

On Asymptotic Expansions of the Distributions of
The Characteristic Roots of Two Matrices *

by

Tseng-chi Chang

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 178

January 1969

* This research was partly supported by the National Science Foundation Contract No. GP-7663. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Asymptotic representation of the distribution of the
 characteristic roots of $S_n S_n^{-1}$ when roots are not all distinct

by

T.C. Chang

ERRATA

<u>Page</u>	<u>Line</u>	<u>Incorrect</u>	<u>Correction</u>
2	1	$\prod_{i=1}^p$	$\prod_{i < j}^p$
3	16	$o[\phi_n(x)]$	$o[\phi_n(x)]$
	20	$o(\phi_n(x))$	$o[\phi_n(x)]$
4	3	$o(\phi(x))$	$o[\phi(x)]$
	13-14	$\cos 2\theta]^{-\frac{n}{2}}$	$\cos 2\theta]^{-\frac{n}{2}}$
14	1	a_i 's	a_1 's
15	4	λ_i 's	λ_1 's
16	10	1 2 ... p	1, 2, ..., p
	12	1 2 ... p	1, 2, ..., p
17	6	\tilde{H}^{**}	\tilde{H}^*
18	6	A Δ	A ~
19	18	$\tilde{K} =$	$\tilde{K} =$
20	1	$:=$	$:$
	13	$ A' A ^{-\frac{1}{2}}$	$ A' A ^{-\frac{1}{2}}$
21	last line	\emptyset	(
25	last line	$:=$	$:$
27	5	$a _p$	$a _p$
	10-11	(1+0)	1+0

33	1-2	$\frac{\pi^{\frac{1}{2}kp}}{\Gamma_k(\frac{p}{2})}$	$\frac{\Gamma_k(\frac{p}{2})}{\pi^{\frac{1}{2}kp}}$
	12-13	$\frac{\pi^{\frac{1}{2}kp}}{\Gamma_k(\frac{p}{2})}$	$\frac{\Gamma_k(\frac{p}{2})}{\pi^{\frac{1}{2}kp}}$
34	9-10	$\pi^{\frac{1}{2}p(p-1) + \frac{1}{2}kp}$	$\pi^{\frac{1}{2}p(p-1) - \frac{1}{2}kp}$
	10-11	$\Gamma_k(\frac{p}{2})$	$[\Gamma_k(\frac{p}{2})]^{-1}$
47	19	$\mu_i = \frac{w_i}{w_p}$	$\mu_i = \frac{w_i}{w_p}$
52	7	l_i 's	l_i 's
53	1	Thomson	Thompson
	2	.2.5	2.5
54	18	Nanda	Nanda
56	4	$(x; 2, k+1)$	$(x; q, k+1)$

CHAPTER I

ASYMPTOTIC REPRESENTATION OF THE DISTRIBUTION OF THE
CHARACTERISTIC ROOTS OF $S_1 S_2^{-1}$ IN THE DISTINCT ROOTS CASE1. Introduction and Summary

Let $S_i: p \times p (i=1,2)$ be independently distributed as Wishart (n_i, p, Σ_i) . Let the characteristic roots of $S_1 S_2^{-1}$ and $\Sigma_1 \Sigma_2^{-1}$ be denoted by $l_i (i=1,2,\dots,p)$ and $\lambda_i (i=1,2,\dots,p)$ respectively such that $l_1 \geq l_2 \geq \dots \geq l_p \geq 0$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$. Then the distribution of l_1, \dots, l_p can be expressed in the form (Khatri [15])

$$(1.1) \quad c |\Lambda|^{-\frac{1}{2}n_1} |L|^{-\frac{1}{2}(n_1-p-1)} \prod_{i < j} (l_j - l_i) \int_{O(p)} |I + \Lambda^{-1} \underline{L} H L H'|^{-\frac{1}{2}(n_1+n_2)} (H' dH)$$

where

$$(1.2) \quad c = 2^{-p} \prod_{i=1}^p \Gamma_p\left(\frac{1}{2}i\right) \Gamma_p\left(\frac{n_1+n_2}{2}\right) \left\{ \Gamma_p\left(\frac{1}{2}p\right) \Gamma_p\left(\frac{1}{2}n_1\right) \Gamma_p\left(\frac{1}{2}n_2\right) \right\}^{-1},$$

$$(1.3) \quad \Gamma_p(t) = \prod_{j=1}^p \Gamma\left(t - \frac{1}{2}j + \frac{1}{2}\right),$$

$L = \text{diag}(l_1, l_2, \dots, l_p)$ and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$. The invariant Haar measure

$$(1.4) \quad (\tilde{H}' d\tilde{H}) = \prod_{i=1}^p h_i' dh_j$$

is defined on the group $O(p)$ of $p \times p$ orthogonal matrices with h_i and dh_j the i th and j th column of \tilde{H} and $d\tilde{H}$ respectively (James [11]). The group $O(p)$ has a volume

$$(1.5) \quad V(p) = \int_{O(p)} (\tilde{H}' d\tilde{H}) = 2^p \prod_{i=1}^p \frac{1}{4^i} (p+1) / \prod_{i=1}^p \Gamma(\frac{1}{2}i).$$

One of the approaches to the evaluation of the integral of (1.1) (from here on we denote it as E) is to expand it as a power series:

$$(1.6) \quad E = \int_{O(p)} \left| \tilde{I} + \tilde{\Lambda}^{-1} \tilde{H} \tilde{L} \tilde{H}' \right|^{-\frac{1}{2}(n_1+n_2)} (\tilde{H}' d\tilde{H})$$

$$= V(p) \sum_{k=0}^{\infty} \frac{1}{k!} \sum_K \frac{C_K(\tilde{\Lambda}^{-1}) C_K(\tilde{L})(n_1+n_2)_K}{C_K(\tilde{I})}$$

with zonal polynomial $C_K(\tilde{T})$ of any $p \times p$ symmetric matrix \tilde{T} defined in James [14]. But its convergence is very slow unless the characteristic roots of the argument matrices are small. In the one sample case G.A. Anderson [1] has obtained a gamma type asymptotic expansion for the distribution of the characteristic roots of the estimated covariance matrix. In this chapter we obtain a beta type

asymptotic representation of the roots distribution of $S_1 S_2^{-1}$ involving linkage factors between sample roots and corresponding population roots. A study is also made of the approximation to the distribution of w_1, \dots, w_p where $w_i = l_i / (1 + l_i)$ ($i=1, 2, \dots, p$). If the roots are distinct the limiting distribution as n_2 tends to infinity has the same form as that of G.A. Anderson [1]. If, moreover, n_1 is assumed also large, then it agrees with Girshick's result [7].

In the following we will assume $l_1 > l_2 > \dots > l_p > 0$ and $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$, and for the simplification of notations we let $A = \Lambda^{-1}$, i.e., $a_i = 1/\lambda_i$ ($i=1, \dots, p$), $0 < a_1 < a_2 < \dots < a_p < \infty$, and $n = n_1 + n_2$.

The definitions below are taken from Erdélyi [6].

DEFINITION. The sequence of functions $\{\phi_n(x)\}$ is an asymptotic sequence as $x \rightarrow \infty$, if for each n

$$\phi_{n+1}(x) = o[\phi_n(x)] \quad \text{as } x \rightarrow \infty,$$

Let $\{\phi_n\}$ be an asymptotic sequence.

DEFINITION. The (formal) series $\sum a_n \phi_n(x)$ is an asymptotic expansion to N terms of $f(x)$ as $x \rightarrow \infty$ if

$$f(x) = \sum_{n=0}^N a_n \phi_n(x) + o(\phi_N(x)) \quad \text{as } x \rightarrow \infty.$$

This is written $f(x) \sim \sum_{n=0}^N a_n \phi_n(x)$.

DEFINITION. The function $\phi(x)$ is an asymptotic representation for $f(x)$ as $x \rightarrow \infty$ if

$$f(x) = \phi(x) + o(\phi(x)) \quad \text{as } x \rightarrow \infty.$$

This is written $f(x) \sim \phi(x)$.

2. The Asymptotic Representation of E when p = 2

An elementary method is used in this section to obtain the asymptotic expansion of E when p = 2.

Let $O^+(2) = \{H \in O(2), |H| = \pm 1\}$ then

$$(2.1) \quad E = 2 \int_{O^+(2)} \left| \frac{I + A H L H'}{\sqrt{1 + A H L H'}} \right|^{\frac{n}{2}} (H' dH).$$

Now let $H = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad -\pi < \theta \leq \pi,$

So that $(H' dH) = d\theta$ and

$$(2.2) \quad E = 4 \left[(1 + a_1 \ell_1) (1 + a_2 \ell_2) \right]^{\frac{n}{2}} \int_{-\pi/2}^{\pi/2} \left[1 + \frac{1}{2} C_{12} (1 - \cos 2\theta) \right]^{\frac{n}{2}} d\theta$$

where

$$(2.3) \quad C_{12} = \frac{(a_2 - a_1)(\ell_1 - \ell_2)}{(1 + a_1 \ell_1)(1 + a_2 \ell_2)}.$$

The integrand of (2.2) has a maximum of unity at $\theta = 0$ and decreases to $(1 + \frac{1}{2}c_{12})^{-\frac{1}{2}n}$ at $\theta = \pm \frac{\pi}{2}$. Write (2.2) as

$$(2.4) \quad 4 \left[\prod_{i=1}^2 (1 + a_i l_i) \right]^{-\frac{n}{2}} \int_{-\pi/2}^{\pi/2} \exp \left\{ -\frac{n}{2} \log \left[1 + \frac{1}{2}c_{12} (1 - \cos 2\theta) \right] \right\} d\theta.$$

Since the integral is mostly concentrated in a small neighborhood of the origin, for large n , we can expand the argument of the exponential function and $\cos 2\theta$ in the usual power series and set the limit to be $\pm \infty$ (see Erdélyi [6]). Thus for large degrees of freedom E is approximately

$$(2.5) \quad 4 \left[\prod_{i=1}^2 (1 + a_i l_i) \right]^{-\frac{n}{2}} \int_{-\infty}^{\infty} \exp \left\{ -\frac{n}{2} c_{12} \theta^2 \right\} d\theta \left\{ 1 + O\left(\frac{1}{n}\right) \right\}$$

or

$$(2.6) \quad E \simeq 4 \left[\prod_{i=1}^2 (1 + a_i l_i) \right]^{-\frac{n}{2}} \left(\frac{2\pi}{nc_{12}} \right)^{\frac{1}{2}} \left\{ 1 + O\left(\frac{1}{n}\right) \right\}.$$

3. The Asymptotic Representation of E in General Case

Lemma (1.1) Let A and L are defined as before then $f(H) = \left| \frac{I + AHLH'}{I + AL} \right|$ $H \in O(p)$ attains its identical minimum value $\left| \frac{I + AL}{I + AL} \right|$ when H is of the form

$$(3.1) \quad \tilde{H} = \begin{pmatrix} +1 & & & & \\ & +1 & & & 0 \\ & & \dots & & \\ & & & & +1 \\ 0 & & & & & +1 \end{pmatrix}$$

$$\begin{aligned} \text{Proof: } df &= d \left| \tilde{I} + \underbrace{AHLH'} \right| \\ &= d \left| \tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right| \\ &= \text{tr} \left(\tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right)^{-1} \left(\underbrace{A^{\frac{1}{2}}dHLH'A^{\frac{1}{2}}} + \underbrace{A^{\frac{1}{2}}HLdH'A^{\frac{1}{2}}} \right) \\ &= 2 \text{tr} \underbrace{LH'A^{\frac{1}{2}}} \left(\tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right)^{-1} \underbrace{A^{\frac{1}{2}}HH'dH}. \end{aligned}$$

Note that $\tilde{H}'dH$ is a skew symmetric matrix, therefore, $df = 0$ implies that $\underbrace{LH'A^{\frac{1}{2}}} \left(\tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right)^{-1} \underbrace{A^{\frac{1}{2}}H}$ is a symmetric matrix. But $\underbrace{H'A^{\frac{1}{2}}} \left(\tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right)^{-1} \underbrace{A^{\frac{1}{2}}H}$ is itself a symmetric matrix and \tilde{L} is a diagonal matrix with distinct positive roots, so $\underbrace{H'A^{\frac{1}{2}}} \left(\tilde{I} + \underbrace{A^{\frac{1}{2}}HLH'A^{\frac{1}{2}}} \right)^{-1} \underbrace{A^{\frac{1}{2}}H}$ has to be a diagonal matrix, say \tilde{D} . Thus $\tilde{I} = \underbrace{A^{\frac{1}{2}}H}(\tilde{L}-\tilde{D}^{-1})\underbrace{H'A^{\frac{1}{2}}}$. This can happen only if \tilde{H} is of the form with ± 1 in one position in a column or a row and zero in other positions. After substituting those stationary values into $f(\tilde{H})$ we obtain a general form

$$(3.2) \quad \prod_{i=1}^p (1 + a_i l_{\sigma_i}),$$

where l_{σ_i} is any permutation of $l_i (i = 1, \dots, p)$. It is easy to see that (3.2) attains its minimum value when $l_{\sigma_i} = l_i (i = 1, 2, \dots, p)$. Or $f(\tilde{H})$ attains its identical minimum value $\left| \tilde{I} + \underbrace{AL} \right|$ when H is of the form of (3.1).

The above lemma enables us to claim that, for large n_1 , the integrand of E is negligible except for small neighborhoods about each of these matrices of (3.1) and E consists of identical contributions from each of these neighborhoods so that

$$(3.3) \quad E \simeq 2^p \int_{N(\tilde{I})} \left| \tilde{I} + \underbrace{AHLH'} \right|^{-\frac{n}{2}} (H' dH),$$

where $N(\tilde{I})$ is a neighborhood of the identity matrix on the orthogonal manifold. Since any proper orthogonal matrix can be written as the exponential of a skew symmetric matrix we transform E under

$$(3.4) \quad \tilde{H} = \exp \tilde{S}, \quad \tilde{S} \text{ a } p \times p \text{ skew symmetric matrix,}$$

So that $N(\tilde{I}) \rightarrow N(\tilde{S} = 0)$. The Jacobian of this transformation has been computed by G.A. Anderson [1],

$$(3.5) \quad J = 1 + \frac{p-2}{24} \operatorname{tr} \tilde{S}^2 + \frac{8-p}{4 \cdot 6!} \operatorname{tr} \tilde{S}^4 + \dots$$

Direct substitution of (3.4) into (3.3) yields

$$(3.6) \quad \left| \tilde{I} + \underbrace{AHLH'} \right|^{-\frac{n}{2}} \\ = \left| \tilde{I} + \underbrace{AL} + \underbrace{ASL} - \underbrace{ALS} + \underbrace{ALS^2} - \underbrace{ASLS} + \dots \right|^{-\frac{n}{2}} \\ = \left| \tilde{I} + \underbrace{AL} \right|^{-\frac{n}{2}} \left| \tilde{I} + (\tilde{I} + \underbrace{AL})^{-1} (\underbrace{ASL} - \underbrace{ALS} + \underbrace{ALS^2} - \underbrace{ASLS} + \dots) \right|^{-\frac{n}{2}}.$$

Lemma (1.2) For any $p \times p$ matrix B and its characteristic roots $b_i (i=1, \dots, p)$, if $\max_{1 \leq i \leq p} |b_i| < 1$ then

$$|I+B|^{-\frac{1}{2}n} = \exp \left\{ -\frac{1}{2}n \operatorname{tr} \left(B - \frac{B^2}{2} + \frac{B^3}{3} - \dots \right) \right\}.$$

Proof: $|I+B|^{-\frac{1}{2}n} = \exp \left\{ -\frac{1}{2}n \log \prod_{i=1}^p (1+b_i) \right\}.$

If $\max_{1 \leq i \leq p} |b_i| < 1$ then

$$|I+B|^{-\frac{1}{2}n} = \exp \left\{ -\frac{1}{2}n \operatorname{tr} \left(B - \frac{B^2}{2} + \frac{B^3}{3} - \dots \right) \right\}.$$

Apply lemma (1.2) to (3.6) and the maximum characteristic roots of $(I+AL)^{-1} (ALS - ALS + \dots)$ can be shown to be less than unity. Since we are only interested in the first term we need to investigate the group of terms up to order of S^2 which is denoted by $\{S^2\}$. Let $R = (I+AL)^{-1}$, then

$$(3.7) \quad \operatorname{tr}\{S^2\} = \operatorname{tr} \left[R(ALS^2 - ASLA) - \frac{1}{2} (RALSRA + RASLRA - RALSRA - RALSRA) \right].$$

after simplification (3.7) reduces to

$$(3.8) \quad \operatorname{tr} \left[R(ALS^2 - ASLS) - (LS - SL) RALSRA \right]$$

or

$$(3.9) \quad \operatorname{tr}\{S^2\} = \sum_{i < j}^p c_{ij} s_{ij}^2$$

where

$$(3.10) \quad C_{ij} = (a_j - a_i) (l_i - l_j) / [(1 + a_i l_i)(1 + a_j l_j)].$$

Direct substitution into E yields

$$(3.11) \quad E \simeq 2^p \prod_{i=1}^p (1 + a_i l_i)^{-\frac{n}{2}} \int_{N(S=0)} \exp\left\{-\frac{n}{2} \sum_{i < j}^p C_{ij} s_{ij}^2\right\} \prod_{i < j} ds_{ij} \left\{1 + O\left(\frac{1}{n}\right)\right\}.$$

For large n the limits for each s_{ij} can be put to $\pm \infty$. We finally have the following theorem.

Theorem: The asymptotic distribution of the roots, $l_1 > l_2 > \dots > l_p > 0$, of $S_1 S_2^{-1}$ for large degrees of freedom $n = n_1 + n_2$ when the roots of $\Sigma_1 \Sigma_2^{-1}$ are $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$ and $a_i = 1/\lambda_i$ ($i=1, \dots, p$), is given by

$$(3.12) \quad C 2^p \prod_{i < j} (l_j - l_i) \prod_{i=1}^p [l_i^{\frac{n_1 - p - 1}{2}} a_i^{\frac{1}{2} n_1} (1 + a_i l_i)^{-\frac{(n_1 + n_2)}{2}}] \prod_{i < j}^p \left[\frac{2\pi}{C_{ij} (n_1 + n_2)} \right]^{\frac{1}{2}}.$$

The asymptotic formula shows that the distribution function of a group of adjacent roots is sensitive only to those other roots which are close to them.

4. A Dual Expansion of E and Some Remarks

If we let $\tilde{W} = L(\tilde{I} + \tilde{L})^{-1}$ in (1.1), i.e., $w_i = l_i / (1 + l_i)$ ($i=1, 2, \dots, p$) where $\tilde{W} = \text{diag}(w_1, \dots, w_p)$, then the joint distribution of w_i 's is given by

$$(4.1) \quad C |\tilde{\Lambda}|^{-\frac{1}{2}n_1} |\tilde{W}|^{\frac{1}{2}(n_1-p-1)} \left| \frac{I-W}{\tilde{\Lambda}} \right|^{-\frac{1}{2}(n_1+p+1)} \prod_{i < j} (w_j - w_i) \\ \int_{O(p)} \left| \frac{I+AHLH'}{\tilde{\Lambda}} \right|^{-\frac{1}{2}(n_1+n_2)} (H' dH) \\ 1 > w_1 > w_2 > \dots > w_p > 0.$$

Application of lemmas (1.1) and (1.2) to (4.1) yields its asymptotic representation

$$(4.2) \quad 2^p C |\tilde{\Lambda}|^{\frac{1}{2}n_1} |\tilde{W}|^{\frac{1}{2}(n_1-p-1)} \left| \frac{I-W}{\tilde{\Lambda}} \right|^{\frac{1}{2}(n_2-p-1)} \\ \prod_{i < j} (w_j - w_i) \prod_{i=1}^p [1 + (a_i - 1)w_i]^{-\frac{1}{2}(n_1+n_2)p} \prod_{i < j} \left[\frac{2\pi}{C_{ij}^*(n_1+n_2)} \right]^{\frac{1}{2}}$$

$$\text{when } C_{ij}^* = \frac{(a_j - a_i)(w_i - w_j)}{[1 + (a_i - 1)w_i][1 + (a_j - 1)w_j]}.$$

Now let us proceed to look at (3.12) once again. The asymptotic distribution of characteristic roots of $S_1 S_2^{-1}$ given there can be rewritten as

$$(4.3) \quad F_1(A) \prod_{i < j} (\ell_i - \ell_j)^{\frac{1}{2}} \prod_{i=1}^p \left[\ell_i^{\frac{n_1-p-1}{2}} (1+a_i \ell_i)^{-\frac{(n_1+n_2)}{2} + p-1} \right] \prod_{i=1}^p d\ell_i$$

where $F_i(A)$ ($i=1,2,3$) depend on a_i but not on ℓ_i ($i=1,2,\dots,p$).

If we make $g_i = \ell_i/n_2$ ($i=1,2,\dots,p$) and let n_2 tends to infinity