The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 942–952

ASYMPTOTIC EXPANSIONS OF THE NON-NULL DISTRIBUTIONS OF THE LIKELIHOOD RATIO CRITERIA FOR MULTIVARIATE LINEAR HYPOTHESIS AND INDEPENDENCE[1]

BY NARIAKI SUGIURA AND YASUNORI FUJIKOSHI

University of North Carolina and Hiroshima University

0. Summary. Asymptotic non-null distribution of the likelihood ratio criterion for testing the linear hypothesis in multivariate analysis is obtained up to the order N^{-2}, where N means the sample size, by using the characteristic function expressed in terms of hypergeometric function with matrix argument. This result holds without any assumption on the rank of non-centrality matrix. Asymptotic non-null distribution of the likelihood ratio criterion for independence between two sets of variates is also obtained up to the order N^{-1}.

1. Expansion of the criterion for multivariate linear hypothesis.

1.1. *Introduction.* Let each column vector of $p \times N$ matrix $X = (X_1, X_2, \cdots, X_N)$ be distributed independently according to a p-variate normal distribution with common covariance matrix Σ. Then the canonical form of the multivariate linear hypothesis is defined by testing the hypothesis

$$H: E[X_j] = 0 \tag{1.1}$$

for $j = 1, 2, \cdots, b$ and $s + 1, \cdots, N$ with $b \leqq s$ against alternatives

$$\begin{aligned} K: E[X_j] &\neq 0 \qquad \text{for some } j \, (1 \leqq j \leqq b) \\ &= 0 \qquad \text{for } j = s + 1, \cdots, N. \end{aligned} \tag{1.2}$$

The likelihood ratio test for this problem is based on the statistic

$$\lambda = (|S_e|/|S_e + S_h|)^{N/2}, \tag{1.3}$$

where $S_e = \sum_{\alpha=s+1}^{N} X_\alpha X_\alpha'$ and $S_h = \sum_{\alpha=1}^{b} X_\alpha X_\alpha'$. The matrix S_e is called the sum of squares and products due to error and has the Wishart distribution $W_p(N - s, \Sigma)$. The matrix S_h is called the sum of squares and products due to departure from the hypothesis and has the noncentral Wishart distribution $W_p(b, \Sigma; \Omega)$, where the non-centrality matrix is $\Omega = \frac{1}{2}\Lambda\Lambda'\Sigma^{-1}$ with $\Lambda = E[X_1, X_2, \cdots, X_b]$ under alternative K.

In this paper, we shall derive the asymptotic expansion of the non-null distribution $P(-2\rho \log \lambda > z)$ up to the order N^{-2}, where a correction factor ρ is determined such that under H, the first remainder term in the asymptotic expansion of $P(-2\rho \log \lambda > z)$ vanishes, that is, ρ is given by $\rho N = N - s + (b - p - 1)/2$ (Anderson [1], p. 208). The non-null distribution depends only on the non-centrality matrix Ω. Posten and Bargmann [10] obtained the same asymptotic

Received 6 February 1968; revised 16 December 1968.
[1] This research was supported by the National Science Foundation Grant No. GU-2059 and the Sakko-kai Foundation.

expansion when the non-centrality matrix Ω is of rank two. Our result holds for arbitrary non-centrality matrix. We shall state some necessary results for hypergeometric function with matrix argument due to Constantine [4] and then prove a lemma for zonal polynomials, which is fundamental for our asymptotic expansion.

1.2. *Preliminaries.* The following two hypergeometric functions are used in this paper.

$$(1.4)\qquad {}_1F_1\,(a; b; Z) = \sum_{k=0}^{\infty} \sum_{(\kappa)} (a)_\kappa (b)_\kappa^{-1} C_\kappa(Z)(k!)^{-1},$$

$${}_2F_1\,(a_1, a_2; b; Z) = \sum_{k=0}^{\infty} \sum_{(\kappa)} (a_1)_\kappa (a_2)_\kappa (b)_\kappa^{-1} C_\kappa(Z)(k!)^{-1},$$

where

$$(a)_\kappa = \prod_{\alpha=1}^{p} (a - (\alpha-1)/2)(a + 1 - (\alpha - 1)/2) \cdots (a + k_\alpha - 1 - (\alpha - 1)/2)$$

and the function $C_\kappa(Z)$ is a zonal polynomial of the $p \times p$ symmetric matrix Z corresponding to the partition $\kappa = \{k_1, k_2, \cdots, k_p\}$ with $k_1 + k_2 + \cdots + k_p = k$ and $k_1 \geqq k_2 \geqq \cdots \geqq k_p \geqq 0$. The symbol $\sum_{(\kappa)}$ means the sum of all such partitions for fixed k. $C_\kappa(Z)$ is a kth degree homogeneous symmetric polynomial of the p characteristic roots of Z. The following formulae are established by Constantine [4].

$$(1.5)\qquad |I - Z|^{-a} = \sum_{k=0}^{\infty} \sum_{(\kappa)} (a)_\kappa C_\kappa(Z)/k!,$$

$$(1.6)\qquad {}_0F_0\,(Z) = \sum_{k=0}^{\infty} \sum_{(\kappa)} C_\kappa(Z)/k! = \operatorname{etr} Z,$$

$$(1.7)\qquad \int_{S>0} \{\operatorname{etr}\,(-Z^{-1}S)\} |S|^{t-(p+1)/2} C_\kappa(ST)\, dS = \Gamma_p(t)(t)_\kappa |Z|^t C_\kappa(TZ),$$

where the first formula (1.5) holds when all characteristic roots of Z are less than one and the last formula (1.7) holds for any $p \times p$ positive definite matrices Z, S and any $p \times p$ symmetric matrix T with $t > (p-1)/2$. The function $\Gamma_p(t)$ is defined by

$$(1.8)\qquad \Gamma_p(t) = \pi^{p(p-1)/4} \prod_{\alpha=1}^{p} \Gamma(t - (\alpha - 1)/2).$$

We shall also use the following asymptotic formula for the gamma function shown by Box [3] (Anderson [1], p. 204).

$$(1.9)\qquad \log \Gamma(x + h) = \log\,(2\pi)^{\frac{1}{2}} + (x + h - \tfrac{1}{2}) \log x - x$$
$$- \sum_{r=1}^{m} (-1)^r B_{r+1}(h)(r(r+1)x^r)^{-1} + O(|x|^{-m-1}),$$

which holds for large $|x|$ and fixed h with Bernoulli polynomial $B_r(h)$ of degree r. For $r = 2$ and 3,

$$(1.10)\qquad B_2(h) = h^2 - h + (\tfrac{1}{6}), \qquad B_3(h) = h^3 - (\tfrac{3}{2})h^2 + \tfrac{1}{2}h.$$

1.3. *A lemma.* In (1.6) replacing Z by xZ, we have

$$(1.11)\qquad \exp\,(x \operatorname{tr} Z) = \sum_{k=0}^{\infty} \sum_{(\kappa)} x^k C_\kappa(Z)/k!$$

and differentiating with respect to x yields

$$\sum_{k=1}^{\infty}\sum_{(\kappa)} C_\kappa(Z)/(k-1)! = (\text{tr } Z) \text{ etr } Z, \tag{1.12}$$
$$\sum_{k=2}^{\infty}\sum_{(\kappa)} C_\kappa(Z)/(k-2)! = (\text{tr } Z)^2 \text{ etr } Z,$$

which was used by Fujikoshi [5] in deriving the asymptotic expansion of the distribution of the generalized variance under the noncentral case. We now prove the following lemma which is fundamental for our asymptotic expansion.

LEMMA. *Let $C_\kappa(Z)$ be a zonal polynomial corresponding to the partition $\kappa = \{k_1, k_2, \cdots, k_p\}$ with $k_1 + k_2 + \cdots + k_p = k$ and $k_1 \geqq k_2 \geqq \cdots \geqq k_p \geqq 0$. Putting*

$$a_1(\kappa) = \sum_{\alpha=1}^{p} k_\alpha(k_\alpha - \alpha), \qquad a_2(\kappa) = \sum_{\alpha=1}^{p} k_\alpha(4k_\alpha^2 - 6\alpha k_\alpha + 3\alpha^2) \tag{1.13}$$

then the following equalities hold.

$$\sum_{k=0}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z) a_1(\kappa)/k! = (x^2 \text{ tr } Z^2) \text{ etr } (xZ), \tag{1.14}$$

$$\sum_{k=1}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z) a_1(\kappa)/(k-1)! = (2x^2 \text{ tr } Z^2 + x^3 \text{ tr } Z^2 \text{ tr } Z) \text{ etr } (xZ), \tag{1.15}$$

$$\sum_{k=0}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z) a_1(\kappa)^2/k! = \{x^4(\text{tr } Z^2)^2 + 4x^3 \text{ tr } Z^3 + x^2 \text{ tr } Z^2 + x^2(\text{tr } Z)^2\} \text{ etr } (xZ), \tag{1.16}$$

$$\sum_{k=0}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z) a_2(\kappa)/k! = \{4x^3 \text{ tr } Z^3 + 3X^2 \text{ tr } Z^2 + 3x^2(\text{tr } Z)^2 + x \text{ tr } Z\} \text{ etr } (xZ). \tag{1.17}$$

PROOF. From (1.5) we can write

$$\begin{aligned} |I - n^{-1}Z|^{-nx} &= \sum_{k=0}^{\infty}\sum_{(\kappa)} (nx)_\kappa C_\kappa(Z)/n^k k! \\ &= \sum_{k=0}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z)(k!)^{-1}[1 + a_1(\kappa)(2nx)^{-1} \\ &\quad + (24n^2x^2)^{-1}\{k - a_2(\kappa) + 3a_1(\kappa)^2\} + O(n^{-3})], \end{aligned} \tag{1.18}$$

which holds for any number x and any positive definite matrix Z with large n. The left-hand side can be expanded asymptotically in another way by the formula

$$-\log|I - n^{-1}Z| = \text{tr } (Z/n) + \tfrac{1}{2}\text{tr } (Z/n)^2 + \tfrac{1}{3}\text{tr } (Z/n)^3 + O(n^{-4}), \tag{1.19}$$

as

$$\begin{aligned} |I - n^{-1}Z|^{-nx} &= \exp\{-nx \log|I - n^{-1}Z|\} \\ &= \{\text{etr } (xZ)\}[1 + (x/2n) \text{ tr } Z^2 \\ &\quad + (x/24n^2)\{3x(\text{tr } Z^2)^2 + 8\text{tr } Z^3\} + O(n^{-3})]. \end{aligned} \tag{1.20}$$

Comparing the coefficients of each term of the order n^{-1} in (1.18) and (1.20), we can obtain the first equality (1.14). Equality (1.15) follows immediately by differentiating (1.14) with respect to x. From the terms of order n^{-2} in (1.18) and (1.20) with the help of the formula (1.12), we have

$$(1.21)\quad \sum_{k=0}^{\infty}\sum_{(\kappa)} x^k C_\kappa(Z)(k!)^{-1}\{3a_1(\kappa)^2 - a_2(\kappa)\} = \{3x^4(\operatorname{tr} Z^2)^2 + 8x^3 \operatorname{tr} Z^3 - x \operatorname{tr} Z\} \operatorname{etr}(xZ).$$

By (1.7) we can also get the following asymptotic formula for any $p \times p$ positive definite matrix Z and large n;

$$(1.22)\quad \Gamma_p(\tfrac{1}{2}n)|Z|^{\frac{1}{2}n}\}^{-1} \int_{S>0} \{\operatorname{etr}(-Z^{-1}S)\}|S|^{\frac{1}{2}(n-p-1)} C_\kappa(2S/n)\, dS = (2/n)^k(\tfrac{1}{2}n)_\kappa C_\kappa(Z) = \{1 + a_1(\kappa)n^{-1} + O(n^{-2})\}C_\kappa(Z).$$

Multiplying on both sides of (1.22) by $x^k a_1(\kappa)/k!$ and using the first formula (1.14), we can get

$$(1.23)\quad \{\Gamma_p(\tfrac{1}{2}n)|Z|^{\frac{1}{2}n}\}^{-1} \int_{S>0} \{\operatorname{etr}(-Z^{-1}S)\}|S|^{\frac{1}{2}(n-p-1)} \operatorname{tr}(2xS/n)^2 \operatorname{etr}(2xS/n)\, dS = \sum_{k=0}^{\infty}\sum_{(\kappa)} \{x^k C_\kappa(Z)/k!\}\{a_1(\kappa) + a_1(\kappa)^2/n + O(n^{-2})\}.$$

The left-hand side can be regarded as the expectation of the statistic $f(V) = \operatorname{tr}(xV)^2 \operatorname{etr} xV$ with respect to the Wishart distribution $W_p(n, Z)$ on nV. We shall expand it asymptotically in another way by using a matrix of differential operators due to James [8], Ito [7], Siotani [12] and etc. Considering the transformation $V \to HVH'$ for some orthogonal matrix H, we may assume that positive definite matrix Z is a diagonal matrix $\Gamma = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$ with the p characteristic roots of Z as its nonzero elements. Then the statistic V converges to Γ in probability as n tends to infinity. So we shall expand the function $f(V)$ in a Taylor series about Γ as $\{\operatorname{etr}(V-\Gamma)\partial\}f(\Sigma)|_{\Sigma=\Gamma}$, where symbol ∂ means the matrix of differential operators having $\frac{1}{2}(1+\delta_{ij})(\partial/\partial\sigma_{ij})$ as its (i,j)th element for a symmetric matrix $\sum = (\sigma_{ij})$ with $\delta_{ij} = 0$ $(i \neq j)$ and $\delta_{ii} = 1$. Taking the expectation with respect to V regarding the matrix ∂ as a constant, we can rewrite the left-hand side of (1.23) as

$$\begin{aligned}
&[\operatorname{etr}\{-\Gamma\partial - \tfrac{1}{2}n \log|I - (2/n)\Gamma\partial|\}]f(\Sigma)|_{\Sigma=\Gamma}\\
&= \{1 + n^{-1} \operatorname{tr}(\Gamma\partial)^2 + O(n^{-2})\} \operatorname{tr}(x\Sigma)^2\cdot\operatorname{etr}(x\Sigma)|_{\Sigma=\Gamma}\\
&= (x^2 \operatorname{tr}\Gamma^2)\operatorname{etr}(x\Gamma)\\
(1.24)\quad &\quad + n^{-1}\{\textstyle\sum_{\alpha=1}^{p}\lambda_\alpha^2\,\partial^2/\partial\sigma_{\alpha\alpha}^2 + \tfrac{1}{2}\sum_{1\leq\alpha<\beta\leq p}\lambda_\alpha\lambda_\beta\,\partial^2/\partial\sigma_{\alpha\beta}^2\}\\
&\quad\cdot\{x^2(\textstyle\sum_{\alpha=1}^{p}\sigma_{\alpha\alpha}^2 + 2\sum_{1\leq\alpha<\beta\leq p}\sigma_{\alpha\beta}^2)\operatorname{etr}(x\Sigma)\}|_{\Sigma=\Gamma} + O(n^{-2})\\
&= (x^2 \operatorname{tr} Z^2)\operatorname{etr}(xZ)\\
&\quad + n^{-1}\{x^4(\operatorname{tr} Z^2)^2 + 4x^3 \operatorname{tr} Z^3 + x^2 \operatorname{tr} Z^2 + x^2(\operatorname{tr} Z)^2\}\operatorname{etr}(xZ)\\
&\quad + O(n^{-2}).
\end{aligned}$$

Comparing the coefficients of each term of order n^{-1} in the last equation of (1.23) and (1.24), we can see that the third formula (1.16) is true. The fourth formula (1.17) is an immediate consequence from the third formula (1.16) and the equality (1.21). Thus the lemma is proved.

1.4. *Approximate non-null distribution.* Now we shall derive the asymptotic expansion of the non-null distribution of the likelihood ratio criterion λ given by (1.3). Constantine [4] showed that the hth moment of the ratio of determinants $|S_e|/|S_e + S_h|$ under K could be expressed by our notation as

$$
\begin{aligned}
& E[|S_e|^h/|S_e + S_h|^h] \\
(1.25) \qquad & = \Gamma_p(h + (N - s)/2)\Gamma_p((N - s + b)/2) \\
& \quad \cdot [\Gamma_p((N - s)/2)\Gamma_p(h + (N - s + b)/2)]^{-1} \\
& \quad \cdot {}_1F_1(h; h + (N - s + b)/2; -\Omega).
\end{aligned}
$$

Put $m = \rho N = N - s + (b - p - 1)/2$ and let m tend to infinity instead of N as in Posten and Bargmann [10]. We can express the characteristic function of $-2\rho \log \lambda$ from (1.25) as

$$
\begin{aligned}
C(t) & = [\Gamma_p(\tfrac{1}{2}m(1 - 2it) - (b - p - 1)/4)\Gamma_p(\tfrac{1}{2}m + (b + p + 1)/4)] \\
(1.26) \qquad & \quad \cdot [\Gamma_p(\tfrac{1}{2}m - (b - p - 1)/4)\Gamma_p(\tfrac{1}{2}m(1 - 2it) + (b + p + 1)/4)]^{-1} \\
& \quad \cdot {}_1F_1(-itm; \tfrac{1}{2}m(1 - 2it) + (b + p + 1)/4; -\Omega) \\
& = C_1(t) \cdot C_2(t).
\end{aligned}
$$

Under the hypothesis H, the non-centrality matrix Ω is equal to zero matrix and the hypergeometric function ${}_1F_1$ is equal to unity. So the first four gamma products give us the characteristic function under H, which we shall denote by $C_1(t)$ and ${}_1F_1$ by $C_2(t)$. The first part $C_1(t)$ can be expanded for large m in the usual manner as shown by Box [3] (Anderson [1], p. 204). Applying the asymptotic formula for gamma function (1.9) to $C_1(t)$, we have

$$
\begin{aligned}
\log C_1(t) = & -\tfrac{1}{2}bp \log (1 - 2it) \\
& + m^{-1} \textstyle\sum_{\alpha=1}^{p} \{B_2((b + p + 3 - 2\alpha)/4) - B_2((-b + p + 3 - 2\alpha)/4)\} \\
& \cdot \{1 - (1 - 2it)^{-1}\} - 2(3m^2)^{-1} \\
& \cdot \textstyle\sum_{\alpha=1}^{p} \{B_3((b + p + 3 - 2\alpha)/4) - B_3((-b + p + 3 - 2\alpha)/4)\} \\
& \cdot \{1 - (1 - 2it)^{-2}\} + O(m^{-3}).
\end{aligned}
$$

The second term of the above expression vanishes, giving

$$
\begin{aligned}
(1.27) \quad C_1(t) = & (1 - 2it)^{-\frac{1}{2}bp} \\
& \cdot [1 + bp(48m^2)^{-1}(b^2 + p^2 - 5)\{(1 - 2it)^{-2} - 1\} + O(m^{-3})].
\end{aligned}
$$

The second factor of $C(t)$ in (1.26) can be written, by the definition (1.4), as

$$
\begin{aligned}
(1.28) \quad C_2(t) = & \textstyle\sum_{k=0}^{\infty} \sum_{(\kappa)} (-imt)_\kappa C_\kappa(-\Omega) \\
& \cdot [(\tfrac{1}{2}m(1 - 2it) + \tfrac{1}{4}(b + p + 1))_\kappa k!]^{-1}.
\end{aligned}
$$

The coefficients can be arranged according to the order of the power of m as

$$
\begin{aligned}
(-itm)_\kappa &= (-itm)^k[1 - a_1(\kappa)(2itm)^{-1} \\
&\quad + (24(itm)^2)^{-1}\{k - a_2(\kappa) + 3a_1(\kappa)^2\} + O(m^{-3})], \\
(\tfrac{1}{2}m(1 &- 2it) + \tfrac{1}{4}(b + p + 1))_\kappa \\
&= \{\tfrac{1}{2}m(1 - 2it)\}^k[1 + ((b + p + 1)k + 2a_1(\kappa)) \\
&\quad \cdot (2m(1 - 2it))^{-1} + (24m^2(1 - 2it)^2)^{-1} \\
&\quad \cdot \{4k + 3(b + p + 1)^2k(k - 1) \\
&\quad + 12(b + p + 1)(k - 1)a_1(\kappa) - 4a_2(\kappa) + 12a_1(\kappa)^2\} \\
&\quad + O(m^{-3})],
\end{aligned}
\tag{1.29}
$$

where $a_1(\kappa)$ and $a_2(\kappa)$ are defined by (1.13). Hence we can write $C_2(t)$ as

$$
\begin{aligned}
&\textstyle\sum_{k=0}^{\infty} \sum_{(\kappa)} (2it/(1 - 2it))^k C_\kappa(\Omega)(k!)^{-1} \\
&\quad \cdot [1 - m^{-1}\{(b + p + 1)k(2(1 - 2it))^{-1} + a_1(\kappa)(2it(1 - 2it))^{-1}\} \\
&+ m^{-2}\{(1 - 4it)k(24(it)^2(1 - 2it)^2)^{-1} \\
&+ (b + p + 1)^2(8(1 - 2it)^2)^{-1}k(k + 1) \\
&+ (b + p + 1)a_1(\kappa)(2(1 - 2it)^2)^{-1}(1 + k/2it) \\
&- (1 - 4it)a_2(\kappa)(24(it)^2(1 - 2it)^2)^{-1} \\
&+ a_1(\kappa)^2(8(it)^2(1 - 2it)^2)^{-1}\} + O(m^{-3})].
\end{aligned}
\tag{1.30}
$$

By the lemma in Section 1.3, with the formulae (1.12), we can simplify the above expression, getting

$$
\begin{aligned}
C_2(t) &= \{\exp(2it(1 - 2it)^{-1}\operatorname{tr}\Omega)\}[1 - m^{-1}\{(b + p + 1)it(1 - 2it)^{-2}\operatorname{tr}\Omega \\
&\quad + 2it(1 - 2it)^{-3}\operatorname{tr}\Omega^2\} \\
&\quad + m^{-2}\{(b + p + 1)^2 it(2(1 - 2it)^3)^{-1}\operatorname{tr}\Omega \\
&\quad + 2[b + p + 2 + (b + p + 1)it]it(1 - 2it)^{-4}\operatorname{tr}\Omega^2 \\
&\quad + [4 + (b + p + 1)^2 it]it(2(1 - 2it)^4)^{-1}(\operatorname{tr}\Omega)^2 \\
&\quad + 8(1 + 2it)it(3(1 - 2it)^5)^{-1}\operatorname{tr}\Omega^3 \\
&\quad + 2(b + p + 1)(it)^2(1 - 2it)^{-5}\operatorname{tr}\Omega^2\operatorname{tr}\Omega \\
&\quad + 2(it)^2(1 - 2it)^{-6}(\operatorname{tr}\Omega^2)^2\} + O(m^{-3})].
\end{aligned}
\tag{1.31}
$$

Combining this result with the expression for $C_1(t)$ in (1.27), we finally obtain the following asymptotic formula for the characteristic function $C(t)$;

$$
\begin{aligned}
C(t) = {} & (1-2it)^{-bp/2}\exp\left(2it(1-2it)^{-1}\operatorname{tr}\Omega\right) \\
& \cdot[1+(2m)^{-1}\{(b+p+1)(1-2it)^{-1}\operatorname{tr}\Omega \\
& -(1-2it)^{-2}[(b+p+1)\operatorname{tr}\Omega-2\operatorname{tr}\Omega^2]-2\,tr\,\Omega^2(1-2it)^{-3}\} \\
& +m^{-2}\{\tfrac{1}{48}bp(b^2+p^2-5)[(1-2it)^{-2}-1] \\
& +\textstyle\sum_{\alpha=2}^{6} g_{2\alpha}(\Omega)(1-2it)^{-\alpha}\}+O(m^{-3})],
\end{aligned}
\tag{1.32}
$$

where each coefficient $g_{2\alpha}(\Omega)$ is given by

$$
\begin{aligned}
g_4(\Omega) = {} & \tfrac{1}{8}(b+p+1)^2\{(\operatorname{tr}\Omega)^2-2\operatorname{tr}\Omega\}+\tfrac{1}{2}(b+p+1)\operatorname{tr}\Omega^2, \\
g_6(\Omega) = {} & \tfrac{1}{4}(b+p+1)^2\operatorname{tr}\Omega-\{1+2(b+p+1)\}\operatorname{tr}\Omega^2 \\
& -\{1+\tfrac{1}{4}(b+p+1)^2\}(\operatorname{tr}\Omega)^2 \\
& +\tfrac{4}{3}\operatorname{tr}\Omega^3+\tfrac{1}{2}(b+p+1)(\operatorname{tr}\Omega)\operatorname{tr}\Omega^2, \\
g_8(\Omega) = {} & \{1+\tfrac{3}{2}(b+p+1)\}\operatorname{tr}\Omega^2+\{1+\tfrac{1}{8}(b+\mathrm{p}+1)^2\}(\operatorname{tr}\Omega)^2 \\
& -4\operatorname{tr}\Omega^3-(b+p+1)(\operatorname{tr}\Omega)\operatorname{tr}\Omega^2+\tfrac{1}{2}(\operatorname{tr}\Omega^2)^2, \\
g_{10}(\Omega) = {} & \tfrac{8}{3}\operatorname{tr}\Omega^3+\tfrac{1}{2}(b+p+1)(\operatorname{tr}\Omega)\operatorname{tr}\Omega^2-(\operatorname{tr}\Omega^2)^2, \\
g_{12}(\Omega) = {} & \tfrac{1}{2}(\operatorname{tr}\Omega^2)^2.
\end{aligned}
\tag{1.33}
$$

By inverting this characteristic function using the well-known fact that $(1-2it)^{-f/2}\exp\{2it\delta^2/(1-2it)\}$ is the characteristic function of the non-central χ^2 distribution with f degrees of freedom and non-centrality parameter δ^2, we obtain the following theorem.

THEOREM 1.1. *The non-null distribution of the likelihood ratio criterion* (1.3) *for the multivariate linear hypothesis can be approximated asymptotically up to the order* m^{-2} *by*

$$
\begin{aligned}
& P(-2\rho\log\lambda<z) \\
& = P(\chi_f^2(\delta^2)<z)+(2m)^{-1}[(b+p+1)\operatorname{tr}\Omega\cdot P(\chi_{f+2}^2(\delta^2)<z) \\
& \quad -\{(b+p+1)\operatorname{tr}\Omega-2\operatorname{tr}\Omega^2\}P(\chi_{f+4}^2(\delta^2)<z) \\
& \quad -2\operatorname{tr}\Omega^2\cdot P(\chi_{f+6}^2(\delta^2)<z)] \\
& \quad +m^{-2}[\tfrac{1}{48}bp(b^2+p^2-5)\{P(\chi_{f+4}^2(\delta^2)<z)-P(\chi_f^2(\delta^2)<z)\} \\
& \quad +\textstyle\sum_{\alpha=2}^{6} g_{2\alpha}(\Omega)\cdot P(\chi_{f+2\alpha}^2(\delta^2)<z)]+O(m^{-3}),
\end{aligned}
\tag{1.34}
$$

where $m=\rho N=N-s+(b-p-1)/2, f=bp, \delta^2=\operatorname{tr}\Omega=\frac{1}{2}\operatorname{tr}\Lambda\Lambda'\Sigma^{-1}$ and the coefficients $g_{2\alpha}(\Omega)$ $(\alpha=2,3,\cdots,6)$ are given by (1.33). The symbol $\chi_f^2(\delta^2)$ means the non-central χ^2-variate with f degrees of freedom and non-centrality parameter δ^2.

If we specialize the rank of Ω to two, we can easily see the agreement between our result and that of Posten and Bargmann [10], after minor changes of notation.

2. Expansion of the criterion for independence.

2.1. *Test criterion for independence.* There is a close connection between the multivariate linear hypothesis and the test for independence between two sets of variates. In fact we can reduce the test for independence to that of the linear hypothesis by considering the conditional distribution of one set of variates, given another set. A monotonicity property of the power function of the test criteria was proved in this way by Anderson and Das Gupta [2]. Thus we can expect that the same may be true in regard to asymptotic expansion.

Let $p \times 1$ vector $X_1, X_2, \cdots, X_N$ be a random sample from a multivariate normal distribution with mean vector μ and covariance matrix Σ. Put $S = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})'$, $\bar{X} = N^{-1} \sum_{\alpha=1}^{N} X_\alpha$ and let us partition Σ and S into p_1 and p_2 rows and columns $(p_1 + p_2 = p)$ as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \qquad S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}.$$

Without loss of generality we may assume $p_1 \leqq p_2$. The likelihood ratio test for the hypothesis of independence $H:\Sigma_{12} = 0$ $(p_1 \times p_2)$ against all alternatives $K:\Sigma_{12} \neq 0$ is given by

$$(2.1) \qquad \lambda = (|S|/|S_{11}|\cdot|S_{12}|)^{N/2} = |I - S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}|^{N/2}.$$

This can be also expressed by $\prod_{j=1}^{p_1} (1 - r_j^2)^{\frac{1}{2}N}$, using the sample canonical correlations r_j where r_j^2 are given by the characteristic root of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ for $j = 1, 2, \cdots, p_1$.

2.2. *Moments of the criterion.* First we shall obtain the moments of the likelihood ratio statistic λ under K in a convenient form for our asymptotic expansion.

THEOREM 2.1. *Under alternative K, the moment of the likelihood ratio statistic λ given by* (2.1) *can be expressed as*

$$\begin{aligned} &E[|S|^h/(|S_{11}|\,|S_{22}|)^h] \\ (2.2) \qquad &= \Gamma_{p_1}(h + \tfrac{1}{2}(N - p_2 - 1))\Gamma_{p_1}(\tfrac{1}{2}(N-1)) \\ &\quad \cdot [\Gamma_{p_1}(\tfrac{1}{2}(N - p_2 - 1))\Gamma_{p_1}(h + \tfrac{1}{2}(N-1))]^{-1} \prod_{j=1}^{p_1} (1 - \rho_j^2)^h \\ &\quad \cdot {}_2F_1(h, h; h + \tfrac{1}{2}(N-1); P^2), \end{aligned}$$

where ρ_j is the population canonical correlation and $P^2 = \text{diag}(\rho_1^2, \rho_2^2, \cdots, \rho_{p_1}^2)$.

PROOF. Considering the conditional distribution of the first p_1 components of the sample, given the p_2 second components $Y(p_2 \times (N-1))$ in the canonical set up of Σ, Constantine [4] showed that the statistic $|S|/(|S_{11}|\,|S_{22}|)$ is expressed by $|VV'|/|UU' + VV'|$, where the $p_1 \times p_1$ matrix VV' has the Wishart distribution $W_{p_1}(N - 1 - p_2, \Gamma)$ and the $p_1 \times p_1$ matrix UU' has the non-central

Wishart distribution $W_{p_1}(p_2, \Gamma; \frac{1}{2}\Gamma^{-1}\Lambda YY'\Lambda')$ with

$$\Lambda(p_1 \times p_2) = \begin{pmatrix} \rho_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \rho_2 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \rho_{p_1} & 0 & \cdots & 0 \end{pmatrix} \quad \text{and}$$

$$(2.3) \qquad \Gamma(p_1 \times p_1) = \begin{pmatrix} 1-\rho_1^2 & 0 & \cdots & 0 \\ 0 & 1-\rho_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1-\rho_{p_1}^2 \end{pmatrix}.$$

The random matrices VV' and UU' are independent for given Y and the $p_2 \times p_2$ matrix YY' has the Wishart distribution $W_{p_2}(N-1, I)$. It follows from (1.25) that

$$(2.4) \qquad \begin{aligned} E[|S|^h/(|S_{11}|\,|S_{22}|)^h \mid YY'] &= \Gamma_{p_1}(h + \tfrac{1}{2}(N - p_2 - 1))\Gamma_{p_1}(\tfrac{1}{2}(N-1)) \\ &\quad \cdot [\Gamma_{p_1}(\tfrac{1}{2}(N - p_2 - 1))\Gamma_{p_1}(h + \tfrac{1}{2}(N-1))]^{-1} \\ &\quad \cdot {}_1F_1(h; h + \tfrac{1}{2}(N-1); -\tfrac{1}{2}\Gamma^{-1}\Lambda YY'\Lambda'). \end{aligned}$$

Applying the Kummer transformation formula ${}_1F_1(a; b; Z) = (\operatorname{etr} Z)\, {}_1F_1(b-a; b; -Z)$ (Herz [6]) to the second factor and taking expectation with respect to YY' by the Wishart distribution $W_{p_2}(N-1, I)$ with the formula (1.7), we can obtain the second factor as

$$(2.5) \qquad \begin{aligned} &\textstyle\sum_{k=0}^{\infty} \sum_{(\kappa)} ((N-1)/2)_\kappa ((N-1)/2)_\kappa \\ &\quad \cdot [(h + \tfrac{1}{2}(N-1))_\kappa |I + \Lambda'\Gamma^{-1}\Lambda|^{(N-1)/2}]^{-1} C_\kappa(P^2)(k!)^{-1}. \end{aligned}$$

Applying again the Kummer transformation formula ${}_2F_1(a_1, a_2; b; Z) = |I - Z|^{b-a_1-a_2}\, {}_2F_1(b-a_1, b-a_2; b; Z)$ due to Herz [6] to the above expression and using the identity $|I + \Lambda'\Gamma^{-1}\Lambda| = |I - P^2|^{-1}$, we can get the moment (2.2).

2.3. *Approximate non-null distribution.* Now we shall derive the asymptotic expansion of the non-null distribution of the likelihood ratio statistic $-2 \log \lambda$ defined by (2.1). Olkin and Siotani [9] have shown that the limiting distribution of $N^{\frac{1}{2}}\{|S|/(|S_{11}|\cdot|S_{22}|) - |\Sigma|/(|\Sigma_{11}|\cdot|\Sigma_{22}|)\}$ is normal with mean zero and variance $4\{|\Sigma|/(|\Sigma_{11}|\cdot|\Sigma_{22}|)\}^2 \cdot \operatorname{tr} \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, so the statistic

$$-2N^{-\frac{1}{2}}\{\log \lambda - \log \textstyle\prod_{j=1}^{p_1} (1 - \rho_j^2)^{N/2}\}$$

has the same limiting distribution with mean zero and variance $4 \sum_{j=1}^{p_1} \rho_j^2$. It may be remarked that this limiting distribution is different from that for the linear hypothesis in Theorem 1.1. Under the hypothesis the test statistic $-2\rho \log \lambda$ is recommended instead of $-2 \log \lambda$, where the correction factor ρ is so chosen that the first remainder term vanishes in the asymptotic expansion of the distribution under H. We have $\rho N = N - (\frac{3}{2}) - (p_1 + p_2)/2$ (Anderson, [1], p. 239).

We now put $m = \rho N$ and let m tend to infinity instead of N. Then the characteristic function of the statistic $-2\rho m^{-\frac{1}{2}}\{\log\lambda - \log\prod_{j=1}^{p_1}(1-\rho_j^2)^{N/2}\}$ is obtained by putting $h = -itm^{\frac{1}{2}}$ in (2.2).

$$
\begin{aligned}
C(t) &= \Gamma_{p_1}(\tfrac{1}{2}m - m^{\frac{1}{2}}it + \tfrac{1}{4}(p_1 - p_2 + 1))\Gamma_{p_1}(\tfrac{1}{2}m + \tfrac{1}{4}(p_1 + p_2 + 1)) \\
&\quad \cdot[\Gamma_{p_1}(\tfrac{1}{2}m + \tfrac{1}{4}(p_1 - p_2 + 1))\Gamma_{p_1}(\tfrac{1}{2}m - m^{\frac{1}{2}}it + \tfrac{1}{4}(p_1 + p_2+1))]^{-1} \qquad (2.6)\\
&\quad \cdot\sum_{k=0}^{\infty}\sum_{(\kappa)}(-m^{\frac{1}{2}}it)_\kappa(-m^{\frac{1}{2}}it)_\kappa[(\tfrac{1}{2}m - m^{\frac{1}{2}}it + \tfrac{1}{4}(p_1 + p_2 + 1))_\kappa]^{-1} \\
&\quad \cdot C_\kappa(P^2)(k!)^{-1}.
\end{aligned}
$$

Applying the asymptotic formula (1.9) for the gamma function to each of the four gamma products, we get

$$\text{(2.7)}\quad \text{First factor} = 1 + itm^{-\frac{1}{2}}p_1p_2 + m^{-1}(it)^2\{p_1p_2 + \tfrac{1}{2}(p_1p_2)^2\} + O(m^{-\frac{3}{2}}).$$

In the same way as in (1.29), we can see that

$$
\begin{aligned}
&(-m^{\frac{1}{2}}it)_\kappa \\
&\quad = (-m^{\frac{1}{2}}it)^k[1 - a_1(\kappa)(2itm^{\frac{1}{2}})^{-1} + (24(it)^2m)^{-1}\{k - a_2(\kappa) + 3a_1(\kappa)^2\} \\
&\qquad + O(m^{-\frac{3}{2}})] \qquad (2.8)\\
&(\tfrac{1}{2}m - m^{\frac{1}{2}}it + \tfrac{1}{4}(p_1 + p_2 + 1))_\kappa \\
&\quad = (\tfrac{1}{2}m)^k[1 - 2itkm^{-\frac{1}{2}} + m^{-1}\{\tfrac{1}{2}k(p_1 + p_2 + 1) \\
&\qquad + 2k(k-1)(it)^2 + a_1(\kappa)\} + O(m^{-\frac{3}{2}})],
\end{aligned}
$$

which implies that $(-m^{\frac{1}{2}}it)_\kappa(-m^{\frac{1}{2}}it)_\kappa/(\frac{1}{2}m - m^{\frac{1}{2}}it + \frac{1}{4}(p_1 + p_2 + 1))_\kappa$ can be expressed as

$$
\begin{aligned}
&(-2t^2)^k[1 + m^{-\frac{1}{2}}\{2itk - (it)^{-1}a_1(\kappa)\} + (12m)^{-1}\{[24(it)^2 + (it)^{-2} \\
&\qquad - 6(p_1 + p_2 + 1)]k + 24(it)^2k^2 - 12(2k+1)a_1(\kappa) - (it)^{-2}a_2(\kappa) \\
&\qquad + 6(it)^{-2}a_1(\kappa)^2\} + O(m^{-\frac{3}{2}})].
\end{aligned}
$$

It follows from the lemma in Section 1.3 that the second factor of the characteristic function $C(t)$ in (2.6) can be written as

$$
\begin{aligned}
&{}_2F_1(-m^{\frac{1}{2}}it, -m^{\frac{1}{2}}it; \tfrac{1}{2}m - m^{\frac{1}{2}}it + \tfrac{1}{4}(p_1 + p_2 + 1); P^2) \\
&\quad = \{\operatorname{etr}(-2t^2P^2)\}[1 + 4m^{-\frac{1}{2}}(it)^3(\mathrm{tr}_2 - \mathrm{tr}_4) \\
&\qquad + m^{-1}\{(it)^2[\mathrm{tr}_4 - \mathrm{tr}_2^2 - (p_1 + p_2 + 1)\mathrm{tr}_2] \qquad (2.9)\\
&\qquad + (it)^4(8\,\mathrm{tr}_2 - 20\,\mathrm{tr}_4 + \tfrac{40}{3}\mathrm{tr}_6) + (it)^6(8\,\mathrm{tr}_4^2 + 8\,\mathrm{tr}_2^2 - 16\,\mathrm{tr}_4\,\mathrm{tr}_2)\} \\
&\qquad + O(m^{-\frac{3}{2}})],
\end{aligned}
$$

where the symbol tr_j is an abbreviation for $\mathrm{tr}\,P^j$. Multiplying the first factor given by (2.7) to the above expression, we have the following asymptotic formula for the characteristic function.

$$
\begin{aligned}
\text{(2.10)}\quad C(t) &= \operatorname{etr}(-2t^2P^2)\cdot[1 + m^{-\frac{1}{2}}\{p_1p_2it + 4(it)^3(\mathrm{tr}\,P^2 - \mathrm{tr}\,P^4)\} \\
&\qquad + m^{-1}\{[p_1p_2 + \tfrac{1}{2}(p_1p_2)^2](it)^2 + \textstyle\sum_{\alpha=1}^{3} l_{2\alpha}(P)(it)^{2\alpha}\} + O(m^{-\frac{3}{2}}),
\end{aligned}
$$

where the coefficients $l_{2\alpha}(P)$ $(\alpha = 1, 2, 3)$ are given by

$$\begin{aligned} l_2(P) &= \operatorname{tr} P^4 + (\operatorname{tr} P^2)^2 - (p_1 + p_2 + 1) \operatorname{tr} P^2, \\ l_4(P) &= 4p_1p_2(\operatorname{tr} P^2 - \operatorname{tr} P^4) + 8 \operatorname{tr} P^2 - 20 \operatorname{tr} P^4 + (40/3) \operatorname{tr} P^6, \\ l_6(P) &= 8(\operatorname{tr} P^4)^2 + 8(\operatorname{tr} P^2)^2 - 16 \operatorname{tr} P^4 \operatorname{tr} P^2. \end{aligned} \tag{2.11}$$

By inverting this characteristic function we can conclude the following theorem.

THEOREM 2.2. *The non-null distribution of the likelihood ratio statistic* $-2\rho \log \lambda$ *given by* (2.1) *with a correction factor* $\rho = 1 - N^{-1}(p_1 + p_2 + 3)/2$ *for testing the independence between two sets of variates with* p_1 *components and* p_2 *components* $(p_1 \leqq p_2)$, *is expanded asymptotically for large* N *in the following way. Put* $\tilde{\lambda} = -(m^{\frac{1}{2}}/\tau)\{\log |S|/(|S_{11}|\cdot|S_{22}|) - \log |\Sigma|/(|\Sigma_{11}|\cdot|\Sigma_{22}|)\}$ *for* $m = \rho N$ *and* $\tau = 2(\operatorname{tr} P^2)^{\frac{1}{2}}$, *where* $P^2 = \operatorname{diag}(\rho_1^2, \rho_2^2, \cdots, \rho_{p_1}^2)$ *with population canonical correlation* ρ_j. *Then we have*

$$\begin{aligned} P(\tilde{\lambda} < z) = \Phi(z) - m^{-\frac{1}{2}}\{ (p_1p_2/\tau)\Phi^{(1)}(z) + (4/\tau^3)\Phi^{(3)}(z)(\operatorname{tr} P^2 - \operatorname{tr} P^4)\} \\ + m^{-1}\{ (p_1p_2/\tau^2)(1 + \tfrac{1}{2}p_1p_2)\Phi^{(2)}(z) \\ + \textstyle\sum_{\alpha=1}^{3} \Phi^{(2\alpha)}(z) l_{2\alpha}(P)/\tau^{2\alpha}\} + O(m^{-\frac{3}{2}}), \end{aligned} \tag{212}$$

where $l_{2\alpha}(P)$ *are given by* (2.11) *and* $\Phi^{(r)}(z)$ *means the rth derivative of the standard normal distribution function* $\Phi(z)$.

In the case $p = 2$ the problem reduces to that of testing that the correlation coefficient vanishes. It would be of interest to compare this formula numerically with another formula due to Ruben [11].

REFERENCES

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

[2] ANDERSON, T. W. and DAS GUPTA, S. (1964). Monotonicity of the power functions of some tests of independence between two sets of variates. *Ann. Math. Statist.* **35** 206–208.

[3] BOX, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika* **36** 317–346.

[4] CONSTANTINE, A. G. (1963). Some non-central distribution problem in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

[5] FUJIKOSHI, Y. (1967). Asymptotic expansion of the generalized variance. (Reported in the Meeting of Math. Society of Japan. October).

[6] HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* **61** 474–523.

[7] ITO, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2-statistic. *Ann. Math. Statist.* **25** 40–75.

[8] JAMES, G. S. (1954). Test of linear hypothesis in univariate and multivariate analysis when the ratios of the population variances are unknown. *Biometrika* **41** 19–43.

[9] OLKIN, I. and SIOTANI, M. (1964). Asymptotic distribution of functions of a correlation matrix. Stanford Univ. Technical Report No. 6.

[10] POSTEN, H. O. and BARGMANN, R. E. (1964). Power of the likelihood-ratio test of the general linear hypothesis in multivariate analysis. *Biometrika* **51** 467–480.

[11] RUBEN, H. (1966). Some new results on the distribution of the sample correlation coefficient. *J. Roy. Statist. Soc. Ser. B* **28** 513–525.

[12] SIOTANI, M. (1957). On the distribution of the Hotelling's T^2-statistic. *Ann. Inst. Statist. Math. Tokyo* **8** 1–14.