On the Moments of Elementary Symmetric Functions
of the Roots of Two Matrices

by

K. C. Sreedharan Pillai
Purdue University

Division of Mathematical Sciences
Department of Statistics
Mimeograph Series No. 17
May, 1964
On the Moments of Elementary Symmetric Functions
of the Roots of Two Matrices
by
K. C. Sreedharan Pillai
Purdue University

1. Summary. A lemma is given first which provides an easy method of expressing the product of an \( s \)th order Vandermonde type determinant and the \( k \)th and \( l \)th \((k, l \geq 0)\) powers of the \( r \)th and \( h \)th \((r, h \leq s)\) elementary symmetric functions \((\text{esf's})\) respectively as a linear compound of determinants. The lemma extends itself readily to the product of powers of any number of esf's up to the \( s \)th. Using this lemma and some reduction formulae for certain special types of Vandermonde type determinants, a second lemma has been proved to show that the moments of esf's in \( s \) non-null characteristic roots \( \lambda_1(0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_s < \infty) \) of a matrix can be easily derived from the respective moments of corresponding esf's in \( s \) non-null roots, \( \theta_i, (0 < \theta_1 \leq \ldots \leq \theta_s < 1) \) of another matrix and vice versa. Illustrations are given explaining both lemma.

2. Introduction. Many of the distribution problems in multivariate analysis are based on the distribution of the non-null characteristic roots of a matrix derived from sample observations taken from multivariate normal populations. This distribution given by Fisher [1], Girshick [2], Hsu [3] and Roy [12] is of the form

\[
(2.1) \quad f(\theta_1, \theta_2, \ldots, \theta_s) = C(s, m, n) \prod_{i=1}^{s} \theta_i^m (1-\theta_i)^n \prod_{i>j} (\theta_i - \theta_j),
\]

\(0 < \theta_1 \leq \ldots \leq \theta_s < 1\)

where
\[ (2.2) \quad C(s, m, n) = \prod_{i=1}^{s/2} \prod_{i=1}^{s} \frac{\Gamma\left(\frac{2m+2n+i+2}{2}\right)}{\Gamma\left(\frac{2m+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \Gamma(1/2)} \]

and \( m \) and \( n \) are defined differently for various situations described in [7], [9].

Now, if \( \lambda_i = \theta_i/(1-\theta_i), \) \( i = 1, 2, \ldots, s, \) the joint distribution of the \( \lambda \)'s is obtained from (2.1) as

\[ (2.3) \quad f(\lambda_1, \lambda_2, \ldots, \lambda_s) = C(s, m, n) \prod_{i=1}^{s} \frac{\lambda_i^{m}/(1+\lambda_i)^{m+n+1}}{\prod_{i=1}^{s} \prod_{i<j} (\lambda_i - \lambda_j)} \]

\[ 0 < \lambda_1 \leq \ldots \leq \lambda_s < \infty. \]

The studies on the first esf in \( \theta \)'s as well as the \( \lambda \)'s have been carried out by Pillai [6], [8], [9], Pillai and Mijares [10] and Pillai and Samson [11]. Mijares [12] has carried out some studies on esf's in general. In this paper, a lemma is proved which enables one to write down easily the moments of \( U^{(s)}_{i, m, n} \) from the respective moments of \( V^{(s)}_{i, m, n} \) and vice versa, where \( U^{(s)}_{i, m, n} \) and \( V^{(s)}_{i, m, n} \) denote the \( i \)th esf's in the \( s \) \( \lambda \)'s and \( s \) \( \theta \)'s respectively. But first, a lemma is given (see next section) on which will be based the proof of the main lemma showing the easy derivation of the moments of \( U^{(s)}_{i, m, n} \) from the respective moments of \( V^{(s)}_{i, m, n} \).

3. Product of a Vandermonde determinant and powers of esf's. In this section we introduce the following lemma:

**Lemma 1.** Let \( D(g_s, g_{s-1}, \ldots, g_1), \) \( (g_j > 0, \ j = 1, 2, \ldots, s), \) denote the determinant
If \( a_r(r \leq s) \) denotes the \( r \)th esf in \( s \) \( x \)'s, then

\[
(3.1) \quad D(g_s', g_{s-1}', \ldots, g_1') = \sum_{\binom{s}{r}} D(g_s, g_{s-1}, \ldots, g_1),
\]

where \( g_j' = g_j + \delta_j \) \( j = 1, 2, \ldots, s, \delta = 0, 1 \) and \( \sum \) denotes the sum over the \( \binom{s}{r} \) combinations of \( s \) \( g \)'s taken \( r \) at a time for which \( r \) indices \( g_j' = g_j + 1 \) such that \( \delta = 1 \) while for other indices \( g_j' = g_j \) such that \( \delta = 0 \).

\[
(3.2) \quad (i) \quad a_r D(g_s, g_{s-1}, \ldots, g_1) = \sum_{\binom{s}{r}} D(g_s', g_{s-1}', \ldots, g_1'),
\]

where \( h \leq s, g_j'' = g_j' + \delta_j \) \( j = 1, 2, \ldots, s, \delta = 0, 1 \) and \( \sum \) denotes summation over the \( \binom{s}{r} \) terms obtained by taking \( h \) at a time of the \( s \) \( g \)'s in each \( D \) in \( \sum \) in (3.2) for which \( h \) indices \( g_j'' = g_j' + 1 \) while for other indices \( g_j'' = g_j' \).

\[
(3.3) \quad (ii) \quad a_r a_h D(g_s, g_{s-1}, \ldots, g_1) = \sum_{\binom{s}{r}} D(g_s'', g_{s-1}'', \ldots, g_1'')
\]

\[
(iii) \quad (a_r)^k (a_h)^l D(g_s, g_{s-1}, \ldots, g_1) (k, l \geq 0) \text{ can be expressed as a sum of } \binom{s}{r} \binom{s}{h} \text{ determinants obtained by performing on } D(g_s, g_{s-1}, \ldots, g_1) \text{ in any}
\]
order (i) \( k \) times and (i) \( l \) times with \( r = h \).

However, if at least two of the indices in any determinant are equal, the corresponding term in the summation vanishes.

Before indicating a proof of the lemma, let us consider an illustration. Let us note first that \([4]\)

\[
(3.4) \quad a_w = \sum (-1)^{i-1} \prod_{i=1}^{w} \frac{p_1^{s_1} p_2^{s_2} \ldots p_w^{s_w} (1^{p_1} 2^{p_2} \ldots w^{p_w} p_1! p_2! \ldots p_w!)}{s_1 s_2 \ldots s_w},
\]

where \( \Sigma \) extends over all non-negative values of \( p_1, \ldots, p_w \) such that

\[
p_1 + 2p_2 + \ldots + wp_w = w,
\]

and

\[
s_k = \sum_{j=1}^{s} x_j^k.
\]

Also note that if we multiply the right side of (3.1) by \( e^{tsu} \), differentiate with respect to \( t \) once and put \( t = 0 \) we get,

\[
(3.5) \quad s_u D(g_s, g_{s-1}, \ldots, g_1) = \sum_{j=1}^{s} D(g_s, g_{s-1}, \ldots, g_{s-1}, g_{s+1}, g_s + w g_{s-1}, \ldots, g_1).\]

Now consider the special case, \( w = 4 \). We get from (3.4)

\[
(3.6) \quad a_4 = \frac{s_1^4}{4!} + \frac{s_2^2}{2} - \frac{s_3^2}{4} + \frac{s_4^3}{3} - \frac{s_4}{4}.
\]

Using the right side of (3.6) and by repeatedly applying (3.5) with varying values of \( u \) (from 1 to 4) we get
\[
(3.7) \quad a_4 D(g_s, g_{s-1}, \ldots, g_j) = b_0 \sum_{j=1}^{s} D(g_s, g_{s-1}, \ldots, g_{j+1}, g_j+1, g_{j-1}, \ldots, g_1) + b_1 \sum_{j'=1}^{s} D(g_s, g_{s-1}, \ldots, g_j+1, g_{j-1}, \ldots, g_{j'+1}, \ldots, g_1) + b_2 \sum_{j'=j''=1}^{s} D(g_s, g_{s-1}, \ldots, g_j+2, \ldots, g_{j''}, +1, \ldots, g_1) + b_3 \sum_{j'=1}^{s} D(g_s, g_{s-1}, \ldots, g_j, +1, g_{j''}, +1, \ldots, g_{j}, +1, \ldots, g_1) + b_4 \sum_{j'=j''=1}^{s} D(g_s, g_{s-1}, \ldots, g_j, +1, g_{j'}, +1, \ldots, g_1) + \ldots, g_{j''}, +1, \ldots, g_1)
\]

where

\[
b_0 = -\frac{1}{4} + \frac{1}{3} = -\frac{1}{4} + \frac{1}{3} + \frac{1}{8} + \frac{1}{41} = 0,
\]

\[
b_1 = \frac{1}{3} - \frac{1}{2} + \frac{1}{41} = 0,
\]

\[
b_2 = \frac{1}{21} \left( -\frac{2}{4} + \frac{12}{41} \right) = 0,
\]

\[
b_3 = \frac{1}{21} \left( \frac{2}{8} - \frac{2}{4} + \frac{6}{41} \right) = 0, \quad b_4 = \frac{1}{41}
\]

and where the indices \( j, j', j'', j''' \) are the only ones which have been increased.

Now since in the last sum the right side of (3.7) there are only \( \binom{8}{4} \) distinguishable terms, it is obvious that \( a_4 D(g_s, g_{s-1}, \ldots, g_1) \) is obtained from (3.7) as a sum of \( \binom{8}{4} \) determinants whose indices are obtained by selecting 4 out of s g's at a time and increasing each of the 4 selected g's by unity.

Now consider the general case (i). Apply (3.4) to (3.1) with \( w = r \).

We can show that the coefficient of the determinant of a specified set of indices
obtained in this operation such that at least one \( g \) on the left side of (3.2) has been increased by more than unity, is equal to zero. For instance, the coefficient of the determinant with one \( g \) index increased by \( r - j > 1 \) while any other \( j \) \( g \)'s increased each by unity is given by

\[
(3.8) \quad (1/j!) \sum_{i=1}^{r-j} \frac{r-j}{\prod_{i=2}^{r-j} (r-j)^{p_{r-j}}} \frac{1}{\prod_{i=1}^{r-j} p_1 \cdots p_{r-j}} = 0
\]

where \( p_1 + 2p_2 + \cdots + (r-j)p_{r-j} = r-j \).

In a similar manner, coefficients of all other determinants with at least one index increased by more than unity can be shown to be equal to zero. There remain, therefore, only determinants in which \( r \) out of \( s \) indices have been increased just by unity. It may be observed that this last set of determinants is obtained from the term \( s_1^r / r! \) in (3.4), \( w=r \), while all other terms arise from more than one term in (3.4), \( w=r \), and their coefficients are obtained as sums of positive and negative values where each sum (coefficient) equals zero.

Now it may be seen that the truth of (ii) in lemma 1 can be observed easily by an application of (3.4) to the right side of (3.2) with \( w = h \).

Similarly (iii) further follows easily by repeated application of (i), as stated in the lemma, \( k + \ell \) times, \( k \) times using (3.4), \( w=r \), and \( \ell \) times using (3.4) with \( w = h \). In addition, it may be pointed out that the method of proof extends itself to generalize (iii) further to include powers of any number of esf's up to \( \frac{h^s}{s!} \).

4. Derivation of moments of \( V^{(s)}_{i,m,n} \) from those of \( V^{(s)}_{i,m,n} \). In this section we prove the main lemma.

Proof. Let
\[ v(m+s-1+q_s, \ldots, m+q_{\lambda}; n) = \]
\[
\int_{0}^{1-m+s-1+q_s} \cdots \int_{0}^{1-m+q_{\lambda}} (1-\theta_s)^n d\theta_s \cdots \int_{0}^{1-m+q_{\lambda}} (1-\theta_1)^n d\theta_1 \]

and let

\[ u(m+s-1+q_s, \ldots, m+q_{\lambda}; n) = \]
\[
\int_{0}^{c_2} \cdots \int_{0}^{c_2} \frac{m+s-1+q_s}{(l+\lambda_s)^P} d\lambda_s \cdots \frac{m+q_{\lambda}}{(l+\lambda_{\lambda})^P} d\lambda_{\lambda} \]

\[ q_j > 0 \quad j = 1, 2, \ldots, s, \]

\[ p = m+n+s+1. \]

Now, from Lemma 1 and (2.1), the kth moment of \( v_{i,m,n}^{(s)} \), \( \mu_k \{ v_{i,m,n}^{(s)} \} \), can be expressed as a linear compound of determinants of the V-type in (4.1) where \( q_s, q_{s-1}, \ldots, q_{\lambda} \) may take different sets of values in different terms. Further, the coefficients of the linear compound would involve as a common factor \( C(s,m,n) \) but otherwise would be independent of \( m \) and \( n \).

Similarly, \( \mu_k \{ u_{i,m,n}^{(s)} \} \) can be shown to be a linear compound of the determinants of the U-type in (4.2) with the coefficients of corresponding terms in this compound the same as in the previous compound, the correspondence of terms being marked by the equality of the vector \( (q_s, q_{s-1}, \ldots, q_{\lambda}) \) in the two compounds.

Now we state the lemma.

**Lemma 2.** \( \mu_k \{ u_{i,m,n}^{(s)} \} \) is derivable from \( \mu_k \{ v_{i,m,n}^{(s)} \} \) by making the following changes in the expression for the latter (obtained by evaluating the linear compound
of V-type determinants): (a) Multiply by \(-1\) all terms except the term in \(n\) in each linear factor involving \(n\) and (b) change \(n\) to \(m+n+1\) after performing (a).

Before proving the lemma let us illustrate it by a couple of special cases.

Using (i) of lemma 1 we get

\[
(4.3) \quad \mu_1^i \left\{ V_{i,m,n}^{(s)} \right\} = \sigma(s,m,n) V(m+s, m+s-1, \ldots, m+s-i+1, m+s-i, \ldots, m+1, m).
\]

The right side of (4.3) can be shown to be equal to

\[
(4.4) \quad \binom{s}{i} \prod_{j=1}^{i} \frac{(2m+i-j+2)}{(2m+2n+2s-j+2)}
\]

based on some particular cases of determinants evaluated in \([10]\). From this result using lemma 2, the first moment of the \(i\)th esf in the \(\lambda\)'s is given by

\[
(4.5) \quad \mu_1^i \left\{ V_{i,m,n}^{(s)} \right\} = \binom{s}{i} \prod_{j=1}^{i} \frac{(2m+i-j+2)}{(2n+j-1)}.
\]

Now consider \(\mu_2^i \left\{ V_{2,m,n}^{(s)} \right\}\). Using (ii) of lemma 1 with \(h = r\) we get

\[
(4.6) \quad \mu_2^i \left\{ V_{2,m,n}^{(s)} \right\} = \sigma(s,m,n) \left\{ V(m+s+1, m+s-3, m+s-3, \ldots, m+1, m) \right\} \cdot
\]

\[
+ V(m+s+1, m+s-2, m+s-2, m+s-4, \ldots, m+1, m) \cdot
\]

\[
+ V(m+s, m+s-1, m+s-2, m+s-3, m+s-5, \ldots, m+1, m) \right\}.
\]

Now substituting the values of the determinants \([10]\) in (4.6)
(4.7) \[ \mu_2 \left\{ \begin{bmatrix} v(s) \\ 2, m, n \end{bmatrix} \right\} = \frac{s(s-1)(2m+s)(2m+s+1)}{3!} \prod_{j=1}^{n} (2m+2n+2s-j+5) \cdot g \]

where \( g = 6m^2 + 2s(s-1)(2s^2+s+8) m^4 + 7s^2 - 8s + 12 \)

+ \( 3n \left\{ 16s(s-1)m^2 + 16(s-1)(8s^2 + 5s + 8)m^2 + 2(s-1)(10s^3 + 12s^2 + 27s + 24)m \right. \)

+ \( 4s^5 + 7s^4 + 12s^3 + 12s^2 - 24s + 36 \left. \right\} + s(s+1)(2n+s+1)(2m+s+2)(m+s)(2m+2s+1) \)

+ \( (s-2)(2m+2s+3)(2m+s+1) \left\{ 4m^2 + 2s(3s+2) m^3 + 2s^3 + 3s^2 + s + 6 \right\} \).

Using lemma 2 we get from (4.7),

(4.8) \[ \mu_2 \left\{ \begin{bmatrix} v(s) \\ 2, m, n \end{bmatrix} \right\} = \frac{s(s-1)(2m+s)(2m+s+1)}{3!} \prod_{j=1}^{n} (2m+j-3) \cdot g \]

where \( g \) is obtained from \( g \) by attaching negative sign to the first degree term in \( n \) and then changing \( n \) to \( m+s+1 \) in all the terms.

**Proof.** Apply theorem 3 of [1] to the V-determinant in (4.1). We get

(4.9) \[ V(m+s-l+q_s, \ldots, m+t_j; n) = (m+s+t_s+n)^{-1} (B(s) + (m+s-l+q_s)C(s)) \]

where

(4.10) \[ B(s) = \sum_{j=s-1}^{l} (-1)^{s-j-1} V(2m+s+j-2+q_s+q_j; 2m+l) X \]

\[ X V(m+s-2+q_s-1, \ldots, m+j+q_j+1, m+j-2+q_j-1, \ldots, m+q_j; n) \]

and

(4.11) \[ C(s) = V(m+s-2+q_s, m+s-2+q_s-1, \ldots, m+q_j; n). \]
Again, applying theorem 4 of [8] to the U-determinant in (4.2) we get

\[(4.12) \quad U(m+s-1+q_s, \ldots, m+q_j; p) = \left[p-(m+s+q_s)\right]^{-1} (q(s)+(m+s-1+q_s)R(s))\]

where

\[(4.13) \quad q(s) = \frac{1}{2} \sum_{j=s-1}^{s} (-1)^{s-j-1} U(2m+s+j-2+q_s+q_j; 2p-1) X \]

\[X \ U(m+s-2+q_s-1, \ldots, m+j+q_j+1, m+j-2+q_j-1, \ldots, m+q_j; p)\]

and

\[(4.14) \quad R(s) = U(m+s-2+q_s, m+s-2+q_s-1, \ldots, m+q_j; p) .\]

First, it may be observed that the factor \(m+s+q_s+n\) in (4.9) becomes the factor \(p-(m+s+q_s)\) in (4.12) by changes (a) and (b) of the lemma. Further, repeated application of theorem 3 of [8] to the right side of (4.9) would reduce it to a linear compound of terms each of which is a product of \(s/2\) simple beta functions of type I (U-type) if \(s\) is even and \((s+1)/2\) beta functions if \(s\) is odd. The coefficients of this linear compound would involve products of functions of \(m\) and \(n\) of the type \((m+j+q_j+n)^{-1}\) and the type \((m+j-1+q_j)\) as in (4.9). Similarly, repeated application of theorem 4 of [8] to the right side of (4.12) would reduce it to a linear compound as above with the exception that simple beta functions involved will be of type II (U-type) instead of type I and \([p-(m+j+q_j)]^{-1}\) will replace \((m+j+q_j+n)^{-1}\). Now it may be observed that changes (a) and (b) of the lemma will make the corresponding coefficients the same in the two linear compounds which are obtained after repeated applications of theorems 3 and 4 of [8] to (4.9) and (4.12) respectively. It remains, therefore, to show that \(C(s,m,n)\) times each term of the linear compound involving products of beta functions of type I reduces to \(C(s,m,n)\) times the corresponding term in the second linear compound involving
the product of beta functions of type II using (a) and (b) of the lemma. Now note that, if \( s \) is even,

\[
(4.15) \quad C(s,m,n) = 2^{-s(s+6)/8} X
\]

\[
X \frac{s/2}{\prod_{i=1}^{s/2} \Gamma(2m+2n+s+2i+1)} \left( \Gamma(2m+2i)\Gamma(2n+2i)\Gamma(1) \right)^{(1.3)(1.3.5)...(1.3.5...(s-3))} \]

and if \( s \) is odd

\[
(4.16) \quad C(s,m,n) = 2^{-(s-1)(s+5)/8} X
\]

\[
X \frac{(s-1)/2}{\prod_{i=1}^{(s-1)/2} \Gamma(2m+2n+s+2i+1)\Gamma(m+n+s+1)} \left( \Gamma(2m+2i)\Gamma(2n+2i)\Gamma(1)\Gamma[(2m+s+1)/2]\Gamma[(2n+s+1)/2] \right)^{(1.3).} \]

\[
... (1.3.5)(1.3.5...(s-2)) \]

Now, for \( s \) even, consider the term

\[
(4.17) \quad C(s,m,n) V(2m+2s+3+q_s+q_{s-1};2n+1) X
\]

\[
X V(2m+2s-7+q_{s-2}+q_{s-3};2n+1) ... V(2m+1+q_2+q_1;2n+1). \]

Substitute in (4.17) the value of \( C(s,m,n) \) from (4.15) and those of the type I beta functions and we get
\begin{equation}
\begin{aligned}
g(m,s) &\left[ (2n+2s-3+q_s+q_{s-1}+3) \ldots (2n+1) \right] \left[ (2n-2s-3+q_s+q_{s-1}+3) \ldots (2n+1) \right] \\
&\ldots \left[ (2n+2s-3+q_s+q_{s-1}+3) \ldots (2n+1) \right] \\
\end{aligned}
\end{equation}

where \( g(m,s) \) is a function of \( m \) and \( s \).

Similarly consider

\begin{equation}
\begin{aligned}
C(s,m,n) U(2n+2s-3+q_s+q_{s-1};2p+1) & U(2n+2s-7+q_s-2+q_{s-3};2p+1) \ldots \\
&\ldots U(2n+1+q_s+q_{s-1};2p+1) .
\end{aligned}
\end{equation}

After substitution of values of \( C(s,m,n) \) and \( U's \) in (4.19) we get

\begin{equation}
\begin{aligned}
g(m,s) &\left[ (2n-q_s-q_{s-1}+3) \ldots (2n+1) \right] \left[ (2n-q_s-2q_{s-3}+7) \ldots (2n+1) \right] \\
&\ldots \left[ (2n+2s-q_s-q_{s-1}+3) \ldots (2n+s-1) \right] \left[ (2n+2n+s+3) \ldots (2n+2n+2s) \right] \\
\end{aligned}
\end{equation}

\begin{equation}
\begin{aligned}
\left[ (2n+2n+s+5) \ldots (2n+2n+2s) \right] \ldots \left[ (2n+2n+2s-1)(2n+2n+2s) \right] .
\end{aligned}
\end{equation}

Now it may be noted that (4.20) can be obtained from (4.18) by (a) and (b) of the lemma. In a similar manner, when \( s \) is even, other corresponding terms of the linear compounds in the two cases can be shown to satisfy (a) and (b) of the lemma.

If \( s \) is odd, we may consider the terms like

\begin{equation}
\begin{aligned}
C(s,m,n)V(2m+2s-3+q_s+q_{s-1};2n+1) \ldots V(2m+3+q_s+q_{s-1};2n+1)V(m+q_s;n) .
\end{aligned}
\end{equation}

Using (4.16) and the values of the \( V's \) and performing (a) and (b) in (4.21) we will get
\[ C(s, m, n) U(2m+2s-3+q_3+q_{s-1};2p-1) \ldots U(2m+3+q_3+q_{2};2p-1) U(m+q_1;p) . \]

Similarly, if \( s \) is odd, we can show that other corresponding terms of the linear compounds in the two cases satisfy (a) and (b).

Hence the lemma.

It may be noted that \( \mu'_{\mathcal{R}} \{ v_{i, m, n} \} \) may similarly be derived from \( \mu'_{\mathcal{R}} \{ u_{i, m, n} \} \) by inverse operations of (b) and (a) of lemma 2. Further, lemma 2 readily extends to the case of product moments (say of the \( r \)th and \( h \)th esf's) in view of (ii) of lemma 1.
REFERENCES


