DISTRIBUTION OF CERTAIN FACTORS USEFUL IN DISCRIMINANT ANALYSIS

R. P. GUPTA AND D. G. KABE

(Received Sept. 9, 1969; revised May 1, 1970)

1. Introduction and summary

Let $A$ be a $(p \times p)$ symmetric positive definite matrix having the noncentral Wishart density.

\begin{equation}
(1.1) \quad f(A) = \exp \left( - \text{tr} \left( \Sigma^{-1} \Omega \right) \right) \frac{\Gamma_p \left( \frac{1}{2} q, \frac{1}{2} \Sigma^{-1} \Omega \Sigma^{-1} A \right)}{W(A | \Sigma | q)}
\end{equation}

where

\begin{equation}
(1.2) \quad W(A | \Sigma | q) = \frac{|A|^{(q-p-1)/2} \exp \left( - \frac{1}{2} \text{tr} \Sigma^{-1} A \right)}{2^{2p/4} \Gamma_p \left( \frac{1}{2} q \right) |\Sigma|^{q/2}}
\end{equation}

and

\begin{equation}
(1.3) \quad \Gamma_p \left( \frac{1}{2} q \right) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma \left( \frac{1}{2} (q+1-i) \right),
\end{equation}

and \( _pF_1(q/2, \Sigma^{-1} \Omega \Sigma^{-1} A/2) \) is a hypergeometric function of matrix arguments, see ([7], p. 733). Let $B$ be another $(p \times p)$ symmetric positive definite matrix, having central Wishart density

\begin{equation}
(1.4) \quad f(B) = W(B | \Sigma | n-q).
\end{equation}

Assuming the matrix $\Omega$ to be of rank $s<p$ we make the transformations

\begin{equation}
(1.5) \quad A = C(I - L)C', \quad B = CC',
\end{equation}

where $C$ is a lower triangular matrix of order $p$. The noncentral multivariate beta density of the $(p \times p)$ matrix $L$ is found by Radcliffe ([7], p. 734) to be

* This research was supported by the National Research Council of Canada grant.
\[ g(L) = |L|^{(q-p-s-1)/2} |I-L|^{(q-p-s-1)/2} \theta(L_{11}) , \]

where \( \theta(L_{11}) \), see Radcliffe ([7], p. 734), involves only the elements of \( s \times s \) matrix \( L_{11} \) and other parameters but not any other elements of \( L \).

The noncentral multivariate beta density of \( L \) is a direct generalization of the noncentral linear beta density of rank one of \( L \) as given by Kshirsagar [5], who used the density of \( L \), to derive the distribution of the test criterion for testing the adequacy of a single hypothetical discriminant function. Radcliffe generalizes Kshirsagar's results and gives the test criterion for testing the adequacy of \( s (\leq p) \) hypothetical discriminant functions. If \( \Gamma'x \), where \( \Gamma' \) is an \( s \times p \) matrix of rank \( s \), denote the \( s \) discriminant functions, then \( A=|L| \) may be factorized as

\[ A = A_1 A_2 |L_{11}| \]

where the direction and collinearity factors \( A_1 \) and \( A_2 \) are

\[ A_1 = \frac{|\Gamma'AB^{-1}(B-A)\Gamma'\Gamma'\Gamma|}{|\Gamma'(B-A)\Gamma|} \]

\[ A_2 = \frac{A}{|L_{11}|} \frac{|\Gamma'(B-A)\Gamma'\Gamma'\Gamma|}{|\Gamma'\Gamma'\Gamma|} \frac{|\Gamma'AB^{-1}(B-A)\Gamma|}{|\Gamma'(B-A)\Gamma|} . \]

It may be noted that the factorization of \( A \), given here, is a generalization of the factorization given by Bartlett [2].

By choosing \( \Gamma'=(I,0) \) where \( I \) is an \( s \times s \) identity matrix and factorizing the density of \( L \) in terms of rectangular coordinates \( T, L=TT' \), \( T \) a lower triangular, Radcliffe [7] expresses the densities of \( A_1 \) and \( A_2 \) in terms of the elements of \( T \). He also gives another factorization of \( A \) as, Radcliffe ([7], p. 732),

\[ A = A_5 A_6 |L_{11}| , \]

where

\[ A_5 = \frac{|B-A||\Gamma'AF' + \Gamma'PA(B-A)^{-1}AF'|}{|B||\Gamma'AF'|} \]

\[ A_6 = \frac{|\Gamma'BF'\Gamma'|}{|\Gamma'(B-A)\Gamma'\Gamma'AF' + \Gamma'PA(B-A)^{-1}AF'|} \]

\( A_5 \) and \( A_6 \) are also useful for testing direction and collinearity of the hypothetical discriminant functions \( \Gamma'x \). Following Kshirsagar's [6] method, Radcliffe expresses \( A_5 \) and \( A_6 \) as functions of the elements of \( T \) and obtains their distributions. We are giving here a shorter and neater proof, which might be of pedagogical interest. Also our main interest is to express \( A_1, A_2, A_5 \) and \( A_6 \) as functions of the elements of \( L \), rather than functions of elements of \( T \). All distributions are derived without
the constant factor, \( K \) is used as a generic symbol for the constant factors of the density functions.

2. Distribution of \( A_1 \) and \( A_2 \)

By partitioning \( L \) and \( I - L \) as

\[
L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \quad I - L = \begin{pmatrix} I - L_{11} & -L_{12} \\ -L_{21} & I - L_{22} \end{pmatrix}
\]

and using (1.6) we write the joint density of \( L_{11}, L_{12}, L_{22} \), as,

\[
f(L_{11}, L_{12}, L_{22}) = K \left| L_{11} \right|^{(s - q - p - 1)/2} \left| I - L_{11} \right|^{(q - p - 1)/2} \theta(L_{11}) \\
\cdot \left| L_{22} - L_{21} L_{11}^{-1} L_{12} \right|^{(s - q - p - 1)/2} \\
\cdot \left| I - L_{22} - L_{21} (I - L_{11})^{-1} L_{12} \right|^{(q - p - 1)/2}.
\]

On setting

\[
Z = L_{22} - L_{21} L_{11}^{-1} L_{12} \\
R = (I - L_{21} (I - L_{11})^{-1} L_{12})^{-1/2} Z (I - L_{21} (I - L_{11})^{-1} L_{12})^{-1/2}.
\]

We find that the joint density of \( L_{11}, L_{12} \), and \( R \) is given by

\[
f(L_{11}, L_{12}, R) = K \left| L_{11} \right|^{(s - q - p - 1)/2} \left| I - L_{11} \right|^{(q - p - 1)/2} \theta(L_{11}) \\
\cdot \left| I - L_{21} (I - L_{11})^{-1} L_{12} \right|^{(s - q - s - 1)/2} \\
\cdot \left| R \right|^{(s - q - p - 1)/2} \left| I - R \right|^{(q - p - 1)/2}.
\]

Again we set \( \Delta = L_{21} (I - L_{11})^{-1} L_{12} \) and assuming \((p - s) \leq s\) we use Hsu’s lemma, (Anderson [1], p. 319, Lemma 13.3.1) to integrate (2.4) with respect to the elements of \( L_{12} \) and find the joint density of \( L_{11}, \Delta \), and \( R \) to be

\[
f(L_{11}, \Delta, R) = K \left| L_{11} \right|^{(s - q - s - 1)/2} \left| I - L_{11} \right|^{(s - s - 1)/2} \theta(L_{11}) \\
\cdot \left| I - \Delta \right|^{(s - q - s - 1)/2} \left| \Delta \right|^{(s - s - 1)/2} \\
\cdot \left| R \right|^{(s - q - p - 1)/2} \left| I - R \right|^{(q - p - 1)/2}.
\]

By setting \( \Gamma'(I, 0) \), it may be easily seen that

\[
A_1 = \left| I - \Delta \right| \quad \text{and} \quad A_2 = \left| R \right|.
\]

It follows from (2.5) that the densities of \( A_1 \) and \( A_2 \) are mutually independent. The densities of \( A_1 \) and \( A_2 \) are identical with those of a product of independent beta variates. This result agrees with the one given by Radcliffe ([7], p. 738), except the fact that we assume \( p \leq 2s \) and Radcliffe assumes \( p \geq 2s \).
3. Distribution of $A_s$ and $A_t$

Noting that,

\begin{equation}
A_s = \frac{|z|}{|z + L_{11}(L_{11}(I - L_{11}))^{-1} L_{11}|}
\end{equation}

$A_t = |z + L_{21}(L_{11}(I - L_{11}))^{-1} L_{11}|$, we set $z = PP'$, where $P$ is a nonsingular matrix of order $(p-s) \times (p-s)$. The joint density of $L_{11}$, $P$ and $L_{12}$ may be obtained by using the result (2.3), and we find that

\begin{equation}
f(L_{11}, P, L_{12}) = K |L_{11}|^{(a - q - p - 1)/2} |I - L_{11}|^{(q - p - 1)/2} \theta(L_{11})
\cdot |I - PP' - L_{21}(L_{11}(I - L_{11}))^{-1} L_{11}|^{(q - p - 1)/2}
\cdot |PP'|^{(a - q - p)/2}.
\end{equation}

Further transforming $L_{11}$ to $\eta$, where $\eta$ is an $(p-s) \times s$, by the relation

\begin{equation}
L_{11} = P\eta
\end{equation}

the joint density of $L_{11}$, $P$ and $\eta$ is found to be

\begin{equation}
f(L_{11}, P, \eta) = K |L_{11}|^{(a - q - p - 1)/2} |I - L_{11}|^{(q - p - 1)/2} \theta(L_{11})
\cdot |P(I + \eta(L_{11}(I - L_{11}))^{-1} \eta')P'|^{(a - q - p + s)/2}
\cdot |I + \eta(L_{11}(I - L_{11}))^{-1} \eta'|^{(a - q - p + s)/2}
\cdot |I - P(I + \eta(L_{11}(I - L_{11}))^{-1} \eta')P'|^{(q - p - 1)/2}.
\end{equation}

Now we set

\begin{equation}
P(I + \eta(L_{11}(I - L_{11}))^{-1} \eta')P' = W
\end{equation}

and using Hsu’s lemma (Anderson [1], Lemma 13.3.1) we find the joint density of $W$, $\eta$ and $L_{11}$ to be

\begin{equation}
f(L_{11}, W, \eta) = K |L_{11}|^{(a - q - p - 1)/2} |I - L_{11}|^{(q - p - 1)/2} \theta(L_{11})
\cdot |I + \eta(L_{11}(I - L_{11}))^{-1} \eta'|^{(a - q)/2}
\cdot |W|^{(a - q - p + s - 1)/2} |I - W|^{(q - p - 1)/2}.
\end{equation}

Further setting

\begin{equation}
\eta(L_{11}(I - L_{11}))^{-1} \eta' = G
\end{equation}

and using (3.6) and Hsu’s lemma we get

\begin{equation}
f(L_{11}, G, W) = K |L_{11}|^{(a - q - s - 1)/2} |I - L_{11}|^{(q - s - 1)/2} \theta(L_{11})
\cdot |I + G|^{(a - q)/2} |G|^{(2a - p - 1)/2}
\cdot |W|^{(a - q - p + s - 1)/2} |I - W|^{(q - p - 1)/2}.
\end{equation}
Again transforming $G$ to $H$ by the transformation

\[(3.9) \quad H = (I+G)^{-1}\]

and noting that the Jacobian of the transformation from $H$ to $G$ is $|I+G|^{-(p-s+1)}$ we obtain the joint density of $L_{11}$, $H$ and $W$ to be

\[(3.10) \quad f(L_{11}, H, W) = K |L_{11}|^{(n-q-p-1)/2} |I-L_{11}|^{(q-s-1)/2} \theta(L_{11})
\cdot |H|^{(n-p-s-1)/2} |I-H|^{(2s-p-1)/2}
\cdot |W|^{(n-q-p+s-1)/2} |I-W|^{(q-p-1)/2}.

Here we note that $A_4 = |H|$ and $A_5 = |W|$. It, thus, follows that the densities of $A_4$ and $A_5$ are independent. This result agrees with the one given by Radcliffe ([7], p. 739), except that we assume $p \leq 2s$ while Radcliffe assumes $p \geq 2s$.

4. Distribution of $A_4$ and $A_5$

We have noted above that the $A_i$ is distributed as a product of $(p-s)$ independent beta variables and as such we must be able to factorize $A_i$ into $(p-s)$ mutually independent beta variables. Consider the factorization of $A_i$ into two parts

\[(4.1) \quad A_i = A_4 A_5,
\]

where $A_4 = a_{11}$, $a_{11}$ being the first element of the matrix $A = L_{11}(L_{11}(I-L_{11}))^{-1} L_{12}$. From (2.5) we find the density of the $(p-s) \times (p-s)$ matrix $A$ to be

\[(4.2) \quad f(A) = K |I - A|^{(n-q-s)/2} |A|^{(2s-p-1)/2}.
\]

Partitioning $A$ and $I - A$ as

\[(4.3) \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad I - A = \begin{pmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & I - a_{22} \end{pmatrix},
\]

where $a_{11}$ is $1 \times 1$, $a_{12}$ is $1 \times (p-s)$, $a_{21}$ is $(p-s-1) \times (p-s-1)$, the joint density of $a_{11}$, $a_{12}$ and $a_{21}$ can be written as

\[(4.4) \quad f(a_{11}, a_{12}, a_{21}) = K a_{11}^{n-p-1/2} \left| a_{21} - \frac{a_{11} a_{12}}{a_{11}} \right|^{(2s-p-1)/2}
\cdot (1-a_{11})^{(n-q-s)/2} \left| I - a_{21} - \frac{a_{11} a_{12}}{1-a_{11}} \right|^{(n-q-s)/2},
\]

Now we set

\[(4.5) \quad M = a_{21} - \frac{a_{11} a_{12}}{a_{11}} \]
and find the joint density of \( A_{11}, M \) and \( A_{21} \) to be

\[
f(A_{11}, A_{21}, M) = K A_{21}^{2s-p-1/2} |M|^{(2s-p-1)/2} \cdot (1 - A_{11})^{(n-q-s)/2} \left| I - M - \frac{A_{21} A_{12}}{A_{11}(1 - A_{11})} \right|^{(n-q-s)/2}
\]

substitute

\[ A_{21} = A_{11} (1 - A_{11})^{1/2} (I - M)^{1/2} \delta, \]

we obtain the joint density of \( A_{11}, M \) and \( \delta \) as

\[
f(A_{11}, M, \delta) = K A_{11}^{2s-p-1/2} A_{12}^{p-1/2} \cdot (1 - A_{11})^{(n-q-s)/2} (1 - A_{11})^{(p-s-1)/2} \left| M \right|^{(2s-p-1)/2} \left| I - M \right|^{(n-q-s+1)/2} \left| I - \delta \delta' \right|^{(n-q-s)/2}
\]

from (4.7) we see that the densities of \( A_1 = A_{11} \) and \( A_2 = |M| \) are independent. We also note that \( A_1 | M = A_2 A_4 = A_1 \). Radcliffe derives the distribution of \( A_4 \) and \( A_1 \) for the particular case \( s=2 \). We also proceed to obtain the results for \( s=2 \). In this case we proceed as follows. From equation (2.4) the density of \( L_{11} \) and \( L_{12} \), for \( s=2 \), is

\[
f(L_{11}, L_{12}) = \theta(L_{11}) \left| L_{11} \right|^{(n-q-p-1)/2} \left| I - L_{11} \right|^{(q-p-1)/2} \left| L_{12} \right|^{(n-p-3)/2} \left| I - L_{12} - L_{11} \right|^{(n-p-3)/2} \left| I - L_{11} \right|^{(n-p-3)/2}.
\]

Let \( L_{12} L_{21} = V \), using Hsu’s lemma, the joint density of \( L_{11} \) and \( V \) is

\[
f(L_{11}, V) = K \left| L_{11} \right|^{(n-q-p-1)/2} \left| I - L_{11} \right|^{(q-p-1)/2} \theta(L_{11}) \left| L_{11} \right|^{(n-p-3)/2} \left| I - L_{11} - V \right|^{(n-p-3)/2} \left| V \right|^{(p-3)/2}.
\]

Further setting

\[
\left\{ \begin{align*}
L_{11}(I - L_{11}) - V &= R \\
L_{11}(I - L_{11}) &= U U^{T}
\end{align*} \right.
\]

where \( U \) is a lower triangular matrix and \( R = U F U^{T} \) we find that the density of the matrix \( F \) is independent of \( L \) and is given by

\[
f(F') = K \left| F' \right|^{(n-p-3)/2} \left| I - F' \right|^{(p-3)/2}.
\]

We further note that \( A_1 = |F| \), \( A_2 = f_{11} \) where \( f_{11} \) is the first element of \( F \). Proceeding on similar lines as in (4.3) and (4.4) and setting \( x_{12} = f_{12} - f_{12}^{1/2} f_{12}^{1/2} \) and \( 1 - x_{12} = (1 - f_{12})^{1/2} (1 - f_{12})^{1/2} f_{12}^{1/2} x_{12} \) the joint density of \( f_{12} \), \( x_{12} \) and \( x_{12} \) can be expressed as
DISTRIBUTION OF CERTAIN FACTORS USEFUL IN DISCRIMINANT ANALYSIS 103

\[ f(f_{11}, x_{22}, x_{12}) = K f_{11}^{(n-p-1)/2} (1 - f_{11})^{(p-4)/2} x_{22}^{(n-p-3)/2} \]
\[ \cdot (1 - x_{22})^{(p-4)/2} (1 - x_{12})^{(p-5)/2}. \]

It follows from (4.12) that beta densities of \( f_{11} = \lambda_4, x_{22} = \lambda_4 \) are independent. This result agrees with the one given by Radcliffe ([7], p. 740).

DALHOUSIE UNIVERSITY, HALIFAX, NOVA SCOTIA CANADA
ST. MARY'S UNIVERSITY, HALIFAX, NOVA SCOTIA CANADA

REFERENCES