$$

$$f = \exp\{\Sigma_{i=1}^{k}(-\tfrac{1}{2}u_i^2 - \tfrac{1}{2}v_i^2 + u_i v_i \rho_i r_i) + \psi\}\Pi_{i=1}^{k} u_i v_i,$$

$$\begin{aligned}\psi = &-\Sigma_{i<j}^{k}\Sigma\{(u_i v_i \rho_i r_i + u_j v_j \rho_j r_j)(\tfrac{1}{2}s_{ij}^2 + \tfrac{1}{2}w_{ij}^2 + \tfrac{1}{2}z_{ij}^2)\\ &+ (\rho_i - \rho_j)(u_i v_i r_i - u_j v_j r_j)t_{ij}^2 + (\rho_i - \rho_j)(u_i v_j r_i - u_j v_i r_j)s_{ij}t_{ij}\\ &+ (u_i v_j \rho_j r_i + u_j v_i \rho_i r_j)s_{ij}w_{ij} + (u_i v_j \rho_j r_j + u_j v_i \rho_i r_i)s_{ij}z_{ij}\\ &+ (\rho_i - \rho_j)(u_j v_j r_j - u_i v_i r_i)t_{ij}w_{ij} + (\rho_i - \rho_j)(u_j v_j r_i - u_i v_i r_j)t_{ij}z_{ij}\\ &+ (u_i v_i \rho_i r_j + u_j v_j \rho_j r_i)w_{ij}z_{ij}\}\\ &-\Sigma_{i=1}^{k}\Sigma_{j=k+1}^{p}\{\tfrac{1}{2}u_i v_i \rho_i r_i(w_{ij}^2 + z_{ij}^2) + u_i v_i \rho_i r_j w_{ij} z_{ij}\}\\ &-\tfrac{1}{2}\Sigma_{i=1}^{k}\Sigma_{j=p+1}^{q} u_i v_i \rho_i r_i z_{ij}^2 + \text{terms of higher order}\end{aligned}$$

and Ξ is the image of Π. The maximum value of f is obtained at the single point ξ, in the interior of Ξ, defined by

$$U = V = \operatorname{diag}((1 - \rho_1 r_1)^{-\frac{1}{2}}, \cdots, (1 - \rho_k r_k)^{-\frac{1}{2}})$$

$$S = W_{11} = Z_{11} = O, W_{12} = O, Z_{21} = O, \text{ and } T = O.$$

Theorem 2.1 is now applied to (5.5). The Hessian of $-\ln f$, Δ, reduces to a product of determinants of matrices which are at most 4×4. After simplifying

$$\begin{aligned}\Delta = &\, 2^{2k}\Pi_{i=1}^{k}[(1 - \rho_i r_i)\{\rho_i r_i/(1 - \rho_i r_i)\}^{q-p}]\Pi_{i=1}^{k}\Pi_{j=k+1}^{p}\{(r_i^2 - r_j^2)\rho_i^2(1 - \rho_i r_i)^{-2}\}\\ &\times \Pi_{\alpha=1}^{l}\Pi_{i<j;\, i,j\in Q(\alpha)}\{\rho_i^5(r_i - r_j)^4(r_i + r_j)(1 - \rho_i r_i)^{-3}(1 - \rho_j r_j)^{-3}\}\\ &\times \Pi_{\alpha<\beta}^{l}\Pi_{i\in Q(\alpha)}\Pi_{j\in Q(\beta)}\{(\rho_i r_i - \rho_j r_j)^4(\rho_i^2 - \rho_j^2)(r_i^2 - r_j^2)(1 - \rho_i r_i)^{-4}(1 - \rho_j r_j)^{-4}\}.\end{aligned}$$

The conditions of Theorem 2.1 can be shown to hold and the result then follows.

Partial checks on some rather complicated algebra are provided by the fact that (5.4) agrees with the known asymptotic behavior of the classical hypergeometric function when $k = p = 1$ (see Luke (1969, Section 7.2)) and it agrees with the results of Glynn and Muirhead (1978) when $q_i = 1$, $(1 \leqslant i \leqslant l)$.

An asymptotic representation of the joint density of the squared canonical correlation coefficients is obtained by substituting (5.4) in (1.1). The result is summarized in the following.

THEOREM 5.2. *An asymptotic representation of the joint density of* $r_1^2, \cdots, r_p^2$ *for large* n *when the population coefficients satisfy* (1.2) *is*

$$(5.6)\quad C\Pi_{i=1}^{p}\left\{(1-\rho_i^2)^{\frac{1}{2}n}(r_i^2)^{\frac{1}{2}(q-p-1)}(1-r_i^2)^{\frac{1}{2}(n-q-p-1)}\right\}$$
$$\times\Pi_{i=1}^{k}\left\{(1-\rho_i r_i)^{-n+\frac{1}{2}(p+q-1)}(\rho_i r_i)^{\frac{1}{2}(p-q)}\right\}$$
$$\times\Pi_{\alpha=1}^{l}\Pi_{i<j;\, i,j\in Q(\alpha)}\left\{(r_i^2-r_j^2)(1-\rho_i r_i)^{-\frac{1}{2}}(1-\rho_j r_j)^{-\frac{1}{2}}(r_i+r_j)^{-\frac{1}{2}}\rho_i^{-\frac{1}{2}}\right\}$$
$$\times\Pi_{\alpha<\beta}^{l}\Pi_{i\in Q(\alpha)}\Pi_{j\in Q(\beta)}\left\{(r_i^2-r_j^2)(\rho_i^2-\rho_j^2)^{-1}\right\}^{\frac{1}{2}}$$
$$\times\Pi_{i=1}^{k}\Pi_{j=k+1}^{p}\left\{(r_i^2-r_j^2)\rho_i^{-2}\right\}^{\frac{1}{2}}\Pi_{k+1;\, i<j}^{p}(r_i^2-r_j^2)$$

where

$$C = e^{-nk}\left(\tfrac{1}{2}n\right)^{nk-\frac{1}{2}k(p+q)}\pi^{\frac{1}{2}(p-k)^2+\frac{1}{2}pk}\Gamma_{p-k}\left\{\tfrac{1}{2}(n-k)\right\}$$
$$\times\left[\Gamma_{p-k}\left\{\tfrac{1}{2}(p-k)\right\}\Gamma_{p-k}\left\{\tfrac{1}{2}(q-k)\right\}\Gamma_p\left\{\tfrac{1}{2}(n-q)\right\}\Gamma_k\left(\tfrac{1}{2}n\right)\right]^{-1}$$
$$\times\Pi_{i=1}^{l}\left[\pi^{-\frac{1}{2}q(i)^2}\Gamma_{q(i)}\left\{\tfrac{1}{2}q(i)\right\}(2\pi/n)^{\frac{1}{4}q(i)(q(i)-1)}\right]^{-1}.$$

Furthermore, the convergence is uniform on any set of r_i*'s and* ρ_i*'s such that the* r_i*'s are strictly bounded away from* 1, 0, *and one another and the distinct* ρ_i*'s are similarly bounded.*

An additional check on the results of this section can be obtained by considering the standardized variables

$$x_i = n^{\frac{1}{2}}(r_i^2-\rho_i^2)/\left\{2\rho_i(1-\rho_i^2)\right\} \qquad 1 \leqslant i \leqslant k$$
$$(5.7)\qquad x_j = nr_j^2 \qquad k+1 \leqslant j \leqslant p.$$

Hsu (1941b) gives the limiting joint density of the x_i's, Suguira (1976) extends Hsu's result by computing the terms through order n^{-1} when the ρ_i are all positive, and Fujikoshi (1977) computes the terms through order $n^{-\frac{1}{2}}$ when some of the ρ_i are 0. By a more careful analysis of the integrals used in the proof of Theorem 5.1 one can show that ${}_2F_1^{(p)} = \phi\{1 + O(n^{-1})\}$ where ϕ is the asymptotic representation (5.4). By using this result and making the change of variables (5.7) in (5.6) one obtains Fujikoshi's result.

6. An asymptotic representation—MANOVA case. Assume, without loss of generality, that $\Theta = \mathrm{diag}(\theta_1, \cdots, \theta_p)$ where $\theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_p \geqslant 0$.

THEOREM 6.1. *If* $\Theta = \mathrm{diag}(\theta_1, \cdots, \theta_p)$ *satisfies* (1.5) *then for large* n_2

$$(6.1)\qquad {}_1F_1^{(p)}\{\tfrac{1}{2}(n_1 + n_2); \tfrac{1}{2}n_1; \tfrac{1}{2}n_2\Theta, L\}$$
$$\sim K_1 \Pi_{\alpha<\beta}^m \Pi_{i \in Q(\alpha)} \Pi_{j \in Q(\beta)} \{(l_i - l_j)(\theta_i - \theta_j)\}^{-\frac{1}{2}}$$
$$\times \Pi_{\alpha=1}^l \Pi_{i<j;\, i,j \in Q(\alpha)} \Big[\{(l_i\theta_i)^{\frac{1}{2}}\xi_i - (l_j\theta_j)^{\frac{1}{2}}\xi_j\}\theta_i^{-1}(l_i - l_j)^{-1}\Big]^{\frac{1}{2}}$$
$$\times \Pi_{i=1}^k \{\xi_i^{n(2)+\frac{1}{2}\{n(1)-p+1\}}(l_i\theta_i)^{\frac{1}{4}\{p-n(1)\}}(l_i\theta_i + 4)^{-\frac{1}{4}}\}$$
$$\times \exp\{\tfrac{1}{4}n_2\Sigma_{i=1}^k (l_i\theta_i)^{\frac{1}{2}}\xi_i\}$$

where

$$\xi_i = (l_i\theta_i)^{\frac{1}{2}} + (l_i\theta_i + 4)^{\frac{1}{2}}$$
$$K_1 = e^{-\frac{1}{2}kn(2)} n_2^a \pi^b 2^c \Gamma_k(\tfrac{1}{2}p)\Gamma_k(\tfrac{1}{2}n_1)\big[\Gamma_k\{\tfrac{1}{2}(n_1 + n_2)\}\big]^{-1}$$
$$\times \Pi_{i=1}^l \Big[\pi^{\frac{1}{2}q(i)^2}\Gamma_{q(i)}\{\tfrac{1}{2}q(i)\}^{-1}\Big],$$
$$a = \tfrac{1}{2}k(n_2 - p + \tfrac{1}{2} + \tfrac{1}{2}k) + \Sigma_{i=1}^l q_i(q_i - 1)/4,$$
$$b = -(k^2 + \Sigma_{i=1}^l q_i^2)/4$$
$$c = \tfrac{1}{2}k(2p - \tfrac{1}{2}k - 3n_2 - n_1 - \tfrac{1}{2}) - \tfrac{1}{2}\Sigma_{i=1}^l q_i^2,$$
$$q(i) = q_i,\ n(2) = n_2, \textit{ and } n(1) = n_1.$$

PROOF. The steps in the proof are essentially the same as the steps in the derivation of the asymptotic representation for large sample size of the ${}_2F_1^{(p)}$ function in the joint density of the canonical correlation coefficients.

The ${}_1F_1^{(p)}$ function can be represented as

$$(6.2)\qquad K_2 \textstyle\int_\Lambda h(x)\{f(x)\}^{n(2)} dx$$

where

$$f(x) = \mathrm{etr}\{-\tfrac{1}{2}V^2 + [VH\Theta_1^{\frac{1}{2}}H'Q'L^{\frac{1}{2}} : O]F\}\det V,$$
$$\Theta_1 = \mathrm{diag}(\theta_1, \cdots, \theta_k),$$
$$h(x) = (\det V)^{n(1)-k}\Pi_{i<j}^k (v_i^2 - v_j^2),$$
$$K_2 = (\tfrac{1}{2}n_2)^{\frac{1}{2}k\{n(1)+n(2)\}}\pi^{\frac{1}{2}k^2}2^k\big[\Gamma_k\{\tfrac{1}{2}(n_1 + n_2)\}\Gamma_k(\tfrac{1}{2}k)\big]^{-1},$$
$$\Lambda = D(V) \times G(q_1, \cdots, q_l) \times V(k,p) \times V(k, n_1),$$
$$D(V) = \{\mathrm{diag}(v_1, \cdots, v_k) : v_1 > v_2 > \cdots > v_k > 0\},$$
$$dx = (dV)(dG)(dQ)(dF),$$

$V = \operatorname{diag}(v_1, \cdots, v_k) \in D(V)$, $H \in G(q_1, \cdots, q_l)$, $Q \in V(k, p)$, and $F \in V(k, n_1)$. It follows by the same argument used in Corollary 4.1 that the maximum of f for fixed V is obtained when

$$(6.3)\quad Q = (q_{ij}) = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdots & 0 & \cdots \end{bmatrix}, \qquad F = (f_{ij}) = \begin{bmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \\ \cdots & 0 & \cdots \end{bmatrix}$$

and

$$H = \operatorname{diag}(H_1, \cdots, H_l)$$

where H_i is a $q_i \times q_i$ orthogonal matrix and F and Q satisfy $q_{ii}f_{ii} = 1$, $(1 \leqslant i \leqslant k)$. It follows from this that

$$\max_{V, H, F, Q} f = \max_V \{\max_{H, F, Q} f\}.$$

A straightforward calculation now shows that the maximum value of f is

$$(6.4)\qquad 2^{-k} e^{-\frac{1}{2}k} \Pi_{i=1}^{k} \left[\xi_i \exp\left\{ \tfrac{1}{4} \xi_i (l_i \theta_i)^{\frac{1}{2}} \right\} \right]$$

and the maximum is obtained at the 2^k points of Λ defined by (6.3) and

$$(6.5)\qquad v_i = \tfrac{1}{2}\xi_i \qquad 1 \leqslant i \leqslant k.$$

Using the invariance properties of f, Q, F, and H can be parametrized by writing

$$[Q : -] = \exp(W) = \exp\begin{bmatrix} W_{11} & W_{12} \\ -W_{12}' & O \end{bmatrix}, \qquad W = (w_{ij})$$

$$[F : -] = \exp(Z) = \exp\begin{bmatrix} Z_{11} & -Z_{21}' \\ Z_{21} & O \end{bmatrix}, \qquad Z = (z_{ij})$$

and

$$H = \exp(S), \qquad S = (s_{ij})$$

where $[Q : -]$ is a $p \times p$ orthogonal matrix whose first k columns are Q, $[F : -]$ is an $n_1 \times n_1$ orthogonal matrix whose first k columns are F, W is $p \times p$ skew symmetric, Z is $n_1 \times n_1$ skew symmetric, and S is $k \times k$ skew symmetric with $s_{ij} = 0$ if $i, j \in Q(\alpha)$ for some α, $(1 \leqslant \alpha \leqslant l)$. Making this change of variables leads to the integral

$$(6.6)\qquad 2^k K_2 \textstyle\int_D g f^{n(2)}$$

where

$$g = \Pi_{i=1}^{k} v_i^{n(1)-k} \Pi_{i<j}^{k} (v_i^2 - v_j^2) J(Q \to W) J(F \to Z) J(H \to S),$$

$$f = \exp\left\{ \Sigma_{i=1}^{k} \left(-\tfrac{1}{2} v_i^2 + v_i \theta_i^{\frac{1}{2}} l_i^{\frac{1}{2}} \right) + \Psi \right\} \Pi_{i=1}^{k} v_i,$$

and

$$\begin{aligned}\Psi = & -\Sigma_{i<j}^{k}\Sigma\Big\{\tfrac{1}{2}\Big(v_i\theta_i^{\frac{1}{2}}l_i^{\frac{1}{2}} + v_j\theta_j^{\frac{1}{2}}l_j^{\frac{1}{2}}\Big)\big(w_{ij}^2 + z_{ij}^2\big) \\ & + \Big(\theta_i^{\frac{1}{2}} - \theta_j^{\frac{1}{2}}\Big)\Big(v_i l_i^{\frac{1}{2}} - v_j l_j^{\frac{1}{2}}\Big)\big(s_{ij}^2 - s_{ij}w_{ij}\big) \\ & - \Big(\theta_i^{\frac{1}{2}} - \theta_j^{\frac{1}{2}}\Big)\Big(v_i l_j^{\frac{1}{2}} - v_j l_i^{\frac{1}{2}}\Big)s_{ij}z_{ij} + \Big(v_i\theta_i^{\frac{1}{2}}l_j^{\frac{1}{2}} + v_j\theta_j^{\frac{1}{2}}l_i^{\frac{1}{2}}\Big)w_{ij}z_{ij}\Big\} \\ & - \Sigma_{i=1}^{k}\Sigma_{j=k+1}^{p}v_i\theta_i^{\frac{1}{2}}\Big\{z_{ij}w_{ij}l_j^{\frac{1}{2}} + \tfrac{1}{2}l_i^{\frac{1}{2}}\big(w_{ij}^2 + z_{ij}^2\big)\Big\} \\ & - \tfrac{1}{2}\Sigma_{i=1}^{k}\Sigma_{j=p+1}^{n(1)}v_i\theta_i^{\frac{1}{2}}l_i^{\frac{1}{2}}z_{ij} + \text{higher order terms.}\end{aligned}$$

The maximum value of f is given by (6.4) and is obtained when the v_i satisfy (6.5), $W = O$, $Z = O$, and $S = O$.

It can be shown that the regularity conditions in the extension of Hsu's lemma are satisfied by (6.6). The Hessian of $-\ln f$ is the product of the determinants of matrices which are at most 3×3 and simplifies to

$$\begin{aligned}& 2^{k\{k+2-p-n(1)\}}\Pi_{i=1}^{k}\Pi_{j=k+1}^{p}\theta_i(l_i - l_j) \\ & \times\Pi_{i=1}^{k}\Big[(l_i\theta_i)^{\frac{1}{2}\{n(1)-p\}}(l_i\theta_i + 4)^{\frac{1}{2}}\xi_i^{n(1)+p-2k-1}\Big] \\ & \times\Pi_{\alpha=1}^{l}\Pi_{i<j;\, i,j\in Q(\alpha)}\big\{\theta_i(l_i - l_j)\big(\xi_i^2 - \xi_j^2\big)\big\} \\ & \times\Pi_{\alpha<\beta}^{l}\Pi_{i\in Q(\alpha)}\Pi_{j\in Q(\beta)}\Big[(l_i - l_j)(\theta_i - \theta_j)\Big\{(l_i\theta_i)^{\frac{1}{2}}\xi_i - (l_j\theta_j)^{\frac{1}{2}}\xi_j\Big\}^2\Big].\end{aligned}$$

The theorem now follows by combining these results.

An asymptotic representation of the joint density of $l_1, \cdots, l_p$ is obtained by substituting (6.1) in (1.3). The result is summarized in the following.

THEOREM 6.2. *If $\Omega = n_2\Theta$ and $\theta_1 \geqslant \cdots \geqslant \theta_p$, the latent roots of Θ, satisfy* (1.5) *then an asymptotic representation for large n_2 of the joint density of $l_1, \cdots, l_p$ is*

$$\begin{aligned}(6.7)\quad & K_3\Pi_{i=1}^{p}\Big[l_i^{\frac{1}{2}\{n(1)-p-1\}}(1 - l_i)^{\frac{1}{2}\{n(2)-p-1\}}\Big]\exp\Big[\tfrac{1}{2}n_2\Sigma_{i=1}^{k}\Big\{\tfrac{1}{2}(l_i\theta_i)^{\frac{1}{2}}\xi_i - \theta_i\Big\}\Big] \\ & \times\Pi_{i=1}^{k}\Big[(l_i\theta_i)^{\frac{1}{4}\{p-n(1)\}}(l_i\theta_i + 4)^{-\frac{1}{4}}\xi_i^{n(2)+\frac{1}{2}\{n(1)-p+1\}}\Big] \\ & \times\Pi_{\alpha=1}^{l}\Pi_{i<j;\, i,j\in Q(\alpha)}\Big[(l_i - l_j)\theta_i^{-1}\Big\{(l_i\theta_i)^{\frac{1}{2}}\xi_i - (l_j\theta_j)^{\frac{1}{2}}\xi_j\Big\}\Big]^{\frac{1}{2}} \\ & \times\Pi_{\alpha<\beta}^{m}\Pi_{i\in Q(\alpha)}\Pi_{j\in Q(\beta)}\Big\{(l_i - l_j)(\theta_i - \theta_j)^{-1}\Big\}^{\frac{1}{2}}\Pi_{k+1;\, i<j}^{p}(l_i - l_j)\end{aligned}$$

where $K_3 = KK_1$ and

$$\xi_i = (l_i\theta_i)^{\frac{1}{2}} + (l_i\theta_i + 4)^{\frac{1}{2}}.$$

Furthermore, (6.7) *holds uniformly on any set of l_i's and δ_i's such that the l_i's are strictly bounded away from one another,* 0, *and* 1 *and the δ_i's are bounded and are bounded away from one another and zero.*

A partial check on the previous results can be obtained by considering the following standardized variables

$$(6.8)\qquad x_i = n_2^{\frac{1}{2}}\sigma_i^{-1}\{l_i(1-l_i)^{-1}-\theta_i\} \qquad 1 \leqslant i \leqslant k$$

and

$$x_j = n_2 l_j(1-l_j)^{-1} \qquad k+1 \leqslant j \leqslant p$$

where

$$\sigma_i = (2\theta_i)^{\frac{1}{2}}(\theta_i+2)^{\frac{1}{2}} \qquad 1 \leqslant i \leqslant k.$$

Hsu (1941a) gives the limiting distribution on the x_i's and Fujikoshi obtains the terms of order $n_2^{-\frac{1}{2}}$. Fujikoshi's results can be obtained by making the change of variables (6.8) in (6.7).

A direct consequence of Theorem 6.2 is the following.

COROLLARY 6.1. *An asymptotic representation of the conditional density for large* n_2 *of* $l_{k+1}, \cdots, l_p$ *given* $l_1, \cdots, l_k$ *is*

$$(6.9)\qquad A\Pi_{i=1}^{k}\Pi_{j=k+1}^{p}(l_i-l_j)^{\frac{1}{2}}\Pi_{i=k+1}^{p}\Big[l_i^{\frac{1}{2}\{n(1)-p-1\}}(1-l_i)^{\frac{1}{2}\{n(2)-p-1\}}\Big] \times \Pi_{k+1;\, i<j}^{p}(l_i-l_j)$$

where A *is a constant.*

Note that (6.9) does not depend on $\theta_1, \cdots, \theta_k$. James (1969) has argued, in a different context, that this suggests the use of (6.9) to test the hypothesis H_k, given by (1.6), that the $p-k$ smallest latent roots of Ω are zero. There are two points worth noting.

(i) Constantine and Muirhead (1976) found an asymptotic representation for the joint density of $l_1, \cdots, l_p$ for large $\omega_1, \cdots, \omega_k$ when $\omega_1 > \omega_2 > \cdots > \omega_k > \omega_{k+1} \geqslant \cdots \geqslant \omega_p \geqslant 0$. While their asymptotic representation is markedly different from that given in Theorem 6.2, an asymptotic representation of the conditional density of $l_{k+1}, \cdots, l_p$ given $l_1, \cdots, l_k$ is still given by (6.9).

(ii) By making the following identifications

$$l_i \to r_i^2,\ n_1 \to q,\ p \to p, \text{ and } n_1 + n_2 \to n$$

in (6.9) we obtain an asymptotic representation of the conditional density of the smallest $p-k$ sample canonical correlation coefficients given the largest k canonical coefficients (Glynn and Muirhead (1978)). In this context, $r_1^2, \cdots, r_p^2$ are the sample canonical correlation coefficients based on a sample of size $n+1$ from a $(p+q)$-variate normal distribution.

The following theorem, which is simply a restatement of Theorem 4 of Glynn and Muirhead (1978) for the MANOVA case, is a direct consequence of (ii)

THEOREM 6.3. *The statistic*

$$L_k = -\{n_2 - k + \tfrac{1}{2}(n_1 - p - 1) + \Sigma_{i=1}^{k} l_i^{-1}\}\ln \Pi_{i=k+1}^{p}(1-l_i)$$

is asymptotically χ^2 *on* $(p-k)(n_1-k)$ *degrees of freedom and* $E_c(L_k) = (p-k)(n_1-k) + O(n_2^{-2})$, *where* E_c *denotes expectation with respect to the density* (6.9).

If the observed values of $l_1, \cdots, l_k$ are all near one, then the multiplying factor in L_k is approximately $-\{n_2 + \frac{1}{2}(n_1 - p - 1)\}$, which is the value suggested by Bartlett.

Acknowledgment. The author is very grateful to Professor Robb Muirhead who suggested this area for thesis research and provided much help and encouragement and to the referees for their helpful comments.

REFERENCES

[1] ANDERSON, G. A. (1965). An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix. *Ann. Math. Statist.* **36** 1153–1173.

[2] BARTLETT, M. S. (1947). Multivariate analysis. *J. R. Statist. Soc.* (*Suppl.*) **9** 176–190.

[3] CHATTOPADYAY, 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.) 117–127. Academic Press, New York.

[4] CHATTOPADYAY, 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] CONSTANTINE, A. G. and MUIRHEAD, R. J. (1976). Asymptotic expansions for distributions of latent roots in multivariate analysis. *J. Multivariate Anal.* **6** 369–391.

[7] FUJIKOSHI, Y. (1977). Asymptotic expansions of the distributions of the latent roots in MANOVA and the canonical correlations. *J. Multivariate Anal.* **7** 386–396.

[8] GLYNN, W. J. and MUIRHEAD, R. J. (1978). Inference in canonical correlation analysis. *J. Multivariate Anal.* **8** 468–478.

[9] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.

[10] HSU, P. L. (1941a). On the limiting distribution of roots of a determinantal equation. *J. London Math. Soc.* **16** 183–194.

[11] HSU, P. L. (1941b). On the limiting distribution of the canonical correlations. *Biometrika* **32** 38–45.

[12] HSU, L. C. (1948). A theorem on the asymptotic behavior of a multiple integral. *Duke Math. J.* **15** 623–632.

[13] JAMES, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.* **25** 40–75.

[14] JAMES, A. T. (1964). Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[15] JAMES, A. T. (1969). Tests of equality of latent roots of the covariance matrix. In *Multivariate Analysis II.* (P. R. Krishnaiah, ed.) 205–218. Academic Press, New York.

[16] LUKE, Y. L. (1969). *The Special Functions and Their Approximations, Vol. I.* Academic Press, New York.

[17] MUIRHEAD, R. J. (1978). Latent roots and matrix variates: a review of some asymptotic results. *Ann. Statist.* **6** 5–33.

[18] MURNAGHAN, F. D. (1938). *The Theory of Group Representations.* John Hopkins Press, Baltimore.

[19] SUGIURA, N. (1976). Asymptotic expansions of the distributions of the canonical correlations. In *Commemoration Volume of the Founding of the Faculty of Integrated Arts and Sciences.* Hiroshima Univ.

DEPARTMENT OF STATISTICS
HARVARD UNIVERSITY
ONE OXFORD STREET
CAMBRIDGE, MASSACHUSETTS 02138