The Annals of Statistics
1993, Vol. 21, No. 3 1379–1384

ONE-SIDED TEST FOR THE EQUALITY OF TWO COVARIANCE MATRICES

BY SATOSHI KURIKI

University of Tokyo

Let $\mathbf{H}$ and $\mathbf{G}$ be independently distributed according to the Wishart distributions $W_m(M, \boldsymbol{\Phi})$ and $W_m(N, \boldsymbol{\Psi})$, respectively. We derive the limiting null distributions of the likelihood ratio criteria for testing H_0: $\boldsymbol{\Phi} = \boldsymbol{\Psi}$ against $H_1 - H_0$ with H_1: $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$; and for testing $H_0^{(R)}$: $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$, $\operatorname{rank}(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \leq R$ (for given R) against $H_1 - H_0^{(R)}$. They are particular cases of the chi-bar-squared distributions.

1. Introduction. Let $\mathbf{H}$ and $\mathbf{G}$ be $m \times m$ random matrices which are independently distributed according to the Wishart distributions $W_m(M, \boldsymbol{\Phi})$ and $W_m(N, \boldsymbol{\Psi})$, respectively, where $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are assumed to be positive definite and $M \geq m$, $N \geq m$. Consider the likelihood ratio test (LRT) for testing H_0: $\boldsymbol{\Phi} = \boldsymbol{\Psi}$, the hypothesis of the equality of two covariance matrices, against the one-sided alternative $H_1 - H_0$ with H_1: $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$, that is, $\boldsymbol{\Phi} - \boldsymbol{\Psi}$ is nonnegative definite. Furthermore consider the LRT for testing the more general null hypothesis that $H_0^{(R)}$: $\boldsymbol{\Phi} \geq \boldsymbol{\Psi}$, rank $(\boldsymbol{\Phi} - \boldsymbol{\Psi}) \leq R$ for a specified R, $0 \leq R < m$, against the alternative $H_1 - H_0^{(R)}$. Note that $H_0^{(0)}$ is equivalent to H_0. The main purpose of this paper is to derive the limiting null distributions of the likelihood ratio criteria for these testing problems.

These testing problems appear in the multivariate variance components model in the balanced case:

$$\mathbf{X}_{ij} = \boldsymbol{\mu} + \mathbf{V}_i + \mathbf{U}_{ij}, \qquad i = 1, \dots, n;\ j = 1, \dots, k, \tag{1.1}$$

where $\mathbf{X}_{ij}$ is an $m \times 1$ observed vector, $\boldsymbol{\mu}$ an unknown mean vector, $\mathbf{V}_i$ an unobserved random effect vector of group i, and $\mathbf{U}_{ij}$ an unobserved measurement error. $\mathbf{V}_i$ and $\mathbf{U}_{ij}$ are assumed to be independently distributed according to the normal distributions $N_m(\mathbf{0}, \boldsymbol{\Theta})$ and $N_m(\mathbf{0}, \boldsymbol{\Psi})$, respectively. The sufficient statistics of the model (1.1) are $\overline{\overline{\mathbf{X}}} = \sum_{i=1}^n \overline{\mathbf{X}}_i / n$ with $\overline{\mathbf{X}}_i = \sum_{j=1}^k \mathbf{X}_{ij}/k$,

$$\mathbf{H} = k \sum_{i=1}^{n} \left(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}\right)\left(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}\right)'$$

and

$$\mathbf{G} = \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\mathbf{X}_{ij} - \overline{\mathbf{X}}_i\right)\left(\mathbf{X}_{ij} - \overline{\mathbf{X}}_i\right)'.$$

Received October 1990; revised June 1992.

AMS 1991 *subject classifications*. Primary 62H10; secondary 62H15.

Key words and phrases. Chi-bar-squared distribution, ordered restricted inference, multivariate variance components model.

Here $\mathbf{H}$ and $\mathbf{G}$ are independently distributed according to $W_m(n-1, \mathbf{\Psi} + k\mathbf{\Theta})$ and $W_m(n(k-1), \mathbf{\Psi})$, respectively. Then testing the hypothesis of no effect: $\mathbf{\Theta} = \mathbf{O}$ reduces to testing H_0 based on $\mathbf{H}$ and $\mathbf{G}$ against the one-sided alternative $H_1 - H_0$. Testing the hypothesis that the effect vectors are linear combinations of R or less dimensional factors reduces to testing $H_0^{(R)}$ against the alternative $H_1 - H_0^{(R)}$.

The likelihood ratio criterion $\Lambda^{(R)}$ for testing $H_0^{(R)}$ against $H_1 - H_0^{(R)}$ was obtained by Anderson, Anderson and Olkin (1986) as

$$\Lambda^{(R)} = \begin{cases} \prod_{i=R+1}^{R^*} \left\{ \dfrac{l_i^{\rho}}{\rho l_i + 1 - \rho} \right\}^{(M+N)/2}, & \text{if } R^* > R, \\ = 1, & \text{otherwise,} \end{cases}$$

where $l_1 > \cdots > l_m$ (> 0) are the latent roots of $(N/M)\mathbf{H}\mathbf{G}^{-1}$, $\rho = M/(M+N)$, and R^* is the number of $l_i > 1$. In this case the limiting null distribution of (-2) times the logarithm of the likelihood ratio criterion is not chi-squared distribution. Anderson (1989) showed that, under $H_0^{(R)}$ and $\operatorname{rank}(\mathbf{\Phi} - \mathbf{\Psi}) = R$, as letting $M, N \to \infty$ with $\rho \to \rho_0$, $0 \le \rho_0 \le 1$, $-2 \log \Lambda^{(R)}$ converges in distribution to

$$Y = \sum_{i=1}^{p} (b_i \vee 0)^2, \tag{1.2}$$

where $p = m - R$, $b_1 > \cdots > b_p$ are the latent roots of a $p \times p$ symmetric random matrix $\mathbf{A}$ with normal density

$$\frac{1}{2^{p/2} \pi^{p(p+1)/4}} \exp\left\{ -\frac{1}{2} \operatorname{tr} \mathbf{A}^2 \right\}$$

and $x \vee y = \max(x, y)$. From this fact Anderson (1989) derived the distribution function of Y for $p = 2$. For $p \ge 3$, Amemiya, Anderson and Lewis (1990) gave the table of the estimated quantiles of the distribution of Y by Monte Carlo simulations. In the following section we derive a method to calculate the distribution function of Y for general p, and give the table of the quantiles of the distribution of Y.

2. Results. The characteristic function of Y is written as

$$\phi(t) = \sum_{r=0}^{p} \int_{B_r \times \overline{B}_{p-r}} \exp\left\{ it \sum_{i=1}^{r} b_i^2 \right\} \varphi(\mathbf{b}) \, db_1 \cdots db_p, \tag{2.1}$$

where

$$\varphi(\mathbf{b}) = d(p) \exp\left\{ -\frac{1}{2} \sum_{i=1}^{p} b_i^2 \right\} \prod_{i<j} (b_i - b_j) \tag{2.2}$$

with

$$d(p) = \frac{\pi^{p(p-1)/4}}{2^{p/2}\Gamma_p(p/2)} = \frac{1}{2^{p/2}\prod_{i=1}^{p}\Gamma(i/2)}$$

is the joint density of $\mathbf{b} = (b_1, \dots, b_p)$, $b_1 > \cdots > b_p$ [Anderson (1984), Theorem 13.3.5], and

$$B_r = \{(b_1, \dots, b_r) | b_1 > \cdots > b_r > 0\},$$
$$\bar{B}_{p-r} = \{(b_{r+1}, \dots, b_p) | 0 > b_{r+1} > \cdots > b_p\}.$$

By the Laplace expansion of the linkage factor $\prod_{i<j}(b_i - b_j)$ contained in the density $\varphi(\mathbf{b})$ in (2.2), which is the Vandermonde determinant $\det(b_i^{p-j})_{1\le i,\, j\le p}$, the R.H.S. of (2.1) becomes

$$d(p)\sum_{r=0}^{p}\sum_{q}(-1)^{\sum_{i=1}^{r}(i+p-q_i)}\int_{B_r}\exp\left\{-\frac{1}{2\theta^2}\sum_{i=1}^{r}b_i^2\right\}\det(b_i^{q_j})_{1\le i,\, j\le r}\, db_1 \cdots db_r$$
$$\times\int_{\bar{B}_{p-r}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{p-r}b_{i+r}^2\right\}\det(b_{i+r}^{\bar{q}_j})_{1\le i,\, j\le p-r}\, db_{r+1} \cdots db_p,$$

where $\theta = (1 - 2it)^{-1/2}$, and Σ_q is summation over all combinations of $q_1 > \cdots > q_r$, $\bar{q}_1 > \cdots > \bar{q}_{p-r}$ such that $\{q_1, \dots, q_r, \bar{q}_1, \dots, \bar{q}_{p-r}\} = \{0, \dots, p-1\}$. Then, putting

$$(2.3)\quad U_r(q_1, \dots, q_r) = \int_{B_r}\exp\left\{-\frac{1}{2}\sum_{i=1}^{r}b_i^2\right\}\det(b_i^{q_j})_{1\le i,\, j\le r}\, db_1 \cdots db_r,$$

$r \ge 1$, and $U_0 = 1$, we see that

$$(2.4)\qquad \phi(t) = d(p)\sum_{r=0}^{p}\sum_{q}U_r(q_1, \dots, q_r)U_{p-r}(\bar{q}_1, \dots, \bar{q}_{p-r})\theta^{Q}$$

with $Q = \sum_{i=1}^{r}q_i + r$. By inverting the characteristic function of (2.4), we get the following theorem.

THEOREM 2.1. *The distribution function of Y in* (1.2) *is given as*:

$$\Pr(Y \le y) = d(p)\sum_{r=0}^{p}\sum_{q}U_r(q_1, \dots, q_r)U_{p-r}(\bar{q}_1, \dots, \bar{q}_{p-r})\Pr(\chi^2(Q) \le y),$$

where $\chi^2(\nu)$ *is a chi-square random variable with* ν *degrees of freedom, and* $\chi^2(0) \equiv 0$.

REMARK 2.1. The distribution function of Y is a mixture of the chi-squared distributions with ν, $0 \le \nu \le p(p+1)/2$, degrees of freedom. This is a particular case of the chi-bar-squared ($\bar{\chi}^2$) distribution which appears as the limiting null distribution of (-2) times the logarithm of the likelihood ratio criterion when the null hypothesis is on the boundary of the parameter space [e.g., Robertson, Wright and Dykstra (1988), Chapter 2; see also Shapiro (1988), Section 6].

In order to calculate the distribution function of Y given in Theorem 2.1, the integral $U_r(q_1,\dots,q_r)$ of (2.3) has to be evaluated. The "pseudo determinant" method of Pillai (1956) is exploited here.

THEOREM 2.2. *$U_r(q_1,\dots,q_r)$ can be evaluated by the following recurrence formula:*

$$\begin{aligned} U_r(q_1,\dots,q_r) &= (-1)^{r-1}U_{r-1}(q_2,\dots,q_r)I(q_1=1) \\ &\quad +(q_1-1)U_r(q_1-2,q_2,\dots,q_r) \\ &\quad +2\sum_{j=2}^{r}(-1)^j\frac{1}{2^{(q_1+q_j)/2}}U_1(q_1+q_j-1) \\ &\quad \times U_{r-2}(q_2,\dots,q_{j-1},q_{j+1},\dots,q_r), \end{aligned} \tag{2.5}$$

$r\ge 2$, $q_1\ge 1$, and

$$U_1(q) = I(q=1)+(q-1)U_1(q-2),\qquad q\ge 1,$$
$$U_1(0)=\sqrt{\pi/2},$$

where $I(\cdot)$ denotes indicator function.

PROOF. Define

$$F_{r,i}(b;q_1,\dots,q_r) = \int_{B_{r,i}(b)}\exp\left\{-\frac{1}{2}\sum_{i=1}^{r}b_i^2\right\}\det(b_i^{q_j})_{1\le i,j\le r}\,db_1\cdots db_r,$$

$0\le i\le r$, with

$$B_{r,i}(b)=\{(b_1,\dots,b_r)|b_1>\cdots>b_i>b>b_{i+1}>\cdots>b_r>0\}.$$

Then, by expanding the first column of $\det(b_i^{q_j})_{1\le i,j\le r}$ in the R.H.S. of (2.3), we have

$$\begin{aligned} U_r(q_1,\dots,q_r) &= \sum_{i=1}^{r}(-1)^{i-1}\int_0^\infty \exp\{-\tfrac{1}{2}b^2\}b^{q_1}\,db \\ &\quad \times F_{r-1,i-1}(b;q_2,\dots,q_r). \end{aligned} \tag{2.6}$$

By integration by parts, and using the relations that

$$\begin{aligned} &\frac{d}{db}F_{r-1,i-1}(b;q_2,\dots,q_r) \\ &\quad = -\sum_{j=2}^{r}(-1)^{i+j}\exp\{-\tfrac{1}{2}b^2\}b^{q_j}\{F_{r-2,i-2}(b;q_2,\dots,q_{j-1},q_{j+1},\dots,q_r) \\ &\qquad +F_{r-2,i-1}(b;q_2,\dots,q_{j-1},q_{j+1},\dots,q_r)\}, \end{aligned}$$

TABLE 1
Quantiles of the distribution of Y

Prob.	p = 2	3	4	5	6	7	8
0.010	0.0000	0.0000	0.0986	0.6832	1.7747	3.3713	5.4720
0.025	0.0000	9.1E − 7	0.3178	1.1788	2.5559	4.4424	6.8352
0.050	0.0000	0.0508	0.6207	1.7319	3.3617	5.5021	8.1492
0.100	0.0000	0.2423	1.1146	2.5241	4.4511	6.8881	9.8313
0.250	0.0957	0.9190	2.3181	4.2400	6.6723	9.6108	13.0531
0.500	0.7717	2.2605	4.2581	6.7572	9.7567	13.2563	17.2561
0.750	2.1535	4.2858	6.8788	9.9539	13.5198	17.5803	22.1373
0.900	4.0457	6.7324	9.8503	13.4372	17.5088	22.0709	27.1272
0.950	5.4845	8.4904	11.9156	15.8047	20.1758	25.0360	30.3894
0.975	6.9229	10.1978	13.8848	18.0330	22.6618	27.7789	33.3888
0.990	8.8211	12.3994	16.3852	20.8310	25.7566	31.1703	37.0766

Prob.	p = 9	10	11	12	13	14
0.010	8.0757	11.1817	14.7894	18.8984	23.5086	28.6196
0.025	9.7322	13.1324	17.0348	21.4389	26.3445	31.7511
0.050	11.3010	14.9559	19.1133	23.7724	28.9330	34.5947
0.100	13.2787	17.2291	21.6817	26.6359	32.0913	38.0478
0.250	16.9980	21.4446	26.3925	31.8415	37.7912	44.2415
0.500	21.7559	26.7557	32.2556	38.2555	44.7554	51.7553
0.750	27.1920	32.7450	38.7967	45.3476	52.3977	59.9472
0.900	32.6794	38.7288	45.2761	52.3218	59.8663	67.9098
0.950	36.2381	42.5835	49.4264	56.7676	64.6072	72.9458
0.975	39.4936	46.0949	53.1935	60.7903	68.8855	77.4795
0.990	43.4779	50.3756	57.7707	65.6639	74.0555	82.9460

and that $\Sigma_{i=0}^{r-2} F_{r-2,i}(b; q_2, \dots) = U_{r-2}(q_2, \dots)$ for any $b > 0$, we can see that the R.H.S. of (2.6) reduces to the R.H.S. of (2.5). The case of $r = 1$ is easy. □

REMARK 2.2. Since $U_r(q_1, \dots, q_r)$ is a skew-symmetric function of $q_1, \dots, q_r$, we can restrict ourselves to $q_1 > \cdots > q_r$, and the second term of the R.H.S. of (2.5) can be replaced by:

$$\begin{array}{ll} (q_1 - 1)U_r(q_1 - 2, q_2, \dots, q_r), & \text{if } q_1 - 2 > q_2, \\ -(q_1 - 1)U_r(q_2, q_1 - 2, q_3, \dots, q_r), & \text{if } q_2 > q_1 - 2 > q_3, \\ 0, & \text{otherwise}. \end{array}$$

The quantiles of the distribution of Y for $2 \le p \le 14$ are given in Table 1.

Acknowledgments. The author wishes to express his thanks to Professors C. Hirotsu, K. Murota and Dr. T. Kubokawa of the University of Tokyo for their advice and encouragement. He also thanks the referees and an Associate Editor whose comments greatly improved the earlier version of the paper.

REFERENCES

AMEMIYA, Y., ANDERSON, T. W. and LEWIS, P. A. W. (1990). Percentage points for a test of rank in multivariate components of variance. *Biometrika* **77** 637–641.

ANDERSON, B. M., ANDERSON, T. W. and OLKIN, I. (1986). Maximum likelihood estimators and likelihood ratio criteria in multivariate components of variance. *Ann. Statist.* **14** 405–417.

ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York.

ANDERSON, T. W. (1989). The asymptotic distribution of the likelihood ratio criterion for testing rank in multivariate components of variance. *J. Multivariate Anal.* **30** 72–79.

PILLAI, K. C. S. (1956). Some results useful in multivariate analysis. *Ann. Math. Statist.* **27** 1106–1114.

ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, Chichester.

SHAPIRO, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *Internat. Statist. Rev.* **56** 49–62.

DEPARTMENT OF MATHEMATICAL ENGINEERING
AND INFORMATION PHYSICS
FACULTY OF ENGINEERING
UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113
JAPAN