Non-central Distributions of the Largest Latent Roots
of Three Matrices in Multivariate Analysis*

by

K.C.S.Pillai and T.Sugiyama**

Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series 129
November, 1967

* This research was supported by the National Science Foundation, Grant No. GP-7663.
** On leave from Aoyama Gakuin University (Japan).
Non-central Distributions of the Largest Latent Roots of Three Matrices in Multivariate Analysis

by

K.C.S. Pillai and T. Sugiyama

Purdue University

1. Introduction and Summary. The cdf of the largest latent root of the generalized E statistic in multivariate analysis in the central case is given by Pillai [9],[10],[11], [13], and also useful formulae [12] approximating at the upper end the cdf of the largest latent root. Further, the above cdf has been obtained by Pillai as a series of incomplete beta functions [9],[10], [14] and also independently by Sugiyama and Fukutomi [17]. Recently, Sugiyama [19] has obtained the cdf of the same, as power series. In the non-central MANOVA case, Hayakawa [7] and Khatri and Pillai [15] have obtained the density in a beta function series form. The purpose of this paper is to find simpler power series expressions than these obtained by the above authors for the non-central density function and the cdf of the largest latent root in the MANOVA situation, both in the generalized beta case and by usual transformation in the generalized F case. We will also obtain similar formulae for the non-central density functions of the largest roots for canonical correlation and equality of two covariance matrices.

2. Non-central distribution of the largest latent root in the MANOVA case. Let \( \sim \) be a \( p \times n_1 \) matrix variate \((p \leq n_1)\) and \( \sim \) a \( p \times n_2 \) matrix variate \((p \leq n_2)\) and the columns be all independently normally distributed with covariance matrix \( \Sigma \sim \), \( E(\sim) = M \) and \( E(\sim) = 0 \). Then it is well known that \( \sim \sim \) is non-central Wishart with \( n_1 \) degrees of freedom and \( \sim \sim \) is central Wishart with

---

1 This research was supported by the National Science Foundation, Grant No. GP-7663.
2 On leave from Aoyama Gakuin University (Japan).
\[ n_2 \text{ degrees of freedom and the covariance matrix } \Sigma, \text{ respectively. The generalized non-central statistics } L \text{ be defined as the latent roots of} \]

\[ L = \left( \frac{U_1 + U_2}{\Sigma_{11} + \Sigma_{22}} \right)^{-\frac{1}{2}} \left( U_1 + U_2 \right)^{-\frac{1}{2}} \]

Let \( 1 > \ell_1 > \ldots > \ell_p > 0 \) be the ordered latent roots of the matrix \( L \), namely the roots of the following determinantal equation

\[ |U_1 - \ell(U_1 + U_2)| = 0, \]

then the joint density function of \( \ell_1, \ldots, \ell_p \) is given by Constantine [2], James [6].

\[
C(p, n_1, n_2) \exp(\text{tr} - \Omega) \left| \frac{\ell_1 - \ell_{p-1}}{\frac{1}{2}(n_1 - p)} \right| \left| I - \frac{1}{2}(n_2 - p) \right| \prod_{i<j} (\ell_i - \ell_j) 
\]

\[
(1) \sum_{k=0}^{\infty} \sum_{k} \frac{(n_1 + n_2)/2}{\binom{n_1}{k}} \frac{C_k(\Omega)C_k(L)}{C_k(I)k!} 
\]

where \( \Omega \) is the non-centrality matrix, \( \frac{1}{2} M' \Sigma^{-1} M \), determinants \( |L| \) and \( |I-L| \) expressed as products of the latent roots of their matrices,

\[ C(p, n_1, n_2) = \pi^{p^2/2} \frac{\Gamma_p((n_1 + n_2)/2)/\Gamma(p/2)\Gamma(n_1/2)\Gamma(n_2/2)}{\Gamma_p(n_1/2)\Gamma_p(n_2/2)} \text{ and } C_k(L) \text{ are zonal polynomials defined in [4], [5]. In this section, we obtain first the density and c.d.f. of } L. \text{ In this connection, we state below two lemmas:}

Lemma 1. Let \( L \) be a diagonal matrix with diagonal elements \( 1 > \ell_2 > \ldots > \ell_m > 0 \), and let \( k \) be a partition of \( k \). Then
\[
\int \prod_{i=2}^{m} (l_i - l_1) \prod_{i<j} (l_i - l_j) \prod_{i=2}^{m} \, dl_i \\
\geq l_2 \geq \ldots \geq l_m > 0
\]

\[
= (mt + k) \left( \frac{\Gamma_m(m/2)}{\pi^{m/2}} \right) \left( \frac{\Gamma_m(t, \kappa)\Gamma_m((m+1)/2, \kappa)}{\Gamma_m(t+(m+1)/2, \kappa)} \right) C_k(I_m).
\]

Lemma 2. Let \( S(p \times p) \) be a symmetric matrix, and \( C_k(S) \) and \( C_\sigma(S) \) be zonal polynomials of degree \( k \) and \( \sigma \) respectively corresponding to the partition

\( \kappa = (k_1 \geq k_2 \geq \ldots \geq k_p \geq 0) \) and \( \sigma = (s_1 \geq s_2 \geq \ldots \geq s_p \geq 0) \). Then

\[
C_k(S) C_\sigma(S) = \sum_{\delta} g_{k, \sigma}^\delta C_\delta(S),
\]

where

\( \delta = (\delta_1 \geq \delta_2 \geq \ldots \geq \delta_p \geq 0), \sum_{i=1}^{p} \delta_i = k+s \) and \( g_{k, \sigma}^\delta \) are constants.

Lemma 1 has been discussed by Sugiyama [18] and [19]. Tables of the coefficients \( g_{k, \sigma}^\delta \) of Lemma 2 are given by Hayakawa [7] and Khatri and Pillai [15] for various values of \( k \) and \( s \).

Now using Lemma 2

\[
|I-L|^{(p-1)/2} C_k(L)
\]

\[
= \sum_{s=0}^{\infty} \sum_{\sigma} ((p+1-n_2)/2) \sigma \sum_{k} C_k(L) / s!
\]

\[
= \sum_{s=0}^{\infty} \sum_{\sigma} \sum_\delta ((p+1-n_2)/2) \sigma \delta \sum_{k} g_{k, \sigma}^\delta C_\delta(L) / s!,
\]

and from (1), we get
(2) \[ c(p, n_1, n_2) \exp(\text{tr}-\Omega) \sum_{i<j} (\frac{\pi}{2}) \sum_{k=0}^{\infty} \frac{((n_1+n_2)/2)_{\infty}}{(n_1/2)_{\kappa}} \frac{c_k(\Omega)}{c_k(\mathcal{I})} \]

\[ \sum_{s=0}^{\infty} \sum_{\sigma, \delta} e_{\kappa, \sigma} (\frac{(p+1-n_2)/2}{\sigma}) c_\delta(\mathcal{I})/s! \]

Now consider the integral

(3) \[ \int \sum_{\ell_i > \ell_j > \ldots > \ell_p > 0} (\frac{n_1-p-1}{2}) \frac{c_\delta(\mathcal{L}) (\ell_i \ell_j) \prod_{i=2}^{p} \delta \ell_i}{(\ell_i \ell_j) \prod_{i=2}^{p} \delta \ell_i} \]

In (3) transform \( q_i = \ell_i / q_1 \), \( i = 2, \ldots, p \) and integrate with respect to \( q_2, \ldots, q_p \), we get

(4) \[ \int_{-\infty}^{\infty} \frac{n_1/2-(p+1)/2}{\prod_{i=2}^{p} (1-q_i) \prod_{i=2}^{p} (q_i-q_j)} c_{\kappa, 1-q_2} \prod_{i=2}^{p} (1-q_i) \prod_{i=2}^{p} (q_i-q_j) \delta \ell_i \]

\[ = \int_{-\infty}^{\infty} \frac{n_1/2+k+s-1}{\prod_{i=2}^{p} (1-q_i) \prod_{i=2}^{p} (q_i-q_j)} \cdot \frac{\Gamma_p((p/2))}{\Gamma_p((n_1+p+1)/2, \delta)} \frac{c_\delta(\mathcal{I})}{c_{\kappa, 1-q_2}} \]

Hence, from (2) and (4) we have the following formula

\[ c(p, n_1, n_2) \exp(\text{tr}-\Omega) \left( \frac{\Gamma_p((p/2)) \Gamma_p((p+1)/2)}{\pi^{p/2}} \right) \]

\[ \sum_{k=0}^{\infty} \frac{((n_1+n_2)/2)_{\infty}}{(n_1/2)_{\kappa}} \frac{c_k(\Omega)}{c_k(\mathcal{I})} \sum_{s=0}^{\infty} \frac{pn_1/2+k+s-1}{\prod_{i=2}^{p} (1-q_i) \prod_{i=2}^{p} (q_i-q_j)} \delta \ell_i \]

\[ \cdot \left( \frac{(p+1-n_2)/2}{\sigma} \frac{n_1/2}{\delta} \right) c_\delta(\mathcal{I}) \]
Further, noting that \( \Gamma_p(a, \delta) = \Gamma_p(a)(a)_\delta \), we obtain the density of the largest latent root in the following form

\[
c_1(p, n_1, n_2) \sum_{k=0}^{\infty} \frac{\binom{n_1+n_2}{k}}{(n_1/2)_k} \frac{C_k(n)}{C_k(I)k!} \sum_{s=0}^{\infty} e_{k, \sigma}^\delta \frac{((p+1-n_2)/2)(n_1/2)_\delta}{((n_1+p+1)/2)_\delta}
\]

where \( 1 > \ell_1 > 0 \), and

\[
c_1(p, n_1, n_2) = \frac{\Gamma_p((p+1)/2)\Gamma_p((n_1+n_2)/2)}{\Gamma_p(n_2/2)\Gamma_p((n_1+p+1)/2)} e^{tr-O}
\]

Further, the c.d.f. of the largest latent root is given by

\[
F(\ell_1 < x) = c_1(p, n_1, n_2) \sum_{k=0}^{\infty} \frac{\binom{n_1+n_2}{k}}{(n_1/2)_k} \frac{C_k(n)}{C_k(I)k!} \sum_{s=0}^{\infty} e_{k, \sigma}^\delta \frac{((p+1-n_2)/2)(n_1/2)_\delta}{((n_1+p+1)/2)_\delta} c_\sigma(I_p) x^{p_1/2+k+s}
\]

Let \( O = 0 \) in (9). Then, since \( e_{0, \sigma}^\delta = 1 \) and \( \delta = \sigma \), we obtain the following formula given by Sugiyama [19]

\[
c_1(p, n_1, n_2) \sum_{k=0}^{\infty} \frac{\binom{n_1+n_2}{k}}{(n_1/2)_k} \frac{C_k(n)}{C_k(I)k!} \sum_{s=0}^{\infty} e_{k, \sigma}^\delta \frac{((p+1-n_2)/2)(n_1/2)_\delta}{((n_1+p+1)/2)_\delta} c_\sigma(I_p) x^{p_1/2+k+s}
\]

And also, in (7), let \( n_2 = p+1 \), and \( x=1 \). Then we have \( O_F(O) = e^{tr-O} \).

Since the roots \( \ell_1, \ldots, \ell_p \) of the generalized beta case are related to the roots \( f_1, \ldots, f_p \) of the generalized F case in the following manner:
\[ \ell_1 = \frac{f_1}{1 + f_1}, \ldots, \ell_p = \frac{f_p}{1 + f_p}, \]

we obtain from (9), the c.d.f. of the largest latent root in the non-central generalized F case in the form

\[
P(f_1 < y) = C_1 (p, n_1, n_2) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{((n_1 + n_2)/2)_k}{(n_1/2)_\kappa} \cdot \sum_{s=0}^{\infty} \sum_{\gamma, \delta} \frac{((p+1-n_2)/2)_{(n_1/2)_\delta}}{s!((n_1+p+1)/2)_\delta} \\
\cdot C_{\delta} (I_{1/\delta}) (y/(1+y))^{pn_1/2+k+s}.
\]

THEOREM 1. Let \( \sim \) be the matrix having non-central Wishart distribution with \( n_1 \) degrees of freedom and matrix of non-centrality parameter \( \Omega \), and \( \sim \) be the matrix having the Wishart distribution with \( n_2 \) degrees of freedom. Then the pdf and the cdf of the largest latent root \( \ell_1 \) of the equation

\[ |U_1 - (U_1 + U_2) \sim | = 0 \]

is given by (6) and (7). And the cdf of the largest latent root \( f_1 \) of the equation

\[ |U_1 - U_2 f | = 0 \]

is given by (9).
3. Distribution of the largest latent root in the canonical correlation case. Let the columns of \( \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \) be \( n \) independent normal \((p+q)\)-dimensional variates \((p \leq q)\) with zero means and covariance matrix

\[
\Sigma = \begin{pmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{12}^t & \Sigma_{22}
\end{pmatrix}.
\]

Let \( R \) be the diagonal matrix with diagonal element \( r_1, r_2, \ldots, r_p \), where \( r_1^2, r_2^2, \ldots, r_p^2 \) are the latent roots of the equation

\[
|X_1 X_2' (X_2 X_2')^{-1} - r^2 X_2 X_2'| = 0,
\]

and also \( P \) be the diagonal matrix with diagonal elements \( \rho_1, \rho_2, \ldots, \rho_p \), where \( \rho_1^2, \rho_2^2, \ldots, \rho_p^2 \) are the latent roots of the equation

\[
|\sum_{12} \sum_{22}^{-1} \rho^2 \sum_{11} | = 0.
\]

Then, the distribution of \( r_1^2, r_2^2, \ldots, r_p^2 \), is given by Constantine [2] in the following form

\[
C(n,p,q) \left| \frac{I-R^2}{\Sigma} \right|^{n/2} \left| \frac{R^2}{\Sigma} \right|^{(p-1)/2} \left| \frac{I-R^2}{\Sigma} \right|^{(n-p-1)/2}
\]

\[
\prod_{i<j} (r_i^2 - r_j^2) \sum_{k=0}^{\infty} \frac{(n/2)_k(n/2)_k}{(q/2)_k} \frac{c_k(R^2)}{c_k(I_p)} \frac{c_k(\Sigma)}{c_k(I_p)^k},
\]

where \( C(n,p,q) = \frac{\Gamma_p(n/2)p^{2/2}}{\Gamma_p(p/2\sigma_p)\Gamma_p(p/2(n-q))\Gamma_p(p/2p)} \).
By the same method as before, namely using lemmas (1) and (2), we have the following

\[
\int_{r_1^2 > r_2^2 > \ldots > r_p^2 > 0} \left| \mathcal{F} \frac{(q-p-1)/2}{|z-F|} \right| \frac{(n-q-p-1)/2}{|z^2|} \prod_{i=2}^{p} \frac{r_i^2}{\hat{r}_i} \, dr_1^2 \ldots dr_p^2
\]

\[= \sum_{s=0}^{\infty} \sum_{\sigma, \delta} g_{\delta}^{\sigma} \left( \frac{(p+q+1-n)/2}{s!} \right) \cdot \left( \frac{pq/2+k+s}{\pi^{2/2}} \right) \cdot \frac{\Gamma_p((q/2, \delta)) \Gamma_p((p+1)/2)}{\Gamma_p((q+p+1)/2, \delta)} \cdot \frac{c_{\delta}(I_p)}{c_{\delta}(I_p^2) \cdot (r_1^2)_{pq/2+k+s-1}}.\]

(11)

Hence, from (10) and (11), we have the following formula

\[c_2(n, p, q) \left| \mathcal{F} \frac{n/2}{|z-F|} \right| \frac{1}{(q/2)_k} \sum_{k=0}^{(n/2)_k} \frac{(n/2)_k}{(q/2)_k} \frac{c_k(I^2)}{c_{\delta}(I^2)_k} \]

\[= \sum_{s=0}^{\infty} \sum_{\sigma, \delta} g_{\delta}^{\sigma} \left( \frac{(pq/2+k+s)/s!}{s!} \right) \sum_{\sigma, \delta} g_{\delta}^{\sigma} \left( \frac{(pq+1-n)/2}{s!} \right) \]

\[\cdot \frac{(q/2)_\delta}{((q+p+1)/2)_\delta} \cdot \frac{c_{\delta}(I_p)}{c_{\delta}(I_p^2) \cdot (r_1^2)_{pq/2+k+s-1}}.\]

(12)

where

\[c_2(n, p, q) = \frac{\Gamma_p((p+1)/2) \Gamma_p(n/2)}{\Gamma_p((n-q)/2) \Gamma_p((q+p+1)/2)} .\]
Integrating (12) from 0 to \( x \) with respect to \( r_1^2 \), we have the following c.d.f of the largest latent root in the canonical correlation case

\[
P(\lambda_1^2 < x) = c_2(n, p, q) \left| I_n \right|^\frac{n}{2} \sum_{k=0}^{\infty} \sum_{\lambda} \frac{n/2 \cdot k/2}{(q/2)_k} \frac{C_{\lambda}(\lambda^2)}{C_{\lambda}(\lambda^2)_k!}
\]

\[
\sum_{s=0}^{\infty} e^{k_s} ((p+q+1-n)/2) \frac{(q/2)_s}{((q+1)/2)_s} \frac{C_{\lambda}(\lambda^2)}{s!} \cdot x^{pq/2+k+s}.
\]

THEOREM 2. Let \( \frac{X_1}{X_2} \) be \( n \) independent normal \( (p+q) \)-dimensional variates \( (p \leq q) \) with zero means and covariance matrix, \( \Sigma \). Then the pdf and the cdf of the largest latent root \( r_1^2 \) of the equation

\[
|X_1 X_2' (X_2 X_2')^{-1} X_2 X_2' - r_1^2 X_1 X_1'| = 0
\]

is given by (12) and (13).

4. Non-central distribution of the largest latent root for test of equality of two covariance matrices. Let \( \Sigma_1 \) and \( \Sigma_2 \) be independently distributed as Wishart \( W(n_1, p, \Sigma_1) \) and \( W(n_2, p, \Sigma_2) \), respectively. Let the latent roots of \( \Sigma_1 \) and \( \Sigma_2 \) be denoted \( \delta_1, \ldots, \delta_p \). Let \( \omega_i = \lambda g_i/(1+\lambda g_i) \), \( i = 1, \ldots, p \),
where $\lambda$ is a given positive constant in the test of the null-hypothesis $H$ that $\lambda \Lambda = I$ and $\Lambda = \text{diag} \left( \delta_1', \ldots, \delta_p' \right)$. Then the joint distribution of $\omega_i$ is given by Khatri [8] in the following form

$$
\mathcal{C}(p, n_1, n_2) \left[ \lambda \Lambda \right]^{-n_1/2} \left[ \sum_{k=0}^{\infty} \sum_{k} \frac{C_k(I-(\lambda \Lambda)^{-1}) C_k(W)}{C_k(I) k!} \right] \sum_{i<j} (\omega_i - \omega_j)
$$

where $W = \text{diag} \left( \omega_1, \ldots, \omega_p \right)$. Then, by the same method as before, we can obtain the density function of the largest latent root $\omega_1$ in the following form

$$
\mathcal{C}_3(p, n_1, n_2) \left[ \lambda \Lambda \right]^{-n_1/2} \sum_{k=0}^{\infty} \sum_{k} \frac{C_k(I-(\lambda \Lambda)^{-1})}{C_k(I) k!} \sum_{s=0}^{\infty} \sum_{\sigma, \delta} \frac{e_\delta^{\sigma} \frac{((p+1-n_2)/2, n_1/2)}{s!((n_1+p+1)/2)^{\delta}}} {C_3(I) \omega_1^{pn_1/2+k+s-1}}
$$

(14)

$$
\sum_{s=0}^{\infty} \sum_{\sigma, \delta} \frac{e_\delta^{\sigma} \frac{((p+1-n_2)/2, n_1/2)}{s!((n_1+p+1)/2)^{\delta}}} {C_3(I) \omega_1^{pn_1/2+k+s-1}}
$$

where $1 > \omega_1 > 0$, and $\mathcal{C}_3(p, n_1, n_2) = \frac{\Gamma_p((p+1)/2) \Gamma_p((n_1+n_2)/2)}{\Gamma_p(n_2/2) \Gamma_p((n_1+p+1)/2)}$.

Let $\lambda \Lambda = I$, namely the central case. Then, since $e_0^{\delta} = 1$ and $\sigma = \delta$, the cdf of $\omega_1$ is

$$
P(\omega_1 < x) = \mathcal{C}_3(p, n_1, n_2) \frac{\Gamma_1((p+1-n_2)/2, n_1/2; (n_1+p+1)/2; \omega_1 I_p)}{\omega_1^{pn_1/2}}
$$

This is the same formula given by Sugiyama [18]. We note that if $(p+1-n)/2$ is an integer, the summation of $s$ will be terminated in finite terms.
Further, \( \theta_{K,\sigma} \)'s are constants which do not exceed unity \([14]\). Again when \( n_2 = p + 1 \) we get

\[
P(\omega_1 < x) = |\lambda_{\Omega}|^{-\frac{1}{2}n_1 - \frac{1}{2}} I_{0}(\frac{1}{2}n_1; x(I - (\lambda_{\Omega})^{-1})) x^{\frac{1}{2}p n_1}.
\]

Let \( x=1 \), \( a=\frac{1}{2}n_1 \), and \( \Omega=I-(\lambda_{\Omega})^{-1} \). Then we have \( I_{0}(a; \Omega) = |I-\Omega|^{-a} \).

**Theorem 3.** Let \( \sim_{1} \) and \( \sim_{2} \) are the matrices having Wishart distributions \( W(n_1,p,\Sigma_{1}) \) and \( W(n_2,p,\Sigma_{2}) \), respectively. Then the pdf of \( \omega_1 = \lambda \sigma_1 / (1+\lambda \sigma_1) \), where \( \sigma_1 \) is the largest latent root of the equation

\[
|\sim_{1} - \sigma \sim_{2}| = 0
\]

is given by \([14]\).

It may be pointed out that Khatri \([8]\) has given the density of \( \sigma_1 \) but \([14]\) does not follow from his result by transformation.
References


