A NOTE ON THE NON-NULL DISTRIBUTION OF THE WILKS STATISTIC IN MANOVA

Y. H. ASOH AND MASASHI OKAMOTO

(Received Aug. 2, 1967; revised Oct. 9, 1968)

1. Introduction

The null distribution of the Wilks statistic [5], the likelihood ratio statistic in multivariate analysis of variance, is known to be expressed as a product of independent beta variates (see, e.g. Anderson [1], Chapter 8). Using random orthogonal transformations due to Wijsman [4], Katti [2] gave an alternative proof of this fact and extended it to the non-null case when the mean matrix has the rank one. The purpose of this note is to show that the general non-null distribution of the statistic can be expressed as a product of conditional beta variates and also to present a stochastic inequality which gives a lower bound of the Wilks statistic.

2. Non-null distribution of the Wilks statistic

Let the columns of a $p\times m$ matrix $X=(x_{ij})$ and a $p\times n$ matrix $Y=(y_{ij})$, $p\leqq n$, be distributed independently in p-variate normal distributions with a common positive definite variance matrix Σ and let $E(X)=M$, $E(Y)=0$. Then the Wilks statistic is defined as

$$\Lambda=\frac{|B_{pn}|}{|A_{pm}+B_{pn}|}\ , \tag{1}$$

where we put $A_{pm}=XX'$ and $B_{pn}=YY'$.

Let $\delta_1^2\geqq\delta_2^2\geqq\cdots\geqq\delta_p^2$ be the roots of the determinantal equation

$$|MM'-\delta^2\Sigma|=0 \tag{2}$$

and δ_i be the non-negative square root of δ_i^2. The sampling distribution of Λ will be called the Wilks distribution with p dimensions, n and m degrees of freedom and noncentrality vector $(\delta_1^2,\cdots,\delta_p^2)$. This distribution or a variate distributed in it will be denoted by $W_p(n, m;\ \delta_1^2,\cdots,\delta_p^2)$ or simply by $W_p(n, m;\ \delta_1^2,\cdots,\delta_r^2)$, where $r=\operatorname{rank}(M)$.

To find the Wilks distribution, we can assume without loss of generality that

$$(3) \qquad \Sigma = I_p, \qquad M = \left\| \begin{array}{ccc:c} \delta_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \delta_r & \\ \hdashline & 0 & & 0 \end{array} \right\| ,$$

so that $\delta_1^2, \cdots, \delta_r^2$ are the positive eigenvalues of MM' (or of $M'M$). As in Katti [2], denote by $x_{i.}$ and $y_{i.}$ the ith rows of the matrices X and Y, respectively. Let

$$(4) \qquad c_1 = y_{1.}'/(y_{1.}y_{1.}')^{1/2}, \qquad c_2 = x_{1.}'/(x_{1.}x_{1.}' + y_{1.}y_{1.}')^{1/2}.$$

Then, there exist orthogonal matrices P and Q of order n and $m+n$, respectively, of the following forms:

$$(5) \qquad P = \| c_1 : \Omega_B(n \times \overline{n-1}) \| , \qquad Q = \left\| \begin{array}{cccc} c_2 & c_4 & \Omega_A(m \times \overline{m-1}) & 0 \\ c_3 & c_5 & 0 & \Omega_B \end{array} \right\|$$

with suitable column vectors c_i, $i=3, 4, 5$. Define

$$(6) \qquad Z = (z_{ij}) = YP, \qquad W = (w_{ij}) = \| X : Y \| Q .$$

Denote by V and T the $(p-1)\times(n-1)$ and $(p-1)\times(m+n-1)$ matrices obtained by deleting the first row and column from Z and W, respectively. Then it follows that

$$(7) \qquad T = \| U : V \| ,$$

$$(8) \qquad U = \left\| \begin{array}{cc} x_{2.} & y_{2.} \\ \vdots & \vdots \\ x_{p.} & y_{p.} \end{array} \right\| R, \qquad R = \left\| \begin{array}{cc} c_4 & \Omega_A \\ c_5 & 0 \end{array} \right\| .$$

If we define finally $B_{p-1,n-1} = VV'$ and $A_{p-1,m} = UU'$, then we can obtain a recurrence relation

$$(9) \qquad \Lambda = \frac{z_{11}^2}{w_{11}^2} \frac{|B_{p-1,n-1}|}{|A_{p-1,m} + B_{p-1,n-1}|} ,$$

which implies that when $r=1$

$$(10) \qquad W_p(n, m; \delta_1^2) = Be\left(\frac{1}{2}n, \frac{1}{2}m; \delta_1^2\right) W_{p-1}(n-1, m) ,$$

a product of a noncentral beta variate and an independent Wilks variate.

Now we shall find the distribution of Λ when $r=2$. Since the right-hand member of (6) represents a random orthogonal transformation, the conditional distribution of all elements of U and V given $x_{1.}$ and $y_{1.}$ is independent normal with unit variance and with the means determined by

$$(11) \qquad E(V)=0\ , \qquad E(U)=\begin{Vmatrix} \delta_2 r_{2.} \\ 0 \end{Vmatrix}=M_{p-1}\ (\text{say})\ ,$$

where $r_{2.}$ denotes the second row of the matrix R defined in (8). Since the second row $q_{2.}$ of the matrix Q defined in (5) satisfies the relation $q_{2.}q'_{2.}=1$ and since the second component c_{22} of the vector c_2 is given by

$$c_{22}=x_{12}/(x_{1.}x'_{1.}+y_{1.}y'_{1.})^{1/2}\ ,$$

we have

$$r_{2.}r'_{2.}=q_{2.}q'_{2.}-c_{22}^2=1-x_{12}^2/(x_{1.}x'_{1.}+y_{1.}y'_{1.})\ .$$

From (11) it follows that all the eigenvalues of the matrix $M_{p-1}M'_{p-1}$ is zero except the only one which is

$$(12) \qquad \Delta_2^2=\delta_2^2\left(1-\frac{x_{12}^2}{x_{1.}x'_{1.}+y_{1.}y'_{1.}}\right)\ .$$

Thus, as a generalization of (10), we can obtain

$$(13) \qquad W_p(n, m;\ \delta_1^2, \delta_2^2)=Be\left(\frac{n}{2}, \frac{m}{2};\ \delta_1^2\right) \times Be\left(\frac{n-1}{2}, \frac{m}{2};\ \Delta_2^2\right) W_{p-2}(n-2, m)\ ,$$

where conditionally (given $x_{1.}$ and $y_{1.}$) the second and the third factors in the right-hand side are independent beta and Wilks variates, respectively.

A further extension to the general non-null case can be obtained in a similar fashion but the distribution is so complicated already for the case $r=2$ that the stochastic inequality given in the following section may be useful for the purpose of evaluating the power of the test.

3. A stochastic inequality for the Wilks statistic

Since $\Delta_2<\delta_2$ by (12) with probability one and since the family of noncentral beta distributions $Be(a, b;\ \lambda)$ with fixed a and b and with $0<\lambda<\infty$ has the monotone decreasing likelihood ratio with respect to the parameter λ, it holds that

$$(14)\qquad Be\left(\frac{n-1}{2}, \frac{m}{2}; \Delta_2^2\right) \succ Be\left(\frac{n-1}{2}, \frac{m}{2}; \delta_2^2\right),$$

where the symbol $\succ$ denotes a relation "stochastically larger". Then (13) implies

$$(15)\qquad W_p(n, m; \delta_1^2, \delta_2^2) \succ Be\left(\frac{n}{2}, \frac{m}{2}; \delta_1^2\right) Be\left(\frac{n-1}{2}, \frac{m}{2}; \delta_2^2\right) \times W_{p-2}(n-2, m).$$

In this formula three factors in the right-hand side are distributed independently, while it was not necessarily the case with (13).

In order to extend this inequality to any value of r to the effect that

$$(16)\qquad W_p(n, m; \delta_1^2, \cdots, \delta_p^2) \succ \prod_{i=1}^{p} Be\left(\frac{n+1-i}{2}, \frac{m}{2}; \delta_i^2\right),$$

we need the following lemma (see, e.g. [3], p. 52):

LEMMA. *For any real symmetric matrix A of order p let $\lambda_i(A)$ denote the ith eigenvalue of A in order of decreasing magnitude. If a $p\times k$ matrix Q satisfies $Q'Q=I_k$, then for any i it holds that $\lambda_i(Q'AQ)\leqq \lambda_i(A)$.*

Now denote by M^* the $(p-1)\times m$ matrix obtained by deleting the first row of the matrix M. Then by (8) we have

$$M_{p-1}=E(U)=\|M^*:0\|R, \qquad R'R=I_m.$$

Therefore the lemma implies that

$$\lambda_i(M'_{p-1}M_{p-1})\leqq \lambda_i(M^{*\prime}M^*)=\delta_{i+1}^2$$

for any i, i.e., a random orthogonal transformation (with the first rows of X and Y fixed) yields smaller eigenvalues of the conditional mean matrix than the original eigenvalues of the unconditional mean matrix. Hence it will be easily seen that mathematical induction leads to (16).

If we can evaluate by some means the distribution function $F_B(x)$ of the product of noncentral beta variates in the right-hand side of (16), then it will give an upper bound of the distribution function $F_W(x)$ of the Wilks statistic. Thus a conservative evaluation of the power of the likelihood ratio test will be provided, which is expected to be a good approximation when n and/or m are large, since $\Delta_2 \to \delta_2$ in probability as $n\to\infty$ and/or $m\to\infty$.

SHIMANE UNIVERSITY
OSAKA UNIVERSITY

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York, 1958.

[2] S. K. Katti, "Distribution of the likelihood ratio for testing multivariate linear hypotheses," *Ann. Math. Statist.*, 32 (1961), 333-335.

[3] C. R. Rao, *Linear Statistical Inference and Its Applications*, John Wiley, New York, 1965.

[4] Robert A. Wijsman, "Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis," *Ann. Math. Statist.*, 28 (1957), 415-423.

[5] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika* 24 (1932), 471-494.