JOURNAL OF MULTIVARIATE ANALYSIS 12, 1–38 (1982)

Some Limit Theorems for the Eigenvalues of a Sample Covariance Matrix

DAG JONSSON

Uppsala University, Uppsala, Sweden

Communicated by P. R. Krishnaiah

Limit theorems are given for the eigenvalues of a sample covariance matrix when the dimension of the matrix as well as the sample size tend to infinity. The limit of the cumulative distribution function of the eigenvalues is determined by use of a method of moments. The proof is mainly combinatorial. By a variant of the method of moments it is shown that the sum of the eigenvalues, raised to k-th power, $k = 1, 2,..., m$ is asymptotically normal. A limit theorem for the log sum of the eigenvalues is completed with estimates of expected value and variance and with bounds of Berry–Esseen type.

1. INTRODUCTION

Let X_{ij}, $i = 1, 2,..., p$, $j = 1, 2,..., n$, be independent normal variables with zero mean and unit variance. (1.1)

Let $Y_{rs} = \sum_{j=1}^{n} X_{rj}X_{sj}$ denote elements of a $p \times p$ matrix, $\mathbf{S}_p^{(n)}$, i.e., $Y_{rs} = Y_{sr}$. Then $\mathbf{S}_p^{(n)}$ has a central Wishart distribution with n degrees of freedom (d.f.). (1.2)

One aim of this paper is to study the asymptotic behaviour of the eigenvalues of $\mathbf{S}_p^{(n)}$ when p and n both tend to infinity, in a way indicated by Arharov [2]. Different functions of the eigenvalues used as test criteria in multivariate analysis are of special interest. Examples of such functions are the sum of the eigenvalues $=$ trace $\mathbf{S}_p^{(n)}$ and the product of the eigenvalues $= |\mathbf{S}_p^{(n)}|$ (the determinant).

Arharov [2] assumed condition (1.1). With a technique introduced by Arnold [3a], Arharov's results can be proved under weaker conditions on the variables X_{ij}. They need not be normal and their moments need not all be finite.

Received January 13, 1978; revised August 6, 1981.

AMS 1980 subject classifications: Primary 62E20, 60F05; Secondary 62H10.

Key words and phrases: Eigenvalues of a sample covariance matrix, cumulative distribution function, method of moments, limit theorems, sums of eigenvalues, generalized variance.

0047-259X/82/010001-38$02.00/0

First, let $n \to \infty$ and then $p \to \infty$. When $n \to \infty$, for fixed p, the matrix $(1/\sqrt{n})(\mathbf{S}_p^{(n)} - n\mathbf{I}_p)$ ($\mathbf{I}_p$ is the $p \times p$ identity matrix) converges in distribution to a stochastic matrix $\mathbf{Z}_p$, whose elements Z_{rs} are normally and for $r \leqslant s$ independently distributed, with $EZ_{rs} = 0$ and

$$\begin{aligned} \operatorname{Var} Z_{rs} &= 2 \qquad \text{if } r = s \\ &= 1 \qquad \text{if } r \neq s,\ r = 1, 2,..., p,\ s = 1, 2,..., p. \end{aligned}$$

Now, let $W_p(x)$ denote the cumulative distribution function (c.d.f.) of the eigenvalues of the matrix $(1/2\sqrt{p})\,\mathbf{Z}_p$, i.e.,

$$W_p(x) = 1/p\ \{\text{the number of eigenvalues} \leqslant x\}.$$

According to Wigner's Semi-Circle Law [16a, b] $\lim_{p\to\infty} EW_p(x) = W(x)$, where W is an absolutely continuous distribution function with density

$$\begin{aligned} w(x) &= (2/\pi)\sqrt{1-x^2} \qquad \text{for } |x| \leqslant 1 \\ &= 0 \qquad\qquad\qquad\quad \text{for } |x| > 1. \end{aligned} \tag{1.3}$$

Under the same conditions, Grenander [8, pp. 177–180] has shown that $W_p(x)$ converges to $W(x)$ in probability, while Arnold [3a] has proved that the convergence is valid with probability 1.

For the purpose of standardization we consider the matrix $(1/n)\,\mathbf{S}_p^{(n)}$ rather than $\mathbf{S}_p^{(n)}$. Let $\lambda_{p1}^{(n)} \leqslant \lambda_{p2}^{(n)} \leqslant \cdots \leqslant \lambda_{pp}^{(n)}$ be the eigenvalues of $(1/n)\,\mathbf{S}_p^{(n)}$ and denote the corresponding c.d.f. by $F_p^{(n)}(x)$. Further, let $\lambda_p^{(n)}$ be a stochastic root of $(1/n)\,\mathbf{S}_p^{(n)}$. This means that $\lambda_p^{(n)}$ is a randomly chosen eigenvalue. Suppose $p = p(N)$ and $n = n(N)$ depend on a variable N, so that

$$p(N) \text{ and } n(N) \to \infty, \qquad \frac{p(N)}{n(N)} \to y, \qquad 0 \leqslant y < \infty, \text{ when } N \to \infty. \tag{1.4}$$

Under these conditions together with (1.1) Arharov [2] has given the first five moments of the asymptotic distribution of $\lambda_p^{(n)}$. However, Marčenko and Pastur [13, Example 1, pp. 511–512] have determined the distribution by use of Stieltjes transforms. They studied matrices of the form $\sum_{j=1}^{n} \mathbf{S}_j$, where $\mathbf{S}_1, \mathbf{S}_2,..., \mathbf{S}_n,...$ are i.i.d. $p \times p$ matrices with elements $Y_{rs}^{(j)} = X_{rj} \cdot X_{sj}$. Evidently $\mathbf{S}_p^{(n)}$ has this form. Independently, Grenander and Silverstein [9] and Wachter [15] have reached the same result. Grenander and Silverstein determined the asymptotic distribution without first stating the limit moments explicitly. Wachter's result is a more general one: Theorem 2.1 of the present paper could be seen as a special case of Wachter's Theorem 1. However, as the methods of this paper on many points differ from the other ones in the mentioned papers, we shall carry through a proof, which is mainly combinatorial, by use of a method of moments (cf. Arharov [2]). In the first theorem it is shown that $F_p^{(n)}(x)$ converges in pobability to the asymptotic distribution, mentioned above. Note that this implies convergence in distribution for the variable $\lambda_p^{(n)}$. The results of this paper were given in 1976 in a technical report by the author [10].

Some preliminaries are needed. Suppose (1.1) is true. Then $EY_{rs} = n \cdot \delta_{rs}$, where Y_{rs} was defined in (1.2). Also, Var Y_{rs} is equal to $2n$ or n according as $r = s$ or $r \neq s$. If $r \neq t$, $s \neq u$ or $r \neq u$, $s \neq t$, then $\operatorname{Cov}(Y_{rs}, Y_{tu}) = 0$. If the Y's belong to the same row or the same column, they are not independent but they are still uncorrelated. If $r \neq s$, then $EY_{rs}^{2k-1} = 0$ and $EY_{rs}^{2k} \sim c_k n^k$ for $k = 1, 2, \ldots$; here c_k is a constant depending on k. Further, $EY_{rr}^k \sim n^k$, $k = 1, 2, \ldots$. Also, $EY_{r_1r_2} \cdot Y_{r_2r_3} \cdot \cdots \cdot Y_{r_{k-1}r_k} \cdot Y_{r_kr_1} = n$ if $r_1, r_2, \ldots, r_k$ are all different. The following terminology will be used repeatedly. The element Y_{rs} as well as the pair (r, s) is said to be *diagonal* if $r = s$ and *off diagonal* if $r \neq s$. Then, for arbitrarily chosen indices $r_1, r_2, \ldots, r_k$,

$$EY_{r_1r_2} \cdot Y_{r_2r_3} \cdot \cdots \cdot Y_{r_{k-1}r_k} \cdot Y_{r_kr_1} \tag{1.5}$$

is equal to the sum of the number of diagonal elements and the number of "independent closed index chains" (an index chain is said to be closed if the first index = the last index) of off diagonal elements.

Notations. $\mathbf{S}_p^{(n)}$, $p(N)$, $n(N)$, $\lambda_{pi}^{(n)}$, $\lambda_p^{(n)}$, etc., will be denoted by $\mathbf{S}$, p, n, λ_i, λ, etc., when there is no danger of confusion.

For discussions about the applications of distributions of eigenvalues of random matrices, the reader is referred to Carmeli [4] and Krishnaiah [12].

2. Limit Theorems for the C.D.F.

THEOREM 2.1. *Under conditions* (1.1) *and* (1.4), $F_p^{(n)}(x)$ *converges in probability to a distribution function* $F_y(x)$ *for every* x. *For* $0 < y \leqslant 1$ *this is a continuous distribution with density*

$$\begin{aligned} f_y(x) &= \frac{\sqrt{(x-a(y))(b(y)-x)}}{2\pi yx} \quad \text{for} \quad a(y) < x < b(y), \\ & \qquad\qquad a(y) = 1 + y - 2\sqrt{y}, \\ & \qquad\qquad b(y) = 1 + y + 2\sqrt{y} \\ &= 0 \quad \text{otherwise}. \end{aligned} \tag{2.1}$$

For $1 < y < \infty$ *it is a mixed distribution with density*

$$f_y(x) = \frac{\sqrt{(x-a(y))(b(y)-x)}}{2\pi yx} \quad \text{for} \quad a(y) < x < b(y),$$

with probability mass for $x = 0$

$$f_y(0) = 1 - \frac{1}{y},$$

and

$$f_y(x) = 0$$

otherwise. For $y = 0$ *it is the degenerated distribution* δ_1. (This case is not regarded in the sequel.)

Remark 2.1. If $y = 1$, the density is

$$f_1(x) = \frac{1}{2\pi} \cdot \sqrt{\frac{4-x}{x}}, \qquad 0 < x < 4. \tag{2.2}$$

This is a general $\beta(\frac{1}{2}, \frac{3}{2})$-distribution over the interval $(0, 4)$. A distribution is said to be general $\beta(r, s)$ over the interval (c, d) if it is continuous with density

$$f(x) = \frac{\Gamma(r+s)}{\Gamma(r) \cdot \Gamma(s)} \cdot \frac{(x-c)^{r-1} \cdot (d-x)^{s-1}}{(d-c)^{r+s-1}}, \qquad c < x < d$$

$$= 0 \qquad \text{otherwise.}$$

In the general case $(0 < y < \infty)$, $x \cdot f_y(x) = \sqrt{(x - a(y))(b(y) - x)}/2\pi y$, which is the density of a general $\beta(\frac{3}{2}, \frac{3}{2})$-distribution over $(a(y), b(y))$.

Proof of Theorem 2.1. The moments of the limit distribution (2.1) are related to the moments of the $\beta(\frac{3}{2}, \frac{3}{2})$-distribution in the following way.

Let $X \in \beta(\frac{3}{2}, \frac{3}{2})$ (over $(0, 1)$), i.e., X has the density

$$f_X(x) = \frac{8}{\pi} \cdot \sqrt{x(1-x)}, \qquad 0 < x < 1.$$

Then $X' = 2(X - \frac{1}{2})$ has density (1.3).

According to Wigner [16a]

$$\begin{aligned} EX'^k &= 0 \qquad \text{if } k \text{ is odd,} \\ &= \frac{k!}{2^k \cdot \left(\frac{k}{2}\right)! \left(\frac{k}{2}+1\right)!} \qquad \text{if } k \text{ is even; } k = 1, 2, \ldots. \end{aligned} \tag{2.3}$$

Set $Y = 2\sqrt{y}\, X' + y + 1 = 4\sqrt{y}(X - \frac{1}{2}) + y + 1$. It can be verified that Y

has a general $\beta(\frac{3}{2}, \frac{3}{2})$-distribution over $(a(y), b(y))$. The kth moment of Y is $(k = 1, 2, \ldots)$

$$EY^k = \sum_{s=0}^{k} \binom{k}{s} 2^s \cdot EX'^s y^{s/2} \cdot (y+1)^{k-s}$$

$$= \sum_{s=0}^{[k/2]} \binom{k}{2s} \binom{2s}{s} \frac{1}{s+1} \cdot y^s \cdot (y+1)^{k-2s}.$$

Denote by $\alpha_k(y)$ the kth moment of distribution (2.1). From Remark 2.1 it is seen that

$$\alpha_k(y) = EY^{k-1} = \sum_{s=0}^{[(k-1)/2]} \binom{k-1}{2s} \binom{2s}{s} \frac{1}{s+1} \cdot y^s \cdot (y+1)^{k-1-2s}. \tag{2.4}$$

This is a polynomial in y of degree $k-1$. The coefficient of y^r is

$$a_{k,r} = \sum_{s=0}^{r} \binom{k-1}{2s} \binom{k-1-2s}{r-s} \binom{2s}{s} \frac{1}{s+1},$$

$$r = 0, 1, \ldots, k-1 \qquad \left(\binom{u}{v} = 0 \text{ if } u < v\right).$$

Note that $a_{k,r} = a_{k,k-1-r}$.

The following lemma gives a simplification.

LEMMA 2.1.

$$a_{k,r} = \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r}, \qquad r = 0, 1, \ldots, k-1; k = 1, 2, \ldots. \tag{2.5}$$

Proof.

$$a_{kr} = \sum_{s=0}^{r'} \frac{(k-1)!}{(2s)!\,(k-1-2s)!} \cdot \frac{(k-1-2s)!}{(r-s)!\,(k-1-r-s)!} \cdot \frac{(2s)!}{s!\,(s+1)!},$$

$$r' = \min(r, k-r-1);$$

$$= \sum_{s=0}^{r'} \left\{ \frac{\binom{k-r-1}{s}\binom{r+1}{r-s}}{\binom{k}{r}} \right\} \cdot \binom{k}{r}\binom{k-1}{r} \cdot \frac{1}{r+1}$$

$$= \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r},$$

as the terms inside the brackets define a hypergeometric distribution.

Thus, the next lemma is proved.

LEMMA 2.2. *The kth moment of limit distribution* (2.1) *is*

$$\alpha_k(y) = \sum_{r=0}^{k-1} a_{k,r} y^r = \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} y^r, \qquad k = 1, 2, \ldots . \tag{2.6}$$

Consider the special case $y = 1$. If X' has a distribution with density (1.3), then $X'^2 \in \beta(\frac{1}{2}, \frac{3}{2})$, i.e., $4X'^2$ has a general $\beta(\frac{1}{2}, \frac{3}{2})$-distribution over $(0, 4)$. Consequently, by (2.3)

$$\alpha_k(1) = E(4X'^2)^k = 4^k \cdot \frac{(2k)!}{4^k \cdot k!\,(k+1)!} = \binom{2k}{k} \frac{1}{k+1}. \tag{2.7}$$

(This can also be calculated from (2.4).)

Let M_k be the kth moment in regard of $F_p^{(n)}(x)$, i.e.,

$$M_k = \frac{1}{p} \sum_{i=1}^{p} \lambda_i^k. \tag{2.8}$$

It will be shown that M_k converges in probability to $\alpha_k(y)$, $k = 1, 2, \ldots$. It will further be shown that the moment sequence $\{\alpha_k(y)\}_{k=1}^{\infty}$ uniquely determines distribution (2.1). We find that

$$\begin{aligned} EM_k &= E\frac{1}{p} \sum_{i=1}^{p} \lambda_i^k = \frac{1}{pn^k} \cdot E \operatorname{trace} \mathbf{S}^k \\ &= \frac{1}{pn^k} \sum_{r_1=1}^{p} \sum_{r_2=1}^{p} \cdots \sum_{r_k=1}^{p} EY_{r_1 r_2} \cdot Y_{r_2 r_3} \cdot \cdots \cdot Y_{r_{k-1} r_k} \cdot Y_{r_k r_1} \qquad (2.9) \\ &= \frac{1}{n^k} \sum_{r_2=1}^{p} \sum_{r_3=1}^{p} \cdots \sum_{r_k=1}^{p} EY_{1 r_2} \cdot Y_{r_2 r_3} \cdot \cdots \cdot Y_{r_{k-1} r_k} \cdot Y_{r_k 1} \end{aligned}$$

(by symmetrical reasons).

By Wigner [16a] it is seen that only terms containing even powers of the variables X_{ij} are relevant. Accordingly it is sufficient to study index combinations, for which every index appears an even number of times. $EY_{1r_2} \cdot Y_{r_2 r_3} \cdot \cdots \cdot Y_{r_k 1}$ is a polynomial in n. The terms in (2.9) may suitably be divided according to the degree of the polynomial. For every degree the terms are then divided according to the number of different indices. Hence

$$n^k EM_k = \sum_{r_2=1}^{p} \sum_{r_3=1}^{p} \cdots \sum_{r_k=1}^{p} EY_{1r_2} \cdot Y_{r_2 r_3} \cdot \cdots \cdot Y_{r_k 1}$$

TABLE I

Number of diagonal elements	Type of index pairs (every chain consists of k pairs)	"Leading term"
k	$(1,1)(1,1)\cdots(1,1)$	n^k
$k-2$	$(1,1)(1,1)\cdots(1,1)(1,r)(r,r)(r,r)\cdots(r,r)$ $\cdots(r,1)(1,1)(1,1)\cdots(1,1)$	$n^{k-1}\cdot p$
$k-3$	$(1,1)\cdots(1,1)(1,r)(r,r)\cdots(r,r)(r,s)(s,s)$ $\cdots(s,s)(s,1)(1,1)\cdots(1,1)$	$n^{k-2}\cdot p^2$
$k-4$	[a] $\cdots(1,r)\cdots(r,s)\cdots(s,t)\cdots(t,1)\cdots$	$n^{k-3}\cdot p^3$
	or	
	$\cdots(1,r)\cdots(r,1)\cdots(1,s)\cdots(s,1)$	$n^{k-2}\cdot p^2$
Etc.		

[a] The diagonal pairs are omitted.

is a polynomial in n and p. Obviously the total degree must be k. Owing to the standardization (division by n^k), terms with total degree lower than k may be neglected. In order to determine the coefficients of n^k, $n^{k-1}\cdot p,\ldots, n\cdot p^{k-1}$, we shall, to begin with, assort the terms in Table (I) according to the number of diagonal elements involved (see (1.5)). It is evident that except for the diagonal elements, there are one or more sequences of elements with a closed index chain, i.e., with index pairs of the type $(1,r)(r,s)(s,1)$, $(u,r)(r,s)(s,t)(t,u)$, etc., where the first index is equal to the last index. Note that in order to determine the maximal degree, it is sufficient to study chains with every off diagonal pair (r,s) appearing once. If, for example, a chain contains the subchain $(1,r)(r,s)(s,1)$, then it is assumed that these three pairs do not appear at any other place in the chain. The number of chains with such repeating pairs may be neglected, but diagonal pairs may occur more than once. This means that the off diagonal pairs of the "indispensable" terms in (2.9) form a closed chain, with all its pairs different, starting and ending with the index "1." The chain may possibly contain one or more closed subchains. (2.10)

EXAMPLE 2.1. The chain $(1,r)(r,s)(s,t)(t,r)(r,1)$ contains the closed subchains $(1,r)(r,1)$ and $(r,s)(s,t)(t,r)$, while the chain $(1,r)(r,s)(s,t)(t,u)(u,1)$ cannot be split up into smaller closed chains.

One more restriction can be made. Consider chains of off diagonal pairs divided into closed subchains, which cannot be further reduced. It is then sufficient to handle chains, whose subchains have the following properties. The second index in all the pairs, except the last one, is a "new" index, i.e., it is an index, which has not appeared before in any subchain. The last pair is a "closing" one, i.e., its second index equals the first index of the subchain. (2.11)

Consider chains divided into maximally reduced closed subchains of off diagonal pairs. Denote every pair in the subchain, except the last one, with an F. Denote the last pair by R_i, where i equals the number of F's preceding R_i in the subchain. So far we have not considered the diagonal pairs. Let these be represented by 0's.

EXAMPLE 2.2. (a) $(1, r)(r, 1)(1, 1)(1, 1)$ is denoted by $FR_1\,00$.

(b) $(1, 1)(1, r)(r, r)(r, s)(s, 1)$ is denoted by $0F0FR_2$.

(c) $(1, r)(r, r)(r, r)(r, s)(s, s)(s, t)(t, r)(r, 1)$ becomes $F00F0FR_2R_1$.

Note that the chain of off diagonal pairs $(1, r)(r, s)(s, t)(t, r)(r, 1)$ contains two closed subchains, viz. $(1, r)(r, 1)$ and $(r, s)(s, t)(t, r)$, i.e., in symbols FR_1 and FFR_2.

Now consider possible arrangements of F's and R's with the restrictions (2.10) and (2.11) taken into account. Evidently, a sequence with r F's may contain at most r R's. Furthermore, among the first m signs in a sequence, whatever value of m, the number of F's is greater than or equal to the sum of the R-indices. When the whole sequence is considered, we have equality in this relation. (2.12)

EXAMPLE 2.3. A sequence with two F's can be built up in the following different ways (the 0's are disregarded).

FFR_1R_1 (chain type $(1, r)(r, s)(s, r)(r, 1)$),

FR_1FR_1 (chain type $(1, r)(r, 1)(1, s)(s, 1)$),

FFR_2 (chain type $(1, r)(r, s)(s, 1)$).

But, for example, FR_2F and FR_1R_1F are not possible, since conditions (2.12) are not fulfilled. There are no corresponding index chains in those cases.

The number of F's corresponding to an index chain is equal to the number of different indices. With the number of F's as a basis of division of possible sequences, we can set up Table II.

Generally: Every sign sequence with the corresponding "leading term" $n^{k-r} \cdot p^r$ consists of r F's, s R's: $R_{i_1}, R_{i_2}, \ldots, R_{i_s}$ and $k-r-s$ 0's. The integers $i_1, i_2, \ldots, i_s$ satisfy $i_1 + i_2 + \cdots + i_s = r$. The number of different arrangements in view of (2.12) is equal to the coefficient of $n^{k-r} \cdot p^r$ in the polynomial $n^k \cdot EM_k$.

TABLE II

Number of F's	Sequence of signs (every sequence consists of k signs	Corresponding "leading term"
0	$00 \cdots 0$	n^k
1	$0 \cdots 0F0 \cdots 0R0 \cdots 0$	$n^{k-1} \cdot p$
2	[a] $\cdots F \cdots F \cdots R_2 \cdots$ or $\cdots F \cdots F \cdots R_1 \cdots R_1 \cdots$ or $\cdots F \cdots R_1 \cdots F \cdots R_1 \cdots$	$n^{k-2} \cdot p^2$
3	[a] $\cdots F \cdots F \cdots F \cdots R_3 \cdots$ or $\cdots F \cdots F \cdots F \cdots R_2 \cdots R_1 \cdots$ or $\cdots F \cdots F \cdots F \cdots R_1 \cdots R_2 \cdots$ or $\cdots F \cdots F \cdots R_2 \cdots F \cdots R_1 \cdots$	$n^{k-3} \cdot p^3$
Etc.	Etc.	

[a] The 0's are omitted.

EXAMPLE 2.4. Suppose $k=4$. Then, for example, the coefficient of $n^2 \cdot p^2$ is 6. The possible arrangements are: $FFR_2 0$, $FF0R_2$, $F0FR_2$, $0FFR_2$, FR_1FR_1, FFR_1R_1.

A general formula is given by the following lemma.

LEMMA 2.3. *The coefficient of $n^{k-r} \cdot p^r$ is equal to $a_{k,r}$, i.e., according to* (2.5) *it is equal to*

$$\frac{1}{r+1}\binom{k}{r}\binom{k-1}{r}, \qquad k=1,2,\ldots;\ r=0,1,\ldots,k-1.$$

Proof. (1) r F's (indistinguishable) can be distributed over k positions in $\binom{k}{r}$ different ways.

(2) When the r F's have been placed out, we have to determine the number of different ways in which $R_{i_1}, R_{i_2}, \ldots, R_{i_s}$ can be distributed over the remaining $k-r$ positions under variation of $i_1, i_2, \ldots, i_s$ (positive integers) with $i_1 + i_2 + \cdots + i_s = r$ for $s = 1, 2, \ldots, r$ (R's with equal indices are indistinguishable). But this is equal to the number of different ways in which r balls can be distributed over $k-r$ cells. This number is $\binom{k-1}{r}$ (Feller [7a, p. 38]). The empty positions are filled with 0's.

(3) The total number of different arrangements of r F's, $R_{i_1}, R_{i_2}, \ldots, R_{i_s}$ (under variation of the indices) and $k-r-s$ 0's is by (1) and (2), $\binom{k}{r}\binom{k-1}{r}$. We shall see that these arrangements in a natural way can be divided into proper subsets with $r+1$ members each, of which exactly one is a possible sequence, i.e., (2.12) is fulfilled.

(4) Consider an arbitrary sequence with r F's, s R's: $R_{i_1}, R_{i_2},\ldots, R_{i_s}$ and $k-r-s$ 0's. Fix the F's and denumerate the surrounding $r+1$ spaces in the following way

$$\text{Sp. } 1, F, \text{Sp. } 2, F,\ldots, \text{Sp. } r, F, \text{Sp. } r+1.$$

If the contents of the spaces are shifted cyclically we generate r additional sequences, which all appear among the $\binom{k}{r}\binom{k-1}{r}$ arrangements in (3).

(5) The $r+1$ sequences within every cycle are different. In order to show this, let a_i denote the index sum of the R's being in space i, $i=1,2,\ldots,r+1$. If the sequences were not different, the cyclical suite $a_1, a_2,\ldots, a_{r+1}$ would be periodical. But $\sum_{i=1}^{r+1} a_i = r$ (= the total index sum), while the suite consists of $r+1$ elements. Then r and $r+1$ would have the same integer divisor, which is impossible.

(6) In every cycle there is at least one possible sequence, for which (2.12) is fulfilled. If the F's are replaced with 1's, we find that a sequence is possible if and only if the partial sums of the corresponding suite

$$-a_1, 1, -a_2, 1, -a_3,\ldots, -a_r, 1, -a_{r+1} \tag{2.13}$$

are all non-negative.

Set $b_i = -(a_1 + a_2 + \cdots + a_i)$, $i=1,2,\ldots,r+1$. Then the partial sums are $b_1, b_1+1, b_2+1, b_2+2,\ldots, b_r+r-1, b_r+r, b_{r+1}+r$, i.e., of type b_i+i-1 or b_i+i. Suppose some of them are negative. Let k be the largest index for which $b_k + k = \min_{1\leqslant i\leqslant r+1}(b_i+i)$ $(k<r+1$, as $b_{r+1}+(r+1)=1)$. Then

$$\begin{aligned} b_k + k &\leqslant b_i + i - 1 \qquad \text{for} \quad i = k+1, k+2,\ldots, r+1 \\ &\leqslant b_i + i \qquad\quad\ \ \text{for} \quad i = 1, 2,\ldots, k. \end{aligned}$$

The suite $-a_{k+1}, 1, -a_{k+2}, 1,\ldots, 1, -a_{r+1}, 1, -a_1, 1,\ldots, 1, -a_k$ has only non-negative partial sums, for

$$-(a_{k+1} + a_{k+2} + \cdots + a_i) + (i-k-1) = (b_i + i - 1) - (b_k + k) \geqslant 0$$

for $i = k+1,\ k+2,\ldots, r+1$, and

$$\begin{aligned} &-(a_{k+1} + a_{k+2} + \cdots + a_{r+1} + a_1 + \cdots + a_i) + (r-k+i) \\ &\quad = (b_{r+1} + r) + (b_i + i) - (b_k + k) = (b_i + i) - (b_k + k) \geqslant 0 \end{aligned}$$

for $i = 1, 2,\ldots, k$.

(7) In every cycle there is exactly one possible sequence, i.e., the corresponding suite has all its partial sums non-negative. Suppose this is true for $-a_1, 1, -a_2, 1,\ldots, 1, -a_{r+1}$ as well as for $-a_k, 1, -a_{k+1}, 1,\ldots, 1, -a_{r+1},$

$1, -a_1, 1,..., 1, -a_{k-1}$. From the first suite it follows that $b_{k-1}+k-2 \geqslant 0$ or $b_{k-1} \geqslant -(k-2)$, while the second one gives $b_{r+1}-b_{k-1}+r-k+1 \geqslant 0$ or $b_{k-1} \leqslant -(k-1)$ (since $b_{r+1}+r=0$), which means a contradiction.

(8) Thus, the number of possible sequences is $(1/(r+1)) \cdot \binom{k}{r}\binom{k-1}{r} = a_{k,r}$, and the proof of the lemma is completed.

There is a direct connection with an urn model: Consider an urn with k balls, of which r are marked with 1's, $k-r-s$ are marked with 0's and the rest with $-i_1, -i_2,..., -i_s$, respectively. Take out balls with equal probabilities one at a time, without replacement. Let $V_1, V_2,..., V_k$ be the successive obtained numbers. Then $P(V_1+V_2+\cdots+V_u \geqslant 0,\ u=1,2,..., k)$ equals the probability that a randomly chosen suite (2.13) has only non-negative partial sums $= 1/(r+1)$, by (4)–(7) (cf. Karlin [11, pp. 244–249, 268]).

To sum up, from Lemma 2.3 it follows that

$$EM_k = \frac{1}{n^k}\left\{\sum_{r=0}^{k-1} a_{k,r} \cdot n^{k-r} \cdot p^r + \text{terms in } n \text{ and } p \text{ with total degree} < k\right\}.$$

Now, let $N \to \infty$ in the way described in (1.4). Then

$$EM_k \to \sum_{r=0}^{k-1} a_{k,r} y^r = \alpha_k(y), \qquad k = 1, 2,... \text{ (see (2.6))}.$$

In order to establish the convergence in probability, it will be shown that $\operatorname{Var} M_k \to 0$, $N \to \infty$,

$$\begin{aligned} n^{2k} \cdot \operatorname{Var} M_k &= \frac{1}{p^2} \sum_{r_1=1}^{p} \cdots\cdots \sum_{r_k=1}^{p} \sum_{r_{k+1}=1}^{p} \cdots\cdots \sum_{r_{2k}=1}^{p} \\ &\quad \cdot \{E(Y_{r_1r_2} \cdot Y_{r_2r_3} \cdot \cdots \cdot Y_{r_kr_1} \\ &\quad \cdot Y_{r_{k+1}r_{k+2}} \cdot Y_{r_{k+2}r_{k+3}} \cdot \cdots \cdot Y_{r_{2k}r_{k+1}}) \\ &\quad - E(Y_{r_1r_2} \cdot Y_{r_2r_3} \cdot \cdots \cdot Y_{r_kr_1}) \\ &\quad \cdot E(Y_{r_{k+1}r_{k+2}} \cdot Y_{r_{k+2}r_{k+3}} \cdot \cdots \cdot Y_{r_{2k}r_{k+1}})\} \\ &= \frac{1}{p^2} \sum_{r_1=1}^{p} \cdots\cdots \sum_{r_{2k}=1}^{p} \sum_{s_1=1}^{n} \cdots\cdots \sum_{s_{2k}=1}^{n} \\ &\quad \cdot \{E(X_{r_1s_1} \cdot X_{r_2s_1} \cdot \cdots \cdot X_{r_{2k}s_{2k}} \cdot X_{r_{k+1}s_{2k}}) \\ &\quad - E(X_{r_1s_1} \cdot X_{r_2s_1} \cdot \cdots \cdot X_{r_ks_k} \cdot X_{r_1s_k}) \\ &\quad \cdot E(X_{r_{k+1}s_{k+1}} \cdot X_{r_{k+2}s_{k+1}} \cdot \cdots \cdot X_{r_{2k}s_{2k}} \cdot X_{r_{k+1}s_{2k}})\}. \end{aligned} \tag{2.14}$$

Let A denote the set of X-variables with index numbers between 1 and k.

Let B denote the set of X's with index numbers between $k+1$ and $2k$. Finally, let C denote the union of A and B. The terms in (2.14) are $\neq 0$ if

(i) the elements in C appear in even powers and A and B have elements in common, and/or

(ii) the elements in A and B appear in even powers and A and B have elements in common.

The total number of indices is $4k$. For every element A and B having in common the number is reduced by 2. Then every time two indices coincide, thus reducing the total number of indices by 1, the number of different elements is reduced by 1, until there are $2k+1$ different indices. At this time there are at least $2k+2$ different elements. We find that the number of different indices is $\leqslant 2k$ if conditions (i) or (ii) are fulfilled. Consequently

$$\operatorname{Var} M_k = \frac{1}{p^2 \cdot n^{2k}} \cdot \{\text{polynomial in } n \text{ and } p \text{ of total degree} \leqslant 2k\} \tag{2.15}$$

which tends to 0, as $N \to \infty$. As a result

$$M_k \xrightarrow{P} \alpha_k(y), \qquad k = 1, 2, \ldots . \tag{2.16}$$

In order to confirm that the moment sequence $\{\alpha_k(y)\}_{k=1}^{\infty}$ uniquely determines limit distribution (2.1), it is sufficient to show that

$$\sum_{k=1}^{\infty} \{\alpha_{2k}(y)\}^{-1/2k} = \infty$$

(Carleman's criterion; see Feller [7b, pp. 227–228]). The $(2k)$th moment has the form $\alpha_{2k}(y) = \sum_{r=0}^{2k-1} a_{2k,r} \cdot y^r$.

In view of (2.7), $y=1$ gives

$$\alpha_{2k}(1) = \sum_{r=0}^{2k-1} a_{2k,r} = \binom{4k}{2k} \frac{1}{2k+1} \leqslant 4^{2k}$$

(this can be shown by induction). It follows that

$$\alpha_{2k}(y) \leqslant \{4(y+1)\}^{2k}, \qquad 0 \leqslant y < \infty, \text{ i.e.,}$$

$$\sum_{k=1}^{\infty} \{\alpha_{2k}(y)\}^{-1/2k} \text{ is divergent.} \tag{2.17}$$

Finally, it is to be shown that

$$\int_{a(y)}^{b(y)} \frac{\sqrt{(x-a(y))(b(y)-x)}}{2\pi y x}\,dx = 1 \qquad \text{for} \quad 0 < y \leqslant 1$$
$$= \frac{1}{y} \qquad \text{for} \quad 1 < y < \infty. \tag{2.18}$$

The transformation $x = -2\sqrt{y}\cos t + y + 1$ gives

$$\int_{a(y)}^{b(y)} \frac{\sqrt{(x-a(y))(b(y)-x)}}{2\pi y x}\,dx$$
$$= \frac{2}{\pi}\int_0^{\pi} \frac{\sin^2 t}{1+y-2\sqrt{y}\cos t}\,dt$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1+y-2\sqrt{y}\cos t}\cdot\left(1-\frac{e^{2it}+e^{-2it}}{2}\right)dt. \tag{2.19}$$

Now, consider the function

$$p_r(t) = \frac{1}{2\pi}\cdot\frac{1-r^2}{1+r^2-2r\cdot\cos t}, \qquad 0 < r < 1.$$

It has the representation

$$p_r(t) = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} r^{|m|}\cdot e^{imt} \qquad \text{(Feller [7b, p. 627])}.$$

(i) The case $0 < y < 1$. With $r = \sqrt{y}$ put into $p_r(t)$ we have

$$\frac{1}{1+y-2\sqrt{y}\cos t} = \frac{1}{1-y}\sum_{m=-\infty}^{\infty} \sqrt{y}^{|m|}\cdot e^{imt}. \tag{2.20}$$

The integral (2.19) takes the form

$$\frac{1}{2\pi(1-y)}\cdot\int_{-\pi}^{\pi}\sum_{m=-\infty}^{\infty}\sqrt{y}^{|m|}\cdot e^{imt}\cdot\left(1-\frac{e^{2it}+e^{-2it}}{2}\right)dt$$
$$= \frac{1}{2\pi(1-y)}\{2\pi-2\pi y\} \qquad \left(\text{since} \int_{-\pi}^{\pi} e^{imt}\,dt = 0 \text{ for } m \neq 0\right)$$
$$= 1.$$

(ii) The case $y = 1$. The limit distribution is a general $\beta(\frac{1}{2}, \frac{3}{2})$-distribution over the interval $(0, 4)$ by Remark 2.1.

(iii) The case $1 < y < \infty$. The integral (2.19) can be written

$$\frac{1}{y} \cdot \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} \frac{1}{1+\dfrac{1}{y}-2\dfrac{1}{\sqrt{y}}\cos t} \cdot \left(1-\frac{e^{2it}+e^{-2it}}{2}\right) dt = \frac{1}{y} \qquad \text{by (i).}$$

In the case $y = 0$ all the moments are $= 1$. The only distribution with this moment sequence is the degenerated δ_1 one.

With (2.16), (2.17) and (2.18) the proof of Theorem 2.1 is now completed (see Feller [7b, p. 269] and Grenander [8, pp. 177–180]).

So far the convergence is stated in probability. A slight modification of conditions (1.4) will also warrant the convergence with probability 1.

THEOREM 2.2. *Suppose $p(1), p(2),\ldots$ is an increasing sequence of positive integers. Then, under conditions* (1.1) *and* (1.4), $F_p^{(n)}(x)$ *converges to* $F_y(x)$ *with probability* 1, *for every* x.

Proof. It is sufficient to show that $M_k \to \alpha_k(y)$ almost surely for every k. By (2.15), $\operatorname{Var} M_k$ is of order $1/p^2$. For every M

$$\sum_{N=1}^{M} \frac{1}{p^2(N)} \leqslant \sum_{p=1}^{M} \frac{1}{p^2}, \qquad \text{i.e.,} \qquad \lim_{M\to\infty} \sum_{N=1}^{M} \frac{1}{p^2(N)} < \infty.$$

From this fact and from Borel–Cantelli's lemma the convergence follows.

The results of Theorem 2.2 can be strengthened further.

THEOREM 2.3. *Under the conditions of Theorem* 2.2, $F_p^{(n)}(x)$ *converges uniformly in* x *to* $F_y(x)$ *with probability* 1.

The proof essentially follows the proof of Glivenko–Cantelli's theorem (see, for instance, Chung [5, pp. 124–125]) and is not given in detail. The fact that the limit distribution is bounded and is without jumps makes the proof simpler.

For the next theorem some concepts are needed. The ρth fractile of the distribution function $F_y(x)$ is defined as the value ξ_ρ, such that $F_y(\xi_\rho) = \rho$, $0 < \rho < 1$. The ρth fractile, $\hat{\xi}_\rho$, of $F_p^{(n)}(x)$ ($F_p^{(n)}(x)$ is based on the eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_p$) is defined as

$$\begin{aligned} \hat{\xi}_\rho &= \lambda_{[\rho p]+1} \qquad \text{if } \rho \cdot p \neq \text{integer} \\ &= \text{any value in } [\lambda_{\rho p}, \lambda_{\rho p+1}] \qquad \text{if } \rho \cdot p = \text{integer}, \ 0 < \rho < 1. \end{aligned}$$

THEOREM 2.4. *Under the conditions of Theorem* 2.2, $\hat{\xi}_\rho \to \xi_\rho$ *with probability* 1, *as* $N \to \infty$.

For the proof, see Rao ([14, (i), p. 423]).

We terminate this chapter with some notes.

Reasons for Restriction to the Case $y \leqslant 1$

According to (1.4), $p/n \to y$ as $N \to \infty$. In principle it is sufficient to assume that $y \leqslant 1$. To see this, let $\mathbf{X}$ denote the $p \times n$ matrix with elements X_{ij} defined in (1.1). Then $\mathbf{S}_p^{(n)} = \mathbf{X} \cdot \mathbf{X}^T$ is a $p \times p$ Wishart matrix with n d.f. Also $\mathbf{X}^T \cdot \mathbf{X}$ is of Wishart type, with dimension $n \times n$ and with p d.f. Suppose that $p \leqslant n$. Then, if $\mu_p^{(n)}$ and $\mu_n^{(p)}$ are stochastic roots of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively, the conditional distribution of $\mu_n^{(p)}$ given that $\mu_n^{(p)} > 0$ is the same as the distribution of $\mu_p^{(n)}$. Therefore, it might be natural to use a symmetric standardization in p and n, for example, define $\lambda_p^{(n)}$ as $\mu_p^{(n)}/\sqrt{np}$ instead of $\mu_p^{(n)}/n$. The density of the limit distribution for $0 < y \leqslant 1$ will then be

$$\begin{aligned} f_y'(x) &= \frac{\sqrt{(x-a'(y))(b'(y)-x)}}{2\pi\sqrt{y}\,x} \\ &\qquad \text{for } a'(y) < x < b'(y);\ a'(y) = \frac{1+y}{\sqrt{y}} - 2,\ b'(y) = \frac{1+y}{\sqrt{y}} + 2 \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

Note that for every fixed $y \leqslant 1$ the density over $(a'(y), b'(y))$ is proportional to the density for $1/y$ over the same interval.

A Note on the Special Case $y = 1$

Suppose conditions (1.1) and (1.4) are fulfilled with $y = 1$. If X' has a distribution with density (1.3), then $(2X')^2$ has the density $f_1(x)$ of (2.2). This relationship with the semi-circle distribution can also be seen in the following way. Suppose for the sake of simplicity that $p = n$ for every N and that $X_{ij} = X_{ji}$ for all i and j. The latter condition does not affect the asymptotic values of the moments as the number of sequences of type (2.9), expressed in X-variables, containing both of X_{ij} and X_{ji}, is negligible. Set

$$\mathbf{U}_n = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{pmatrix}.$$

If ρ_n is a stochastic root of $(1/\sqrt{n})\mathbf{U}_n$, then ρ_n^2 is a stochastic root of $(1/n)\mathbf{U}_n^2$, which is equivalent with $(1/n)\mathbf{S}_n^{(n)}$. According to Wigner [16a]

$$\lim_{n\to\infty} E(\rho_n^2)^k = \lim_{n\to\infty} E(\rho_n)^{2k}$$

$$\left(\text{note that } \frac{\rho_n}{2} \text{ gives the semi-circle distribution over } [-1, 1]\right)$$

$$= \binom{2k}{k}\frac{1}{k+1}$$

$$= \alpha_k(1) \qquad \text{by (2.7)}$$

$$= \lim_{n\to\infty} E(\lambda_n^{(n)})^k,$$

i.e., ρ_n^2 and $\lambda_n^{(n)}$ has the same limit distribution. (This is Corollary 1 by Arnold [3b].)

The Limit Distribution when the Elements of **S** *Are Independent*

Under conditions (1.2) the elements of **S** are not independent. It might be of interest to see how the dependence between the Y's affects the limit distribution. Suppose that the structure of the elements is preserved, but that they are independent on and above the main diagonal. Formally, denote the matrix by $\mathbf{S}^*$ and let

$$Y_{rs}^* = \sum_{j=1}^{n} X_{rj}^{(s)} \cdot X_{sj}^{(r)} \tag{2.21}$$

be its elements, where $X_{rj}^{(s)}$ for $1 \leqslant r, s \leqslant p$, $1 \leqslant j \leqslant n$ are independent and $\in N(0; 1)$. Let $F_p^{(n)*}(x)$ be the c.d.f. of the eigenvalues of $(1/n)\mathbf{S}^*$. An analogue to Theorem 2.1 is

THEOREM 2.5. *Under conditions* (2.21) *and* (1.4), $F_p^{(n)*}(x)$ *converges in probability to a distribution function* $F_y^*(x)$ *for every* x. *This is, for* $0 < y < \infty$, *a continuous distribution with density*

$$f_y^*(x) = \frac{\sqrt{(x-a^*(y))(b^*(y)-x)}}{2\pi y}$$

$$\text{for } a^*(y) < x < b^*(y);\ a^*(y) = 1 - 2\sqrt{y},\ b^*(y) = 1 + 2\sqrt{y}$$

$$= 0 \qquad \text{otherwise.} \tag{2.22}$$

Proof. The limit distribution is a general $\beta(\frac{3}{2}, \frac{3}{2})$-distribution over $(a^*(y), b^*(y))$. Let $X \in \beta(\frac{3}{2}, \frac{3}{2})$ (over $(0, 1)$). Then $X' = 2(X - \frac{1}{2})$ has a Wigner distribution with density (1.3) and $Y' = 2\sqrt{y}\,X' + 1 = 4\sqrt{y}(X - \frac{1}{2}) + 1$ has density (2.22). The kth moment of Y' is

$$\alpha_k^*(y) = EY'^k = \sum_{r=0}^{[k/2]} \frac{1}{r+1}\binom{k}{r}\binom{k-r}{r} y^r, \qquad k = 1, 2, \ldots \text{ (cf. (2.6)).}$$

Let M_k^* be the kth moment in regard of $F_p^{(n)*}(x)$. Here we shall only show that

$$EM_k^* \to \alpha_k^*(y), \qquad N \to \infty. \tag{2.23}$$

The proof is analogous to the proof of Theorem 2.1. However, only a few of the terms displayed in Table I are relevant for this case, viz. the terms which, besides the diagonal elements, only contain second powers of the off diagonal elements. The latter ones are forming chains of the same type as those occurring in Wigner's proof of the Semi-Circle Law [16a]. An analogue to Table I is Table III.

The number of ways in which $2r$ diagonal elements can be placed out among $k-2r$ off diagonal elements is $\binom{k}{2r}$. Thus, by (15c) of [16a] (Wigner) the coefficient of $n^{k-r} \cdot p^r$ is

$$\frac{(2r)!}{r!\,(r+1)!} \cdot \binom{k}{2r} = \frac{1}{r+1}\binom{k}{r}\binom{k-r}{r}.$$

Now, if $N \to \infty$ in the way described in (1.4), (2.23) follows. The rest of the proof is carried through in line with the proof of Theorem 2.1 with necessary modifications.

TABLE III

Number of diagonal elements	Type of index pairs (every chain consists of k pairs)	"Leading term"[a]
k	$(1, 1)(1, 1) \cdots (1, 1)$	n^k
$k-2$	[b] $\cdots (1, r) \cdots (r, 1) \cdots$	$n^{k-1} \cdot p$
$k-4$	[b] $\cdots (1, r) \cdots (r, 1) \cdots (1, s) \cdots (s, 1)$ or $\cdots (1, r) \cdots (r, s) \cdots (s, r) \cdots (r, 1)$	$n^{k-2} \cdot p^2$
Etc.		

[a] Note that $EY_{rs}'^{2k} \sim n^k$ if $r \neq s$; $EY_{rs}'^{2k-1} = 0$ if $r \neq s$; $EY_{rr}'^{k} \sim n^k$.
[b] The diagonal pairs are omitted.

3. Limit Theorems for the C.D.F. under Weaker Conditions

In Section 2 the X_{ij}'s were assumed to be normally distributed. Now, suppose that X_{ij}, $i = 1, 2,..., p$, $j = 1, 2,..., n$, are i.i.d. random variables with

$$EX_{ij} = 0, \qquad \operatorname{Var} X_{ij} = 1, \qquad EX_{ij}^4 < \infty. \tag{3.1}$$

THEOREM 3.1. *Under conditions* (3.1) *and* (1.4) *the statements of Theorem* 2.1 *are still true, provided* $y > 0$.

Proof. A truncation method used in the proof of Arnold's theorem [3a] is applied with slight modifications. X_{ij} corresponds to a_{ij} in Arnold's paper. The following lemmas will be needed.

LEMMA 3.1. *Set* $M = M(N) = \max(p, n)$, *where* p *and* n *satisfy* (1.4), $y > 0$. *Then for every distribution function* $H(x)$

$$\int |x|^k \, dH(x) < \infty \Rightarrow \lim_{N\to\infty} p^{(r-k-s)/2} \cdot n^{-s/2} \cdot \int_{|x|<M^{1/2}} x^r \, dH(x) = 0$$
$$\text{for} \quad s = 1, 2,..., r-k-1$$

for every r *and* k (*integers*) *with* $r \geqslant k + 1 \geqslant 2$.

Proof. By use of

$$\int_{|x|<M^{1/2}} x^r \, dH(x) = \int_{|x|<M^{1/4}} x^r \, dH(x) + \int_{M^{1/4}\leqslant |x|<M^{1/2}} x^r \, dH(x)$$

we find that

$$\begin{aligned}
&\left| \int_{|x|<M^{1/2}} x^r \, dH(x) \right| \\
&\quad \leqslant \int_{|x|<M^{1/4}} |x|^{r-k-s} \cdot |x|^s \cdot |x|^k \, dH(x) \\
&\quad\quad + \int_{M^{1/4}\leqslant |x|<M^{1/2}} |x|^{r-k-s} \cdot |x|^s \cdot |x|^k \, dH(x) \\
&\quad \leqslant M^{(r-k-s)/4} \cdot M^{s/4} \cdot \int_{|x|<M^{1/4}} |x|^k \, dH(x) \\
&\quad\quad + M^{(r-k-s)/2} \cdot M^{s/2} \cdot \int_{M^{1/4}\leqslant |x|<M^{1/2}} |x|^k \, dH(x),
\end{aligned}$$

i.e.,

$$\begin{aligned}
&\left| p^{-(r-k-s)/2} \cdot n^{-s/2} \cdot \int_{|x|<M^{1/2}} x^r \, dH(x) \right| \\
&\quad \leqslant \left(\frac{M}{p^2}\right)^{(r-k-s)/4} \cdot \left(\frac{M}{n^2}\right)^{s/4} \cdot \int_{|x|<M^{1/4}} |x|^k \, dH(x) \\
&\quad + \left(\frac{M}{p}\right)^{(r-k-s)/2} \cdot \left(\frac{M}{n}\right)^{s/2} \cdot \int_{M^{1/4} \leqslant |x| < M^{1/2}} |x|^k \, dH(x). \qquad (3.2)
\end{aligned}$$

Since $M/p = \max(1, n/p)$, $M/n = \max(p/n, 1)$ and $\lim_{N\to\infty}(p/n) = y$, $y > 0$, it follows that $\overline{\lim}_{N\to\infty}(M/p) \leqslant 1 + 1/y < \infty$ and $\overline{\lim}_{N\to\infty}(M/n) \leqslant 1 + y < \infty$. Thus

$$\frac{M}{p^2} \to 0, \qquad \frac{M}{n^2} \to 0, \qquad N \to \infty.$$

We further note that

$$\int_{|x|<M^{1/4}} |x|^k \, dH(x) \leqslant \int |x|^k \, dH(x) < \infty$$

and

$$\int_{M^{1/4} \leqslant |x| < M^{1/2}} |x|^k \, dH(x) \leqslant \int_{M^{1/4} \leqslant |x|} |x|^k \, dH(x) \to 0,$$

$$N \to \infty, \text{ since } M^{1/4} \to \infty.$$

Consequently, the right member of $(3.2) \to 0$ as $N \to \infty$ and the lemma is proved.

LEMMA 3.2. *Suppose that the conditions of Lemma* 3.1 *are true. Then*

$$\left.\begin{aligned} &\int |x|^k \, dH(x) < \infty \\ &\int x \, dH(x) = 0 \end{aligned}\right\} \Rightarrow \lim_{N\to\infty} p^{(k-1-s)/2} \cdot n^{s/2} \cdot \int_{|x|<M^{1/2}} x \, dH(x) = 0,$$

$$s = 0, 1, \ldots, k-1; \; k = 1, 2, \ldots .$$

Proof. As $\int x \, dH(x) = 0$ it follows that

$$\begin{aligned}
&\left| p^{(k-1-s)/2} \cdot n^{s/2} \cdot \int_{|x|<M^{1/2}} x \, dH(x) \right| \\
&\quad = \left| p^{(k-1-s)/2} \cdot n^{s/2} \cdot \int_{|x|\geqslant M^{1/2}} x \, dH(x) \right|
\end{aligned}$$

$$\leqslant p^{(k-1-s)/2} \cdot n^{s/2} \cdot \int_{|x| \geqslant M^{1/2}} |x|^k \cdot \frac{1}{|x|^s} \cdot \frac{1}{|x|^{k-1-s}} \, dH(x)$$

$$\leqslant \left(\frac{p}{M}\right)^{(k-1-s)/2} \cdot \left(\frac{n}{M}\right)^{s/2} \cdot \int_{|x| \geqslant M^{1/2}} |x|^k \, dH(x). \tag{3.3}$$

But p/M and n/M are both $\leqslant 1$ for every N. Thus, the last member of (3.3) tends to 0, since $\int_{|x| \geqslant M^{1/2}} |x|^k \, dH(x) \to 0$.

Especially, if $H(x)$ is the distribution function of X_{ij}, introduced in (3.1), then Lemmas 3.1 and 3.2 with $k=2$ give

$$\text{(i)} \quad \left.\begin{array}{l} \int x^2 \, dH(x) < \infty \\ \\ \int x \, dH(x) = 0 \end{array}\right\} \Rightarrow \begin{array}{l} \lim\limits_{N\to\infty} n^{1/2} \cdot \int_{|x| < M^{1/2}} x \, dH(x) = 0 \\ \text{and} \quad \lim\limits_{N\to\infty} p^{1/2} \cdot \int_{|x| < M^{1/2}} x \, dH(x) = 0. \end{array}$$

$$\text{(ii)} \quad \int x^2 \, dH(x) < \infty \Rightarrow \lim_{N\to\infty} p^{-(r-s-2)/2} \cdot n^{-s/2} \cdot \int_{|x| < M^{1/2}} x^r \, dH(x) = 0,$$

$r = 3, 4, \ldots; s = 1, 2, \ldots, r-3$.

Furthermore, it is true that

$$\text{(iii)} \quad \int x^2 \, dH(x) = 1 \Rightarrow \lim_{N\to\infty} \int_{|x| < M^{1/2}} x^2 \, dH(x) = 1.$$

Statements (i), (ii) and (iii) are applied in the same way as (5) of Arnold's paper [3a], with X_{ij} instead of a_{ij}.

Here we consider

$$pn^k EM_k = E \sum_{i_1=1}^{p} \cdots \sum_{i_k=1}^{p} \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} X_{i_1j_1} \cdot X_{i_2j_1} \cdot X_{i_2j_2} \cdot X_{i_3j_2} \cdots X_{i_{k-1}j_{k-1}} \cdot X_{i_kj_{k-1}} \cdot X_{i_kj_k} \cdot X_{i_1j_k}.$$

Set $r_1 = i_1$, $r_2 = j_1$, $r_3 = i_2$, $r_4 = j_2, \ldots, r_{2k-1} = i_k$, $r_{2k} = j_k$, which gives

$$pn^k EM_k = E \sum_{r_1=1}^{p} \sum_{r_2=1}^{n} \sum_{r_3=1}^{p} \sum_{r_4=1}^{n} \cdots \sum_{r_{2k-1}=1}^{p} \sum_{r_{2k}=1}^{n} X_{r_1r_2} \cdot X_{r_3r_2} \cdots X_{r_{2k-1}r_{2k}} \cdot X_{r_1r_{2k}}.$$

The change of indices has given a closed index chain $r_1, r_2, \ldots, r_{2k}, r_1$. (Every second index pair is reversed, but it is easily seen that this fact does not lead to any difficulties.) It is sufficient to study terms where the X_{ij}'s appear quadratically, i.e., there are k different X_{ij}'s in every term. Set

$$\begin{aligned} X'_{ij} &= X_{ij} \qquad \text{if } |X_{ij}| < M^{1/2} \\ &= 0 \qquad \text{if } |X_{ij}| \geqslant M^{1/2}. \end{aligned} \tag{3.4}$$

Let

$$M'_k = \frac{1}{pn^k} \sum_{r_1=1}^{p} \sum_{r_2=1}^{n} \cdots\cdots \sum_{r_{2k}=1}^{n} X'_{r_1r_2} \cdot X'_{r_3r_2} \cdot \cdots \cdot X'_{r_1r_{2k}}.$$

It is seen that

$$\lim_{N\to\infty} EM'_k = \lim_{N\to\infty} EM_k = \alpha_k(y), \qquad k = 1, 2, \ldots.$$

Note that M_k in Section 2 was expressed in terms of Y_{rs}'s instead of X_{ij}'s just in order to simplify the assorting of the different index combinations.

Analogous modifications of (c), (d) and (e) of [3a] complete the proof of Theorem 3.1. For example, p_n of (e) here takes the form

$$\begin{aligned} p_n &= \sum_{i=1}^{p} \sum_{j=1}^{n} P(X_{ij} \neq X'_{ij}) \\ &= \sum_{i=1}^{p} \sum_{j=1}^{n} P(|X_{ij}| \geqslant M^{1/2}) \\ &\leqslant M^2\{1 - H(M^{1/2} - 0) + H(-M^{1/2})\} \\ &\leqslant \int_{|x| \geqslant M^{1/2}} x^4 dH(x) \to 0 \qquad \text{if } \int x^4\, dH(x) < \infty. \end{aligned} \tag{3.5}$$

Note that the existence of the 4th moment is necessary here. It is not requested in the former part of the proof (only the 2nd moment).

The correspondence to Theorem 2.2 is

THEOREM 3.2. *With condition* (3.1) *instead of* (1.1) *and with the additional condition* $EX_{ij}^6 < \infty$, *the statements of Theorem* 2.2 *are still true, provided* $y > 0$.

The details of the proof are omitted.

Remark. The results of this section are still true if the elements $\sum_{j=1}^{n} X_{rj} \cdot X_{sj}$ of $\mathbf{S}_p^{(n)}$ are replaced with $\sum_{j=1}^{n}(X_{rj} - \bar{X}_r)(X_{sj} - \bar{X}_s)$, where $\bar{X}_r = (1/n) \sum_{j=1}^{n} X_{rj}$.

4. Limit Theorems for Sums of Eigenvalues

Let $A_k = \operatorname{trace} \mathbf{S}^k = n^k \sum_{i=1}^p \lambda_i^k = pn^k M_k$, $k = 1, 2, \ldots$ (see (2.8)). Arharov [2] (Theorem 2) has set up a central limit theorem for $A_1, A_2, \ldots, A_m$. The theorem will be restated here together with a detailed proof. Arharov's version is corrected on several points.

THEOREM 4.1. *Set*

$$\eta_k = \frac{1}{(n+p)^k}(A_k - EA_k), \qquad k = 1, 2, \ldots, m.$$

Then, under conditions (1.1) *and* (1.4), *the stochastic vector* $(\eta_1, \eta_2, \ldots, \eta_m)$ *converges in distribution to a certain m-variate normal distribution*; $m = 1, 2, \ldots$.

Proof. Here we shall use a multidimensional analogue of the moment method, exploited in the proof of Theorem 2.1. Consider the mixed moments

$$B_{s_1, s_2, \ldots, s_m} = E(A_1 - EA_1)^{s_1}(A_2 - EA_2)^{s_2} \cdots (A_m - EA_m)^{s_m}; s_1, s_2, \ldots, s_m$$

are non-negative integers. Set

$$s = s_1 + s_2 + \cdots + s_m, \qquad \alpha = s_1 + 2s_2 + \cdots + ms_m$$

and

$$G\begin{pmatrix} i_1 i_2 \cdots i_k \\ j_1 j_2 \cdots j_k \end{pmatrix} = X_{i_1 j_1} \cdot X_{i_2 j_1} \cdot X_{i_2 j_2} \cdot X_{i_3 j_2} \cdot \cdots \cdot X_{i_k j_k} \cdot X_{i_1 j_k}, \qquad k = 1, 2, \ldots.$$

Then

$$\begin{aligned} B_{s_1, s_2, \ldots, s_m} = E \sum_{i_1, i_2, \ldots, i_\alpha} \sum_{j_1, j_2, \ldots, j_\alpha} & \left[\left\{ G\begin{pmatrix} i_1 \\ j_1 \end{pmatrix} - EG\begin{pmatrix} i_1 \\ j_1 \end{pmatrix} \right\} \right. \\ & \left. \cdot \cdots \cdot \left\{ G\begin{pmatrix} i_{s_1} \\ j_{s_1} \end{pmatrix} - EG\begin{pmatrix} i_{s_1} \\ j_{s_1} \end{pmatrix} \right\} \right] \\ & \times \left[\left\{ G\begin{pmatrix} i_{s_1+1} & i_{s_1+2} \\ j_{s_1+1} & j_{s_1+2} \end{pmatrix} - EG\begin{pmatrix} i_{s_1+1} & i_{s_1+2} \\ j_{s_1+1} & j_{s_1+2} \end{pmatrix} \right\} \right. \\ & \left. \cdot \cdots \cdot \left\{ G\begin{pmatrix} i_{s_1+2s_2-1} & i_{s_1+2s_2} \\ j_{s_1+2s_2-1} & j_{s_1+2s_2} \end{pmatrix} - EG\begin{pmatrix} i_{s_1+2s_2-1} & i_{s_1+2s_2} \\ j_{s_1+2s_2-1} & j_{s_1+2s_2} \end{pmatrix} \right\} \right] \\ & \cdot \cdots \cdot \left[\left\{ G\begin{pmatrix} i_{\alpha-ms_m+1} & i_{\alpha-ms_m+2} \cdot \cdots \cdot i_{\alpha-m(s_m-1)} \\ j_{\alpha-ms_m+1} & j_{\alpha-ms_m+2} \cdot \cdots \cdot j_{\alpha-m(s_m-1)} \end{pmatrix} \right. \right. \end{aligned}$$

$$-EG\begin{pmatrix} i_{\alpha-ms_m+1} & i_{\alpha-ms_m+2} & \cdots & i_{\alpha-m(s_m-1)} \\ j_{\alpha-ms_m+1} & j_{\alpha-ms_m+2} & \cdots & j_{\alpha-m(s_m-1)} \end{pmatrix}\Bigg\}$$

$$\cdots\cdots\Bigg\{G\begin{pmatrix} i_{\alpha-m+1} & i_{\alpha-m+2} & \cdots & i_{\alpha} \\ j_{\alpha-m+1} & j_{\alpha-m+2} & \cdots & j_{\alpha} \end{pmatrix}$$

$$-EG\begin{pmatrix} i_{\alpha-m+1} & i_{\alpha-m+2} & \cdots & i_{\alpha} \\ j_{\alpha-m+1} & j_{\alpha-m+2} & \cdots & j_{\alpha} \end{pmatrix}\Bigg\}\Bigg]. \tag{4.1}$$

It is to be shown that these mixed moments divided by $(n+p)^\alpha$ converge to the corresponding moments in a certain m-variate normal distribution. The following lemmas will then be needed.

LEMMA 4.1. *$B_{s_1,s_2,\ldots,s_m}$ is a polynomial in n and p of total degree $= \alpha$ if s is even, and $< \alpha$ if s is odd. If, moreover, s is even, then*

$$B_{s_1,s_2,\ldots,s_m} = \sum B_{m_1m_2}(n,p)\cdot B_{m_3m_4}(n,p)\cdot\cdots\cdot B_{m_{s-1}m_s}(n,p) + \Delta(n,p),$$

where the summation is taken over all possible partitions of s_1 1's, s_2 2's,..., s_m m's into pairs (m_1, m_2), (m_3, m_4),..., (m_{s-1}, m_s) and where $\Delta(n,p)$ is a polynomial of degree $<\alpha$. $B_{m_{2i-1}m_{2i}}(n,p)$ is a polynomial of degree $m_{2i-1}+m_{2i}$, $i = 1, 2,\ldots, s/2$, and is uniquely determined by its indices.

Lemma 4.1 is the same as Lemma 2 by Arharov [2] apart from some details.

Remark 4.1. In forming the partitions, we disregard the order between pairs and between digits inside pairs. All the digits are distinguishable, even digits with the same value. For instance 1, 1, 1, 1 has 3 partitions: (1, 1), (1, 1), which can be formed in 3 ways $((1^1, 1^2), (1^3, 1^4); (1^1, 1^3), (1^2, 1^4); (1^1, 1^4), (1^2, 1^3))$. 1, 1, 2, 2 has 3 partitions: (1, 1), (2, 2); (1, 2), (1, 2) (2 ways), 1, 1, 1, 2, 2, 3 has 15 partitions: (1, 1), (1, 2), (2, 3) (6 ways); (1, 1), (2, 2), (1, 3) (3 ways); (1, 2), (1, 2), (1, 3) (6 ways).

Proof of Lemma 4.1. The degree of $B_{s_1,s_2,\ldots,s_m}$ equals the number of different indices among $i_1, i_2,\ldots, i_\alpha, j_1, j_2,\ldots, j_\alpha$. However, all these indices cannot be chosen different, for as soon as a factor in (4.1) only contains indices not appearing in any other factor, then $B_{s_1,s_2,\ldots,s_m} = 0$. We find that the maximum degree for even s's is obtained if the factors are divided into pairs, whose index sets are unique. Inside every pair, the maximum degree is obtained by setting certain indices equal. If more than two factors contain the same index or if two or more pairs have an index in common, then the degree must be $<\alpha$.

Every pair has an expectation of the form const. $\cdot$ $E(A_k - EA_k)^2$ or const. $\cdot$ $E(A_k - EA_k)(A_l - EA_l)$. The first one is a polynomial of degree $2k$, the second one a polynomial of degree $k + l$.

This is true because the X_{ij}'s must occur quadratically together with the fact that the factors inside a pair must have at least one X-variable in common. As a consequence $B_{s_1, s_2, \ldots, s_m}$ is a polynomial of degree

$$\begin{aligned}(m_1 + m_2) + (m_3 + m_4) + \cdots + (m_{s-1} + m_s)\\ = (1 + 1 + \cdots + 1) + (2 + 2 + \cdots + 2) + \cdots + (m + m + \cdots + m)\\ (s_1 \text{ 1's}, s_2 \text{ 2's}, \ldots, s_m\ m\text{'s}) = s_1 + 2s_2 + \cdots + ms_m = \alpha.\end{aligned}$$

The coefficients are then obtained by summing over the possible partitions into $s/2$ pairs.

If s is odd, it is not possible to form even pairs. More than two factors will then have indices in common or some pair of factors will have indices in common with other factors. This means that the degree of $B_{s_1, s_2, \ldots, s_m}$ must be $< \alpha$.

The next step is to determine the mixed moments of a multivariate normal distribution. This can be done by use of characteristic functions. First some preliminaries.

Let $g(t) = \exp(y(t))$, where $y(t)$ is a quadratic form in $t = (t_1, t_2, \ldots, t_m)$, i.e., $y(t) = \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} t_i t_j$; the c_{ij}'s are real constants. Suppose that $c_{ij} = c_{ji}$ for all i and j. Then

$$y'_k = \frac{\partial y(t)}{\partial t_k} = \frac{\partial}{\partial t_k} \left\{ 2 \sum_{j \neq k} c_{kj} t_k t_j + c_{kk} \cdot t_k^2 \right\} = 2 \sum_{j=1}^{m} c_{kj} t_j,$$

$$y''_{kk} = \frac{\partial^2 y(t)}{\partial t_k^2} = 2c_{kk},$$

$$y''_{kl} = \frac{\partial^2 y(t)}{\partial t_l \partial t_k} = 2c_{kl}.$$

All derivatives of higher order vanish.

Now, consider derivatives of all orders of $g(t)$:

$$g_{r_1, r_2, \ldots, r_s} = \frac{\partial^s g(t)}{\partial t_{r_1} \cdot \partial t_{r_2} \cdot \cdots \cdot \partial t_{r_s}}, \qquad s = 1, 2, \ldots;$$

the indices $r_1, r_2, \ldots, r_s$ need not all be different.

LEMMA 4.2. *If s is even, then*

$$g_{r_1,r_2,\ldots,r_s} = \exp(y(t)) \cdot \Big\{ y'_{r_1} \cdot y'_{r_2} \cdot \cdots \cdot y'_{r_s}$$

$$+ \sum y''_{m_1m_2} \cdot y'_{m_3} \cdot \cdots \cdot y'_{m_s}$$

$$+ \sum y''_{m_1m_2} \cdot y''_{m_3m_4} \cdot y'_{m_5} \cdot \cdots \cdot y'_{m_s}$$

$$+ \cdots + \sum y''_{m_1m_2} \cdot y''_{m_3m_4} \cdot \cdots \cdot y''_{m_{s-1}m_s} \Big\} .$$

If s is odd, the last summation is replaced by

$$\sum y''_{m_1m_2} \cdot y''_{m_3m_4} \cdot \cdots \cdot y''_{m_{s-2}m_{s-1}} \cdot y'_{m_s}.$$

Here $(m_1, m_2, \ldots, m_s)$ *is an arrangement of* $(r_1, r_2, \ldots, r_s)$. *In every sum,* $(m_1, m_2, \ldots, m_s)$ *varies so that all possible index combinations occur.*

Proof. The statement is true for $s = 1$, since $g_{r_1} = \exp(y(t)) \cdot y'_{r_1}$. If the statement is true for s, then it is true for $s + 1$, as

$$g_{r_1,r_2,\ldots,r_{s+1}} = \frac{\partial}{\partial t_{r_{s+1}}} g_{r_1,r_2,\ldots,r_s}$$

$$= g_{r_1,r_2,\ldots,r_s} \cdot y'_{r_{s+1}} + \exp(y(t))$$

$$\times \Big\{ \sum_{i=1}^{s} y''_{r_{s+1}i} \cdot y'_1 \cdot \cdots \cdot y'_{i-1} \cdot y'_{i+1} \cdot \cdots \cdot y'_{r_s}$$

$$+ \Big(\sum y''_{m_1m_2} \cdot y''_{m_3m_{s+1}} \cdot y'_{m_4} \cdot \cdots \cdot y'_{m_s}$$

$$+ \cdots + \sum y''_{m_1m_2} \cdot y'_{m_3} \cdot \cdots \cdot y'_{m_{s-1}} \cdot y''_{m_sm_{s+1}} \Big)$$

$$+ \Big(\sum y''_{m_1m_2} \cdot y''_{m_3m_4} \cdot y''_{m_5m_{s+1}} \cdot y'_{m_6} \cdot \cdots \cdot y'_{m_s}$$

$$+ \cdots + \sum y''_{m_1m_2} \cdot y''_{m_3m_4} \cdot y'_{m_5} \cdot \cdots \cdot y'_{m_{s-1}} \cdot y''_{m_sm_{s+1}} \Big) \Big\}$$

$$+ \cdots ,$$

i.e., we get all possible combinations of first and second derivatives of $y(t)$ in respect to $t_{r_1}, t_{r_2}, \ldots, t_{r_s}, t_{r_{s+1}}$.

COROLLARY 4.1. *Let* $X_1, X_2, \ldots, X_m$ *be jointly normally distributed* *random variables, all with expectation zero, and* $\operatorname{Cov}(X_i, X_j)$ *denoted by* σ_{ij}. *Then*

$$
\begin{aligned}
&EX_{r_1} \cdot X_{r_2} \cdot \cdots \cdot X_{r_s} \\
&\quad = \sum \sigma_{m_1 m_2} \cdot \sigma_{m_3 m_4} \cdot \cdots \cdot \sigma_{m_{s-1} m_s} \qquad \text{if } s \text{ is even} \\
&\quad = 0 \qquad \text{if } s \text{ is odd};
\end{aligned}
$$

$r_1, r_2, \ldots, r_s$ *are selected among the first m positive integers* (they need not all be different). *The summation is taken over all possible partitions of* $r_1, r_2, \ldots, r_s$ *into pairs* (m_1, m_2), $(m_3, m_4), \ldots, (m_{s-1}, m_s)$ (see Remark 4.1).

Proof. The characteristic function of $(X_1, X_2, \ldots, X_m)$ is of the type $g(t)$ with $c_{ij} = -\frac{1}{2}\sigma_{ij}$, i.e.,

$$
g(t) = \exp\left(-\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_{ij} \cdot t_i \cdot t_j\right),
$$

$y_k' = 0$ for $t = 0$ and $y_{kl}'' = -\sigma_{kl}$; $k, l = 1, 2, \ldots, m$.

We have the following relationship between the moments and the characteristic function:

$$
\begin{aligned}
&EX_{r_1} \cdot X_{r_2} \cdot \cdots \cdot X_{r_s} \\
&\quad = (i)^s \cdot \left.\frac{\partial g(t)}{\partial t_{r_1} \cdot \partial t_{r_2} \cdot \cdots \cdot \partial t_{r_s}}\right|_{t=0} \\
&\quad = \sum y_{m_1 m_2}'' \cdot y_{m_3 m_4}'' \cdot \cdots \cdot y_{m_{s-1} m_s}'' \cdot (i)^s \qquad \text{if } s \text{ is even} \\
&\quad = 0 \qquad \text{if } s \text{ is odd} \\
&\quad = \sum \sigma_{m_1 m_2} \cdot \sigma_{m_3 m_4} \cdot \cdots \cdot \sigma_{m_{s-1} m_s} \qquad \text{if } s \text{ is even} \\
&\quad = 0 \qquad \text{if } s \text{ is odd.}
\end{aligned}
$$

The statement of Corollary 4.1 can be written in the following form (this is Lemma 3 by Arharov [2]).

COROLLARY 4.1′. *Under the conditions of Corollary* 4.1

$$
\begin{aligned}
&EX_1^{s_1} \cdot X_2^{s_2} \cdot \cdots \cdot X_m^{s_m} \\
&\quad = \sum \sigma_{m_1 m_2} \cdot \sigma_{m_3 m_4} \cdot \cdots \cdot \sigma_{m_{s-1} m_s} \qquad \text{if } s = s_1 + s_2 + \cdots + s_m \text{ is even} \\
&\quad = 0 \qquad \text{if } s \text{ is odd.}
\end{aligned}
$$

The summation is taken over all partitions of s_1 1's, s_2 2's,..., s_m *m*'s *into pairs* (m_1, m_2), $(m_3, m_4), \ldots, (m_{s-1}, m_s)$ (see Remark 4.1).

A formula for determining the coefficient of the term $\sigma_{m_1m_2} \cdot \sigma_{m_3m_4} \cdot \cdots \cdot \sigma_{m_{s-1}m_s}$ (i.e., the number of partitions which gives this term) is given by the following lemma. It is stated without proof.

LEMMA 4.3. *The coefficient of* $\sigma_{m_1m_2} \cdot \sigma_{m_3m_4} \cdot \cdots \cdot \sigma_{m_{s-1}m_s}$ *in* $EX_1^{s_1} \cdot X_2^{s_2} \cdot \cdots \cdot X_m^{s_m}$, *where* $s = s_1 + s_2 + \cdots + s_m$ *is even, is*

$$\frac{s_1!\, s_2! \cdot \cdots \cdot s_m!}{2^v \cdot f_1! \cdot \cdots \cdot f_u!};$$

v is the number of pairs with equal elements, i.e., with $m_{2i-1} = m_{2i}$, u is the number of different pairs among $(m_1, m_2), (m_3, m_4), \ldots, (m_{s-1}, m_s)$ and $f_1, f_2, \ldots, f_u$ are the frequences for the u pairs. The total sum of coefficients is $1 \cdot 3 \cdot \cdots \cdot (s-1) =$ the coefficient of $\sigma_{ii}^{s/2}$ in EX_i^s.

The last step of the proof of Theorem 4.1 is to establish the convergence of the moments $B_{s_1,s_2,\ldots,s_m}$. By Lemma 4.1

$$\begin{aligned}
&E\eta_1^{s_1} \cdot \eta_2^{s_2} \cdot \cdots \cdot \eta_m^{s_m} \\
&= \frac{B_{s_1,s_2,\ldots,s_m}}{(n+p)^\alpha} \\
&= \sum \frac{B_{m_1m_2}(n,p)}{(n+p)^{m_1+m_2}} \cdot \frac{B_{m_3m_4}(n,p)}{(n+p)^{m_3+m_4}} \cdot \cdots \cdot \frac{B_{m_{s-1}m_s}(n,p)}{(n+p)^{m_{s-1}+m_s}} \\
&\quad + \frac{\Delta(n,p)}{(n+p)^\alpha} \qquad \text{if } s \text{ is even} \\
&= \frac{\text{Polynomial of degree} < \alpha}{(n+p)^\alpha} \qquad \text{if } s \text{ is odd.}
\end{aligned}$$

Every factor

$$\frac{B_{m_{2i-1}m_{2i}}(n,p)}{(n+p)^{m_{2i-1}+m_{2i}}}$$

converges to a constant as $N \to \infty$, since

$$\frac{n}{n+p} \to \frac{1}{1+y} \qquad \text{and} \qquad \frac{p}{n+p} \to \frac{y}{1+y}.$$

Denote the limit by

$$\sigma_{m_{2i-1}m_{2i}} = \lim_{N\to\infty} E\eta_{m_{2i-1}}\eta_{m_{2i}}.$$

We observe that the matrix $\{E\eta_i \eta_j\}$ is a covariance matrix and consequently non-negative definite. Then, the matrix $\{\sigma_{ij}\}$ must have the same property and it is a covariance matrix of a certain normal distribution. By Lemma 4.1 and Corollary 4.1 the mixed moments of $(\eta_1, \eta_2, ..., \eta_m)$ converge to the corresponding moments of this normal distribution. This implies that $(\eta_1, \eta_2, ..., \eta_m)$ converges in distribution to the normal distribution, since the moments determine the distribution uniquely (cf. the proof of Theorem 2.1 together with Feller [7b, p. 529]).

Remark 4.2. If $y = 0$ then $\sigma_{ij} = 0$ for every i and j, i.e., the limit distribution is a degenerated one.

The conditions of Theorem 4.1 can be weakened.

THEOREM 4.2. *Suppose that* X_{ij}, $i = 1, 2, ..., p$, $j = 1, 2, ..., n$ *are i.i.d. random variables with* $EX_{ij} = 0$, $\operatorname{Var} X_{ij} = 1$, $EX_{ij}^4 = \mu_4 < \infty$. *Then the statements of Theorem* 4.1 *are still true, provided* $y > 0$. *However, the asymptotic covariances* σ_{ij} *are not the same as in Theorem* 4.1 *if* $\mu_4 \neq 3$, *since the coefficients of the polynomial* $B_{m_{2i-1}, m_{2i}}$ *depend on* μ_4.

Proof. The truncation method by Arnold [3a] is used in the same way as in the proof of Theorem 3.1. Denote by $(\eta_1', \eta_2', ..., \eta_m')$ the correspondence of $(\eta_1, \eta_2, ..., \eta_m)$ when X_{ij} is replaced with X_{ij}' defined in (3.4). Then the vector $(\eta_1', \eta_2', ..., \eta_m')$ converges in distribution to the actual m-variate normal distribution. It remains to show that $(\eta' - \eta_1, \eta_2' - \eta_2, ..., \eta_m' - \eta_m)$ converges in probability to zero. But

$$P(\max_{1 \leqslant r \leqslant m} |\eta_r' - \eta_r| > \varepsilon) \leqslant \sum_{i=1}^{p} \sum_{j=1}^{n} P(X_{ij} \neq X_{ij}'),$$

which tends to zero, $N \to \infty$, since the 4th moment of X_{ij} is finite (cf. (3.5)).

5. THE GENERALIZED VARIANCE

Let $\mathbf{X}_1$, $\mathbf{X}_2, ..., \mathbf{X}_{n+1}$ be i.i.d. normal vectors with p components, with mean value vector $\boldsymbol{\mu}_p$ and covariance matrix $\boldsymbol{\Sigma}_p$.

Consider the matrix

$$\mathbf{C}_p^{(n+1)} = \frac{1}{n} \sum_{j=1}^{n+1} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^T.$$

Suppose $p \leqslant n$. According to Anderson [1, pp. 170–172] the generalized variance $|\mathbf{C}_p^{(n+1)}|$ under conditions (1.1) has the same distribution as

$$|\mathbf{\Sigma}_p| \cdot \frac{1}{n^p} \cdot |\mathbf{S}_p^{(n)}| = |\mathbf{\Sigma}_p| \cdot \frac{1}{n^p} \cdot \prod_{j=1}^{p} U_j, \tag{5.1}$$

where $U_1, U_2, \ldots, U_p$ are independent χ^2 variables with $n-p+1$, $n-p+2, \ldots, n$ d.f., respectively.

Now, consider the distribution of $|\mathbf{S}_p^{(n)}|$ under the following conditions (conditions (1.4) modified):

Suppose p and $n \to \infty$, $p < n$ for every N and

$$\lim_{N\to\infty} \frac{p}{n} = y, \qquad 0 < y < 1. \tag{5.2}$$

THEOREM 5.1. *Under conditions* (1.1) *and* (5.2), $(1/(n-1)_p)|\mathbf{S}_p^{(n)}|$ *converges in distribution to a lognormal distribution, as* $N \to \infty$; $(n-1)_p = (n-1)(n-2)\cdots(n-p)$. *The parameters are* $\mu = 0$ *and* $\sigma^2 = -2\log(1-y)$, *i.e.*,

$$\frac{1}{\sqrt{-2\log(1-y)}} \log \frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}$$

is asymptotically $\in N(0; 1)$.

Proof. The characteristic function (c.f.) of $\log U_j$ is ([1, p. 171])

$$\begin{aligned} \varphi_j(t) &= E e^{it\log U_j} = E U_j^{it} \\ &= \frac{\Gamma(\frac{1}{2}(n-p+j)+it)}{\Gamma(\frac{1}{2}(n-p+j))} \cdot 2^{it}. \end{aligned} \tag{5.3}$$

Thus, $\log(|\mathbf{S}_p^{(n)}|/(n-1)_p)$ has the c.f.

$$\varphi(t) = \prod_{j=1}^{p} \frac{\Gamma(\frac{1}{2}(n-p+j)+it)}{\Gamma(\frac{1}{2}(n-p+j))} \cdot \frac{2^{it}}{(n-p-1+j)^{it}}.$$

The following expansion formula is valid for $|\arg(z+h)| < \pi$, $|h|$ bounded ([6, p. 47])

$$\begin{aligned} \log\Gamma(z+h) = {} & \left(z+h-\frac{1}{2}\right)\log z - z + \frac{1}{2}\log(2\pi) \\ & + \frac{h^2}{2z} - \frac{h}{2z} + \frac{1}{12z} + O\left(\frac{|h|}{|z|^2}\right). \end{aligned}$$

Set $z = \frac{1}{2}(n-p+j)$ and $h = it$. Then

$$\begin{aligned}\varphi(t) = \exp\Bigg\{&\sum_{j=1}^{p} \log \Gamma\left(\frac{1}{2}(n-p+j)+it\right) \\ &- \sum_{j=1}^{p} \log \Gamma\left(\frac{1}{2}(n-p+j)\right) \\ &- it \sum_{j=1}^{p} \log \frac{n-p+j}{2} - it \log \frac{n-p}{n}\Bigg\} \\ = \exp\Bigg\{&it\left(-\sum_{j=1}^{p} \frac{1}{n-p+j} - \log\left(1-\frac{p}{n}\right)\right) \\ &+ \frac{(it)^2}{2} \cdot \sum_{j=1}^{p} \frac{2}{n-p+j} + \sum_{j=1}^{p} R_{Nj}(t)\Bigg\},\end{aligned}$$

where the remainder term $R_{Nj}(t) = O(|t|/(n-p+j)^2)$, i.e.,

$$\sum_{j=1}^{p} R_{Nj}(t) = O\left(\sum_{j=1}^{p} \frac{|t|}{(n-p+j)^2}\right) = O\left(\frac{|t|}{n-p}\right).$$

Further

$$-\log\left(1-\frac{p}{n+1}\right) < \sum_{j=1}^{p} \frac{1}{n-p+j} < -\log\left(1-\frac{p}{n}\right), \text{ i.e.,} \tag{5.4}$$

$$\sum_{j=1}^{p} \frac{1}{n-p+j} \to -\log(1-y), \text{ as } N \to \infty, \text{ and}$$

$$\varphi(t) \to \exp\left\{\frac{(it)^2}{2}(-2\log(1-y))\right\} = (1-y)^{t^2},$$

which is the c.f. of a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = -2\log(1-y)$.

COROLLARY 5.1. *Under conditions* (1.1) *and* (5.2)

$$\frac{1}{p}\sum_{j=1}^{p} \log \lambda_j = \int_0^\infty \log x \, dF_p^{(n)}(x)$$

converges in probability to

$$\int_0^\infty \log x \, dF_y(x) = \frac{1}{y}\left\{(1-y)\log\frac{1}{1-y} - y\right\}, \qquad \textit{as } N \to \infty.$$

Proof. (i) By Theorem 5.1

$$\frac{1}{p}\log\frac{|\mathbf{S}|}{(n-1)_p}=\frac{1}{p}\log\prod_{j=1}^{p}\lambda_j\cdot\frac{n^p}{(n-1)_p}\xrightarrow{p}0.$$

This means that

$$\begin{aligned}\frac{1}{p}\sum_{j=1}^{p}\log\lambda_j &\xrightarrow{p}\lim_{N\to\infty}\frac{1}{p}\log\frac{(n-1)_p}{n^p}\\
&=\lim_{N\to\infty}\frac{n}{p}\cdot\frac{1}{n}\sum_{j=1}^{p}\log\left(1-\frac{j}{n}\right)\\
&=\frac{1}{y}\int_0^y\log(1-x)\,dx\\
&=\frac{1}{y}\left\{(1-y)\log\frac{1}{1-y}-y\right\}.\end{aligned}$$

(ii) It remains to show that

$$\int_0^\infty\log x\,dF_y(x)=\frac{1}{y}\left\{(1-y)\log\frac{1}{1-y}-y\right\}.$$

The transformation $x=-2\sqrt{y}\cos t+y+1$ gives (cf. (2.18)–(2.20))

$$\begin{aligned}&\int_0^\infty\log x\,dF_y(x)\\
&\quad=\int_{a(y)}^{b(y)}\log x\frac{\sqrt{(x-a(y))(b(y)-x)}}{2\pi yx}dx\\
&\quad=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\log(1+y-2\sqrt{y}\cos t)}{1+y-2\sqrt{y}\cos t}\cdot\left(1-\frac{e^{2it}+e^{-2it}}{2}\right)dt\\
&\quad=\frac{1}{2\pi(1-y)}\int_{-\pi}^{\pi}\Bigg[2\log(1-y)+\left\{\left(y+\frac{1}{y}\right)\log(1-y)\right.\\
&\qquad\left.+(1-y)\right\}(e^{-2it}+e^{2it})+\text{negligible terms}\Bigg]\\
&\qquad\cdot\left(1-\frac{e^{2it}+e^{-2it}}{2}\right)dt\\
&\quad=\frac{1}{y}\left\{(1-y)\log\frac{1}{1-y}-y\right\}.\end{aligned}$$

The special case $y=1$ is considered in the next theorem.

THEOREM 5.1a. (i) *Under conditions* (1.1) *and* (5.2), *but with* $y=1$,

$$\frac{1}{-2\log(1-p/n)}\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}$$

is asymptotically $\in N(0;1)$, $N\to\infty$.

(ii) *Suppose* $p=n$ *for every* N. *Then, under conditions* (1.1)

$$\frac{1}{\sqrt{2\log n}}\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)!}$$

is asymptotically $\in N(0;1)$, $N\to\infty$.

Proof. (i) The c.f. of

$$\frac{1}{\sqrt{-2\log(1-p/n)}}\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}$$

is

$$\begin{aligned}\varphi&\left(\frac{t}{\sqrt{-2\log(1-p/n)}}\right)\\ &=\exp\left\{\frac{it}{\sqrt{-2\log(1-p/n)}}\left(-\sum_{j=1}^{p}\frac{1}{n-p+j}-\log\left(1-\frac{p}{n}\right)\right)\right.\\ &\quad+\frac{(it)^2}{-4\log(1-p/n)}\sum_{j=1}^{p}\frac{2}{n-p+j}\\ &\quad\left.+\sum_{j=1}^{p}R_{Nj}\left(\frac{|t|}{\sqrt{-2\log(1-p/n)}}\right)\right\}.\end{aligned}$$

The remainder term $\to 0$, as $N\to\infty$, since $-2\log(1-p/n)\to\infty$. From inequality (5.4) and the fact that

$$-\log\left(1-\frac{p}{n+1}\right)\geqslant-\log\left(1-\frac{p}{n}\right)-\log 2$$

we find that

$$\varphi\left(\frac{t}{\sqrt{-2\log(1-p/n)}}\right)\to\exp\left\{\frac{(it)^2}{2}\right\}.$$

(ii) The c.f. of

$$\frac{1}{\sqrt{2\log n}}\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)!}$$

is

$$\exp\left\{it\frac{1}{\sqrt{2\log n}}\left(-\sum_{k=1}^{n}\frac{1}{k}+\log n\right)\right.$$
$$\left.+\frac{(it)^2}{2\cdot 2\log n}\sum_{k=1}^{n}\frac{2}{k}+O\left(\frac{|t|}{\sqrt{2\log n}}\right)\right\},$$

which evidently $\to \exp\{(it)^2/2\}$ as $N\to\infty$.

COROLLARY 5.2. *Under the conditions of* (5.2)

$$\frac{|\mathbf{C}_p^{(n+1)}|}{|\mathbf{\Sigma}_p|}\cdot\frac{n^p}{(n-1)_p}$$

converges in distribution to a lognormal distribution with parameters $\mu=0$ *and* $\sigma^2=-2\log(1-y)$, *as* $N\to\infty$.

The limit distribution of $|\mathbf{S}_p^{(n)}|/(n-1)_p$ in Theorem 5.1 has the expected value $e^{\mu+(1/2)\sigma^2}=1/(1-y)$ and the variance

$$e^{2\mu}\cdot e^{\sigma^2}(e^{\sigma^2}-1)=\frac{2y-y^2}{(1-y)^4},$$

while

$$E\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}=\frac{(n)_p}{(n-1)_p}=\frac{1}{1-y^*},\qquad y^*=\frac{p}{n},$$

and

$$\operatorname{Var}\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}=\frac{n^2}{(n-p)^2}\left\{\frac{n+1}{n-p+1}\cdot\frac{n+2}{n-p+2}-1\right\}$$
$$\approx\frac{2y^*-y^{*2}}{(1-y^*)^4}.$$

We have the following estimates for $\log(|\mathbf{S}_p^{(n)}|/(n-1)_p)$.

THEOREM 5.2. *If* $1\leqslant p<n$, *then*

(i) $$-\frac{1}{3n}\cdot\frac{y^*}{1-y^*}\leqslant E\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}\leqslant\frac{1}{n}\cdot\frac{y^*}{1-y^*}.$$

(ii) $$-\frac{2}{n}\cdot\frac{y^*}{1-y^*}\leqslant \operatorname{Var}\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}+2\log(1-y^*)$$
$$\leqslant\frac{3}{n}\cdot\frac{y^*}{1-y^*},\qquad y^*=\frac{p}{n}.$$

Proof. If $X\in\chi^2(f)$, then $E\log X=\psi(f/2)+\log 2$ and $\operatorname{Var}\log X=\psi'(f/2)$, where $\psi(f)=(\partial/\partial z)\log\Gamma(z)|_{z=f}$. Thus, by (5.1)

$$E\log|\mathbf{S}_p^{(n)}|=\sum_{j=1}^{p}\psi(\tfrac{1}{2}(n-p+j))+p\log 2,$$

$$\operatorname{Var}\log|\mathbf{S}_p^{(n)}|=\sum_{j=1}^{p}\psi'(\tfrac{1}{2}(n-p+j)).$$

(i) Expansion of $\psi(z)$ into series gives for $z\geqslant 1$

$$\psi(z)=\log z-\frac{1}{2z}-\frac{\theta}{12z^2},\qquad 0<\theta<1.$$

Thus

$$E\log\frac{|\mathbf{S}_p^{(n)}|}{(n-1)_p}=-\log(1-y^*)$$
$$-\sum_{j=1}^{p}\frac{1}{n-p+j}-\frac{\theta}{3}\sum_{j=1}^{p}\frac{1}{(n-p+j)^2}. \tag{5.5}$$

An upper bound is, by (5.4),

$$-\log(1-y^*)+\log\left(1-\frac{p}{n+1}\right)=\log\frac{n}{n-p}\cdot\frac{n-p+1}{n+1}$$
$$\leqslant\frac{p}{n(n-p)}=\frac{1}{n}\cdot\frac{y^*}{1-y^*}. \tag{5.6}$$

A lower bound of (5.5) is

$$-\frac{1}{3}\int_{n-p}^{n}\frac{1}{x^2}\,dx=-\frac{1}{3n}\frac{y^*}{1-y^*}.$$

(ii) Expansion of $\psi'(z)$ into series gives for $z\geqslant 1$

$$\psi'(z)=\frac{1}{z}+\left(\frac{1}{2z^2}+\frac{1}{6z^3}\right)\theta,\qquad 0<\theta<1.$$

It follows that

$$
\begin{aligned}
&\operatorname{Var}\log|\mathbf{S}_p^{(n)}| \\
&\quad = 2\cdot\sum_{j=1}^{p}\frac{1}{n-p+j} + 2\theta\sum_{j=1}^{p}\frac{1}{(n-p+j)^2} \\
&\qquad + \frac{4\theta}{3}\sum_{j=1}^{p}\frac{1}{(n-p+j)^3} \qquad (5.7)\\
&\quad \leqslant 2\int_{n-p}^{n}\frac{1}{x}\,dx + 2\int_{n-p}^{n}\frac{1}{x^2}\,dx + \frac{4}{3}\int_{n-p}^{n}\frac{1}{x_3}\,dx \\
&\quad = -2\log(1-y^*) + \frac{2}{n}\cdot\frac{y^*}{1-y^*} + \frac{2}{3}\cdot\frac{2y^*-y^{*2}}{n^2(1-y^*)^2} \\
&\quad \leqslant -2\log(1-y^*) + \frac{3}{n}\cdot\frac{y^*}{1-y^*} \quad \text{since } \frac{2-y^*}{n(1-y^*)} \leqslant \frac{3}{2} \text{ for } 1 \leqslant p < n.
\end{aligned}
$$

A lower bound of (5.7) is

$$
2\int_{n-p-1}^{n-1}\frac{1}{x}\,dx = -2\log\left(1-\frac{p}{n+1}\right) \geqslant -\frac{2}{n}\cdot\frac{y^*}{1-y^*} - 2\log(1-y^*).
$$

by (5.6).

EXAMPLE 5.1. Let $p=5$, $n=20$, i.e., $y^*=0{,}25$. The correct mean value is 0.003. The bounds are –0.006 and 0.017. The correct variance is 0.592. The approximating value is $-2\log 0.75 = 0.575$. The bounds are 0.542 and 0.625.

EXAMPLE 5.2. Let $p=10$, $n=40$, i.e., $y^*=0.25$. The correct mean value is 0.0014. The bounds are –0.0028 and 0.0083. The correct variance is 0.568. The approximating value is 0.575, while the bounds are 0.559 and 0.600.

The next theorem concerns the convergence rate of $\log|\mathbf{S}_p^{(n)}|$.

THEOREM 5.3. *Let $G_p^{(n)}(x)$ be the distribution function of*

$$
\frac{\log|\mathbf{S}_p^{(n)}| - E\log|\mathbf{S}_p^{(n)}|}{\sqrt{\operatorname{Var}\log|\mathbf{S}_p^{(n)}|}} = \frac{\sum_{j=1}^{p}(\log U_j - E\log U_j)}{\sqrt{\sum_{j=1}^{p}\operatorname{Var}\log U_j}}.
$$

Then

$$\sup_{-\infty<x<\infty} |G_p^{(n)}(x) - \Phi(x)| \leqslant \frac{1}{\sqrt{n}} \cdot \frac{C}{\sqrt{(1-y^*)\log(1/(1-y^*)) - y^*/n}},$$

where $y^* = p/n$ and C is an absolute constant, $1 \leqslant p < n$.

Proof. According to the Berry–Esseen theorem (Feller, [7b, p. 544])

$$\sup_{-\infty<x<\infty} |G_p^{(n)}(x) - \Phi(x)| \leqslant 6 \cdot \frac{\sum_{j=1}^p E\,|\log U_j - E\log U_j|^3}{(\sum_{j=1}^p \operatorname{Var}\log U_j)^{3/2}}. \tag{5.8}$$

By (5.3), $\log U_j$ has the c.f.

$$\varphi_j(t) = \frac{\Gamma(\frac{1}{2}(n-p+j)+it)}{\Gamma(\frac{1}{2}(n-p+j))} \cdot 2^{it}.$$

The rth cumulant, $\mathscr{K}_r^{(j)}$, i.e., the coefficient of $(it)^r/r!$ in the Taylor series expansion of $\log \varphi_j(t)$, is

$$\mathscr{K}_r^{(j)} = \frac{d^{r-1}}{dz^{r-1}}\,\psi(z)\,\bigg|_{z=\frac{1}{2}(n-p+j)}, \qquad \text{where } \psi(z) = \frac{d}{dz}\log\Gamma(z), \quad r = 2, 3,\ldots.$$

Let $\mu_r^{(j)}$ denote $E(\log U_j - E\log U_j)^r$, $r = 2, 3,\ldots$. We need the following relations between cumulants and moments:

$$\mu_2^{(j)} = \mathscr{K}_2^{(j)} = \psi'(\tfrac{1}{2}(n-p+j))$$
$$\mu_4^{(j)} = \mathscr{K}_4^{(j)} + 3(\mathscr{K}_2^{(j)})^2 = \psi'''(\tfrac{1}{2}(n-p+j)) + 3(\psi'(\tfrac{1}{2}(n-p+j)))^2.$$

Taylor series expansion for $z \geqslant 1$ gives

$$\psi'(z) = \frac{1}{z} + \theta \cdot \frac{1}{z^2}, \qquad 0 < \theta < 1,$$
$$\geqslant \frac{1}{z}$$
$$\psi'''(z) = \frac{2}{z^3} + 3\,\frac{\theta'}{z^4}, \qquad 0 < \theta' < 1,$$

i.e.,

$$\psi'''(z) + 3(\psi'(z))^2 \leqslant \frac{c_1}{z^2} \qquad \text{for some constant } c_1.$$

Thus

$$\mu_2^{(j)} \geqslant \frac{2}{n-p+j} \qquad \text{and} \qquad \mu_4^{(j)} \leqslant \frac{4c_1}{(n-p+j)^2}.$$

Summation over j gives

$$\begin{aligned}\sum_{j=1}^{p} \operatorname{Var} \log U_j &\geqslant -2 \log \left(1 - \frac{p}{n+1}\right) \\ &\geqslant -2 \log(1-y^*) - \frac{2}{n} \frac{y^*}{1-y^*};\end{aligned}$$

$$\begin{aligned}&\sum_{j=1}^{p} E\,|\log U_j - E \log U_j|^3 \\ &\quad\leqslant \sum_{j=1}^{p} (\mu_4^{(j)})^{3/4}, \qquad \text{by Hölder's inequality,} \\ &\quad\leqslant c_2 \sum_{j=1}^{p} \frac{1}{(n-p+j)^{3/2}} \qquad \text{for some constant } c_2, \\ &\quad\leqslant c_2 \int_{n-p}^{n} \frac{1}{x^{3/2}}\,dx \\ &\quad= 2c_2 \frac{1-\sqrt{1-y^*}}{\sqrt{n}\sqrt{1-y^*}} \\ &\quad\leqslant 2c_2 \frac{y^*}{\sqrt{n}\sqrt{1-y^*}\,(2-(n/(n+1))\,y^*)}.\end{aligned}$$

Thus, the right member of (5.8) has the upper bound

$$\frac{c_3}{\sqrt{n}} \frac{y^*}{\left(\sqrt{1-y^*}\,(2-(n/(n+1))\,y^*)(-\log(1-(n/(n+1))\,y^*)) \atop \sqrt{-\log(1-y^*) - y^*/n(1-y^*)}\right)}$$

for some constant c_3.

But

$$-\log\left(1 - \frac{n}{n+1} y^*\right) \geqslant \frac{2(n/(n+1))\,y^*}{2-(n/(n+1))\,y^*} \geqslant \frac{y^*}{2-(n/(n+1))\,y^*},$$

which gives the upper bound

$$\frac{C}{\sqrt{n}} \frac{1}{\sqrt{(1-y^*)\log(1/(1-y^*)) - y^*/n}}.$$

The bound is primarily of order $1/\sqrt{n}$ except when y^* is near 0 or near 1, i.e., when p is very small or near n.

ACKNOWLEDGMENT

My teacher, Professor Carl-Gustav Esseen, introduced me to the topic of this paper. I wish to thank him for his encouraging support, valuable advice and constructive criticism.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] ARHAROV, L. V. (1971). Limit theorems for the characteristical roots of a sample covariance matrix. *Soviet Math. Dokl.* **12**, 1206–1209.

[3] [a] ARNOLD, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.* **20**, 262–268. [b] ARNOLD, L. (1971). On Wigner's Semicircle Law for the Eigenvalues of Random Matrices. *Z. Wahrsch. Verw. Gebiete* **19**, 191–198.

[4] CARMELI, M. (1974). Statistical theory of energy levels and random matrices in physics. *J. Statist. Phys.* **10**, 259–297.

[5] CHUNG, K. L. (1968). A Course in Probability Theory. Harcourt, Brace & World, New York.

[6] ERDÉLYI, A., MAGNUS, W., OBERHETTINGER, F., AND TRICOMI, F. (1953). *Higher Transcendental Functions*, I. McGraw–Hill, New York.

[7] [a] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications*, I. Wiley, New York. [b] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, II. Wiley, New York.

[8] GRENANDER,U. (1963). *Probabilities on Algebraic Structures*. Almqvist & Wiksell, Stockholm.

[9] GRENANDER, U., AND SILVERSTEIN, J. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* **32**(2), 499–519.

[10] JONSSON, D. (1976). Some limit theorems for the eigenvalues of a sample covariance matrix. Technical Report No. 6, Department of Mathematics, Uppsala University, Uppsala.

[11] KARLIN,S. (1969). *A First Course in Stochastic Processes*. Academic Press, New York.

[12] KRISHNAIAH, P. R. (1978). Some recent developments on real multivariate distributions. *Develop. Statist.* **1**, 135–169.

[13] MARČENKO, V. A., AND PASTUR, L. A. (1967). Distributions of eigenvalus of some sets of random matrices. *Math. USSR-Sb.* **1**, 507–536.

[14] RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

[15] WACHTER, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6**, 1–18.

[16] [a] WIGNER, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* **62**, 548–564. [b] WIGNER, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Ann. of Math.* **67**, 325–327.