ON THE LIKELIHOOD RATIO TEST OF A NORMAL MULTIVARIATE TESTING PROBLEM II[1]

By N. Giri

University of Montreal

0. Introduction and summary. Let the random vector $X = (X_1 \cdots X_p)'$ have a multivariate normal distribution with unknown mean ξ and unknown nonsingular covariance matrix Σ. Write $\bar{\Gamma} = \Sigma^{-1}\xi = (\Gamma_1, \Gamma_2, \Gamma_3)'$, where Γ_1, Γ_2 and Γ_3 are subvectors of $\bar{\Gamma}$ containing first q, next $p' - q$ and last $p - p'$ components of $\bar{\Gamma}$ respectively. We will consider here, the problem of testing the hypothesis $H_0 : \Gamma_3 = \Gamma_2 = 0$ against the alternative $H_1 : \Gamma_2 = 0, \Gamma_2 \neq 0$ when $p > p' > q$ and ξ, Σ are both unknown. The origin of the problem and its likelihood ratio test have been discussed in Giri (1964a, 1964b). It has also been shown there that for $p = p' > q$, the likelihood ratio test of H_0 against H_1 is uniformly most powerful invariant similar. In this paper we will prove that the likelihood ratio test of H_0 against H_1 is uniformly most powerful invariant similar for the general case, i.e. $p > p' > q$. A corollary that the likelihood ratio test is uniformly most powerful similar among the group of tests with power depending only on $\Delta(H_1)$ (defined in Section 1) will follow from this.

The problem of testing H_0 against H_1 remains invariant under the group G of $p \times p$ nonsingular matrices

$$g = \begin{pmatrix} g_{11} & 0 & 0 \\ g_{12} & g_{22} & 0 \\ g_{13} & g_{23} & g_{33} \end{pmatrix}$$

operating as $(X, \xi, \Sigma) \rightarrow (gX, g\xi, g\Sigma g')$, where g_{11}, g_{22} and g_{33} are $q \times q$, $(p' - q) \times (p' - q)$ and $(p - p') \times (p - p')$ submatrices of g respectively. We may restrict our attention to the space of minimal sufficient statistic $(\bar{X}, S)$ of (ξ, Σ). A maximal invariant in $(\bar{X}, S)$ under G is $R = (R_1, R_2, R_3)'$ and a corresponding maximal invariant in (ξ, Σ) under G is $\Delta = (\delta_1, \delta_2, \delta_3)'$ where $R_i \geqq 0$, $\delta_i \geqq 0$ are defined in Section 1. In terms of maximal invariant the above problem reduces to that of testing $H_0 : \delta_3 = \delta_2 = 0, \delta_1 > 0$ against $H_1 : \delta_3 = 0, \delta_2 > 0, \delta_1 > 0$ when ξ, Σ are both unknown.

It has been shown in Giri (1964a) that on the basis of N random observations the likelihood ratio test of H_0 against H_1 is given by reject H_0, if $Z = (1 - R_1 - R_2)/(1 - R_1) \leqq Z_0$, where the constant Z_0 is determined in such a way that the test has size α and under H_0, Z has beta-distribution with parameters $(N - p')/2, (p' - q)/2$.

In Section 1 we will find the maximal invariants R and Δ along with the distribution of R. Actually, we will first find here the maximal invariant in $(\bar{X}, S)$

Received 23 June 1964.

[1] Research supported by ONR contract No. Nonr-401(03) at Cornell University.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

under a more general group G_k (to be defined in Section 1) and its distribution. The maximal invariant R under G and its distribution will follow from this as a special case. In Section 2 we will prove the theorem that the likelihood ratio test is uniformly most powerful invariant similar.

1. The maximal invariant and its distribution. As remarked in the previous Section, we first consider the group G_k of $p \times p$ nonsingular matrices

$$g_k = \begin{pmatrix} g_{11} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{k1} & g_{k2} & \cdot & \cdots & g_{kk} \end{pmatrix}$$

where g_{ii} is $d_i \times d_i$ submatrix of g_k and $\sum_{i=1}^k d_i = p$ and find the maximal invariants in $(\bar{X}, S)$ and in (ξ, Σ) under G_k. Then we find the distribution of the maximal invariant in $(\bar{X}, S)$. The maximal invariant under G and its distribution will follow from this by taking $k = 3$, $d_1 = q$, $d_2 = p' - q$ and $d_3 = p - p'$. The reason for considering the above case is that we get a more general result at practically no extra cost.

Let $X^\alpha = (X_{\alpha 1} \cdots X_{\alpha p})'$, $\alpha = 1 \cdots N$ be N random observations on X, $N\bar{X} = \sum_{\alpha=1}^N X_\alpha$, $S = \sum_{\alpha=1}^N (X^\alpha - \bar{X})(X^\alpha - \bar{X})'$. Denote for any p-vector Y and any $p \times p$ matrix B; $Y = (Y_1 \cdots Y_k)'$, $Y_{[i]} = (Y_1 \cdots Y_i)'$ and

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1k} \\ \vdots & \vdots & \vdots \\ B_{k1} & \cdots & B_{kk} \end{pmatrix}, \qquad B_{[ii]} = \begin{pmatrix} B_{11} & \cdots & B_{1i} \\ \vdots & \vdots & \vdots \\ B_{i1} & \cdots & B_{ii} \end{pmatrix};$$

where $Y_i (i = 1 \cdots k)$ $d_i \times 1$ subvectors of Y and $B_{ii} (i = 1 \cdots k)$ and $d_i \times d_i$ submatrices of B. If a function ϕ of $(\bar{X}, S)$ is invariant under G_k in the usual fashion, then $\phi(\bar{X}, S) = \phi(g_k\bar{X}, g_k S g_k')$ for all $\bar{X}$, S and g_k. Since S is positive definite symmetric with probability 1 for all ξ, Σ, there exists an F and G_k such that $S = FF'$. Letting

$$g_k = \pm \begin{pmatrix} g_1^* & 0 & \cdot & 0 \\ 0 & g_2^* & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & g_k^* \end{pmatrix} F^{-1}$$

where g_i^* is $d_i \times d_i$ orthogonal matrix having the property that $g_i^*(F^{-1}\bar{X})_i$ is a $d_i \times 1$ subvector whose first component is C_i and the rest are zeroes, we see that ϕ is a function of $|C_i|$ or equivalently C_i^2, $i = 1, \cdots k$. Now,

$$\sum_{j=1}^i C_j^2 = N\bar{X}_{[i]}' S_{[ii]}^{-1} \bar{X}_{[i]}, \qquad (i = 1, \cdots k). \tag{1.1}$$

The k-vector $(C_1^2 \cdots C_k^2)$ is thus a maximal invariant if it is invariant under G_k which is easily seen to be the latter. We will find it more convenient to work with the equivalent statistic $R = (R_1 \cdots R_k)'$ where,

$$\sum_{j=1}^i R_j = \sum_{j=1}^i C_j^2 / (1 + \sum_{j=1}^i C_j^2) \qquad (i = 1, \cdots k). \tag{1.2}$$

It may be verified that $R_i \geqq 0$ and $\sum_{j=1}^{k} R_j \leqq 1$. A corresponding maximal invariant $\Delta = (\delta_1 \cdots \delta_k)'$ in the parametric space of (ξ, Σ) under G_k is easily seen to be given by $\sum_{j=1}^{i} \delta_j = N\xi'_{[i]}\Sigma_{[ii]}^{-1}\xi_{[i]}$, $i = (1, \cdots k)$. Here $\delta_i \geqq 0$.

We now compute $f_\Delta(r)$, the probability density function of R and $f_0(r)$, the probability density function of R when $\delta_i = 0 (i = 1, \cdots k)$. Since f_Δ depends on (ξ, Σ) only through Δ, we may put $\Sigma = I$ and take $N^{\frac{1}{2}}\xi = N^{\frac{1}{2}}P = N^{\frac{1}{2}}(P_1 \cdots P_k)$ such that $NP_iP_i' = \delta_i$. From Giri, Kiefer and Stein (1963), p. 1529, it follows that,

$$(1.3) \quad f_0(r) = \Gamma(\tfrac{1}{2}N)/[\Gamma(\tfrac{1}{2}(N-p))\textstyle\prod_{i=1}^{k}\Gamma(\tfrac{1}{2}d_i)] \cdot \prod_{i=1}^{k} r_i^{\frac{1}{2}d_i - 1}(1 - \sum_{i=1}^{k} r_i)^{\frac{1}{2}(N-p)-1}.$$

To compute $f_\Delta(r)$, we compute the ratio $f_\Delta(r)/f_0(r)$ using Stein's method (1956) based on invariant measure. It can be verified that a left invariant measure under G_k in $(N^{\frac{1}{2}}\bar{X}, S)$ is $d(N^{\frac{1}{2}}\bar{X}, S) = (\text{set } S)^{-\frac{1}{2}(p+2)}\, d(N^{\frac{1}{2}}\bar{X})\, dS$ and a left invariant Haar measure in G_k is $\mu(dg_k) = \prod_{i=1}^{k} |\det (g_{ii})|^{-\sigma_i}\, dg_k$, where $\sigma_i = \sum_{j=1}^{i} dj (i = 1, \cdots k)$ and $\sigma_0 = 0$. From (2.4), p. 186, Giri (1964a),

$$(1.4) \quad f_\Delta(r)/f_0(r) = \frac{\int p_{\xi,I}(g_k\, \bar{x}, g_k\, sg_k') \prod_{i=1}^{k} |\det g_{ii}|^{-\sigma_i} \prod_{i \leqq j} dg_{ij}}{\int p_{0,I}(g_k\, \bar{x}, g_k\, sg_k') \prod_{i=1}^{k} |\det g_{ii}|^{-\sigma_i} \prod_{i \leqq j} dg_{ij}}$$

where, $p_{\xi,I}(\bar{x}, s) = C(\det s)^{\frac{1}{2}N} \exp[-\frac{1}{2}\operatorname{tr}(s + N(\bar{x} - \xi)(\bar{x} - \xi)')$ with $C = N^{\frac{1}{2}p}2^{\frac{1}{2}Np}\prod^{\frac{1}{4}p(p+1)}\prod_{i=1}^{p}\Gamma(\frac{1}{2}(N-i))$ and $p_{0,I}$ is $p_{\xi,I}$ with $\xi = 0$. Let A be the matrix belonging to G_k for which $A(S + N\bar{X}\bar{X}')A' = I$. Then $A'A = (S + N\bar{X}\bar{X}')^{-1} = S^{-1} - NS^{-1}\bar{X}\bar{X}'S^{-1}/(1 + N\bar{X}'S^{-1}\bar{X})$, so that $N\bar{X}'A'A\bar{X} = N\bar{X}'S^{-1}\bar{X}/(1 + N\bar{X}'S^{-1}\bar{X}) = \sum_{j=1}^{k} R_j$. Since $A_{[ii]}(S_{[ii]} + N\bar{X}_{[i]}\bar{X}'_{[i]})A'_{[ii]} = I_{[i]}$, we obtain similarly $N\bar{X}'_{[i]}A'_{[ii]}A_{[ii]}X'_{[i]} = N\bar{X}'_{[i]}S_{[ii]}^{-1}\bar{X}_{[i]}/(1 + N\bar{X}'_{[i]}S_{[ii]}^{-1} \cdot \bar{X}_{[i]}) = \sum_{j=1}^{i} R_j$. So, we can now define $N^{\frac{1}{2}}A\bar{X}$ as vector $Y = (Y_1 \cdots Y_k)$ where $Y_i'Y_i = R_i (i = 1, \cdots k)$. Writing $gA^{-1} = h$, $f_\Delta(r)/f_0(r)$ can be expressed as I_1/I_0 where

$$(1.5) \quad I_1 = \int \exp[-\tfrac{1}{2}\textstyle\sum_{j=1}^{k}\delta_j]\prod_{i=1}^{k}|\det(h_{ii}h'_{ii})|^{\frac{1}{2}(N-\sigma_i-1)} \cdot \exp[-\frac{1}{2}\operatorname{tr}\{\sum_{j \leqq i=1}^{k} h_{ij}h'_{ij} - 2\sum_{j \leqq i} P_iY_jh'_{ij}\}]\, dh,$$

where the integration is from $-\infty$ to ∞ in each variable and I_0 is the value of I_1 when $\delta_i = 0 (i = 1, \cdots k)$. Proceeding exactly the same way as in Giri (1964a) pp. 186–187 we obtain,

$$(1.6) \quad f_\Delta(r)/f_0(r) = \exp[-\tfrac{1}{2}(\textstyle\sum_{j=1}^{k}\delta_j - \sum_{j=1}^{k} R_j \sum_{i>j}\delta_i)] \cdot \prod_{i=1}^{k}\phi(\frac{1}{2}(N - \sigma_{i-1}), \frac{1}{2}d_i, \frac{1}{2}R_i\delta_i)$$

where ϕ is the confluent hypergeometric function

$$\phi(a, b, x) = \textstyle\sum_{j=0}^{\infty}[\Gamma(a+j)\Gamma(b)/\Gamma(a)\Gamma(b+j)j!]x^j.$$

Hence from (1.3) and (1.6),

$$f_\Delta(r) = \exp\left[-\tfrac{1}{2}\left(\sum_{j=1}^k \delta_j - \sum_{j=1}^k R_j \sum_{i>j} \delta_i\right)\right] \prod_{i=1}^k \phi\left(\tfrac{1}{2}(N - \sigma_{i-1}), \cdot \tfrac{1}{2}d_i, \tfrac{1}{2}r_i\delta_i\right)\left[\Gamma(\tfrac{1}{2}N)/\Gamma(\tfrac{1}{2}(N-p)) \prod_{i=1}^k \Gamma(\tfrac{1}{2}d_i)\right] \cdot \prod_{i=1}^k R_i^{\frac{1}{2}d_i - 1}\left(1 - \sum_{i=1}^k R_i\right)^{\frac{1}{2}(N-p)-1}. \tag{1.7}$$

From now on, we will specialize $k = 3$, $d_1 = q$, $d_2 = p' - q$ and $d_3 = p - p'$. The group G_k in this case reduces to G. Hence the maximal invariant in $(\bar{X}, S)$ under G is $R = (R_1, R_2, R_3)'$. The corresponding maximal invariant in (ξ, Σ) is $\Delta = (\delta_1, \delta_2, \delta_3)'$. In terms of Γ and Σ, the components of Δ can be written as

$$\delta_1 = N(\Sigma_{11}\Gamma_1 + \Sigma_{12}\Gamma_2 + \Sigma_{13}\Gamma_3)'\Sigma_{11}^{-1}(\Sigma_{11}\Gamma_1 + \Sigma_{12}\Gamma_2 + \Sigma_{13}\Gamma_3),$$

$$\delta_1 + \delta_2 = N\begin{pmatrix}\Sigma_{11}\Gamma_1 + \Sigma_{12}\Gamma_2 + \Sigma_{13}\Gamma_3 \\ \Sigma_{21}\Gamma_1 + \Sigma_{22}\Gamma_2 + \Sigma_{23}\Gamma_3\end{pmatrix}' \Sigma_{[22]}^{-1} \begin{pmatrix}\Sigma_{11}\Gamma_1 + \Sigma_{12}\Gamma_2 + \Sigma_{13}\Gamma_3 \\ \Sigma_{21}\Gamma_1 + \Sigma_{22}\Gamma_2 + \Sigma_{23}\Gamma_3\end{pmatrix},$$

$$\delta_3 = N\Gamma_3'\left(\Sigma_{33} - \begin{pmatrix}\Sigma_{13} \\ \Sigma_{23}\end{pmatrix}' \Sigma_{[22]}^{-1} \begin{pmatrix}\Sigma_{13} \\ \Sigma_{23}\end{pmatrix}\right)\Gamma_3.$$

Thus the maximal invariant in (ξ, Σ) under G when H_1 is true is $\Delta(H_1) = (\delta_1, \delta_2, 0)$. The corresponding maximal invariant under H_0 takes on the value $\Delta(H_0) = (\delta_1, 0, 0)$. Let $f_{\Delta(H_i)}(r)$ be the probability density of R under $H_i (i = 0, 1)$. Then

$$f_{\Delta(H_0)}(r) = \exp\left[-\tfrac{1}{2}\delta_1\right]\phi(\tfrac{1}{2}N, \tfrac{1}{2}q, \tfrac{1}{2}r_1\delta_1) \cdot \{\Gamma(\tfrac{1}{2}N)/[\Gamma(\tfrac{1}{2}(N-p))\Gamma(\tfrac{1}{2}q)\Gamma(\tfrac{1}{2}(p'-q))\Gamma(\tfrac{1}{2}(p-p'))]\} \cdot r_1^{\frac{1}{2}q-1} r_2^{\frac{1}{2}(p'-q)-1} r_3^{\frac{1}{2}(p-p')-1}\left(1 - \sum_{i=1}^3 r_i\right)^{\frac{1}{2}(N-p)-1}. \tag{1.8}$$

REMARKS. 1. From (1.8), R_1 is sufficient for $\Delta(H_0)$.

2. The probability distribution of Z and R_1 under H_0 is

$$\exp\left(-\frac{1}{2}\delta_1\right)\sum_{r=0}^{\infty}\frac{(\frac{1}{2}r_1\delta_1)^r r_1^{\frac{1}{2}q-1}(1-r_1)^{\frac{1}{2}(N-q)-1}Z^{\frac{1}{2}(N-p')-1}(1-Z)^{\frac{1}{2}(p'-q)-1}}{r!\,B(\frac{1}{2}(N-q), \frac{1}{2}q + r)\,B(\frac{1}{2}(N-p'), \frac{1}{2}(p'-q))}$$

where B is the beta function. Hence under H_0 Z is independent of R_1.

3. The distribution of R_1 under H_0 is

$$\exp(-\tfrac{1}{2}\delta_1)\phi(\tfrac{1}{2}N, \tfrac{1}{2}q, \tfrac{1}{2}r_1\delta_1)[\Gamma(\tfrac{1}{2}N)/\Gamma(\tfrac{1}{2}q)\Gamma(\tfrac{1}{2}(N-q))]r_1^{\frac{1}{2}q-1}(1-r_1)^{\frac{1}{2}(N-q)-1}.$$

Hence, from Giri (1964a) p. 188, the family of distributions $\{P_{\delta_1}(r_1), \delta_1 \geqq 0\}$ is boundedly complete.

2. The uniformly most powerful invariant test of H_0 against H_1. Let $\phi(r)$ be any invariant level α test of H_0 against H_1. Since R_1 is boundedly complete, it is well known that ϕ has Neyman structure with respect to R_1 (see for example Lehmann (1959), p. 134), i.e.

$$E_{H_0}(\phi(r)/R = r_1) = \alpha. \tag{2.1}$$

Now, to find the uniformly most powerful test we need the ratio,

$$\begin{aligned}
&f_{\Delta(H_1)}(r/R_1 = r_1)/f_{\Delta(H_0)}(r/R_1 = r_1)\\
&\quad = [f_{\Delta(H_1)}(r)/f_{\Delta(H_0)}(r)][f_{\Delta(H_0)}(r_1)/f_{\Delta(H_1)}(r_1)]\\
&\quad = [f_{\Delta(H_0)}(r_1)/f_{\Delta(H_1)}(r_1)]\cdot\exp[-\tfrac{1}{2}\delta_2(1 - r_1)\\
&\quad\quad \cdot\textstyle\sum_{r=0}^{\infty}[(r_2\tfrac{1}{2}\delta_2)^r/r!][\Gamma(\tfrac{1}{2}(N - q) + r)\Gamma(\tfrac{1}{2}(p' - q))]/\\
&\quad\quad [\Gamma(\tfrac{1}{2}(N - q))\Gamma(\tfrac{1}{2}(p' - q) + r)].
\end{aligned} \tag{2.2}$$

Since the distribution of R given $R_1 = r_1$ is independent of $\Delta(H_0)$, the Condition (2.1) essentially reduces the problem to that of testing the simple hypothesis $\delta_2 = 0$ against the alternative $\delta_2 > 0$ on each surface $R_1 = r_1$. In this conditional situation, by Neyman and Pearson's fundamental lemma, the most powerful level α invariant test $\phi(r/R_1 = r_1)$ for testing $\delta_2 = 0$ against $\delta_2 = \delta_2$ is, from (2.2), given by

$$\begin{aligned}
\phi(r/R_1 = r_1) &= 1, \quad \text{if} \sum_{r=0}^{\infty}\frac{(r_2\,\frac{1}{2}\delta_2)^r}{r!}\\
&\qquad\cdot\frac{\Gamma(\frac{1}{2}(N - q) + r)\,\Gamma(\frac{1}{2}(p' - q))}{\Gamma(\frac{1}{2}(p' - q) + r)\,\Gamma(\frac{1}{2}(N - q))} \geqq C(r_1);\\
&= 0, \quad \text{otherwise};
\end{aligned} \tag{2.3}$$

where $C(r_1)$ is chosen in such a way that $E_{\delta_2=0}\phi(r/R_1 = r_1) = \alpha$. Since $R_2 = (1 - R_1)(1 - Z)$ and Z is independent of R_1 under H_0 the above condition reduces to $\phi(r/R_1 = r_1) = 1$ if $Z \leqq C'$; $= 0$, otherwise; where C' is independent of r_1 and is given by $E_{\delta_2=0}(\phi(r/R_1 = r_1)) = \alpha$. Furthermore, ϕ is independent of δ_2. Hence we have the following theorem:

THEOREM 2.1. *Given the observations $X^1 \cdots X^N (N > p)$, the likelihood ratio test of $H_0 : \Gamma_3 = \Gamma_2 = 0$ against the alternative $H_1 : \Gamma_3 = 0, \Gamma_2 \neq 0$ when ξ, Σ are both unknown, is uniformly most powerful invariant similar.*

Since there exists a right invariant Haar measure on G (i.e., $d\mu(g) = (\det g_{11})^p\cdot(\det g_{22})^{(p'-q)}\cdot(\det g_{33})^{(p-p')}\cdot dg$) from the above theorem and from Lehmann (1959), p. 226, it follows that the likelihood ratio test in this case is uniformly most powerful similar among the group of tests with power depending only on $\Delta(H_1)$.

REFERENCES

[1] GIRI, N. (1964a). On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Statist.* **35** 181–189.

[2] GIRI, N. (1964b). Correction to "On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Statist.* **35** 1388.

[3] GIRI, N., KIEFER, J. and STEIN, C. (1963). Minimax character of Hotelling's T^2-test in the simplest case. *Ann. Math. Statist.* **34** 1524–1535.

[4] LEHMANN, E. L. (1959). *Testing Statistical Hypothesis.* Wiley, New York.

[5] STEIN, C. M. (1956). Some problems in multivariate analysis, Part 1. Technical Report No. 6, Department of Statistics, Stanford Univ.