APPROXIMATIONS FOR THE DISTRIBUTIONS OF THE EXTREME LATENT ROOTS OF THREE MATRICES

ROBB J. MUIRHEAD AND YASUKO CHIKUSE

(Received Sept. 9, 1974; revised July 1, 1975)

Summary

In this paper we present simple approximations for the distributions of the extreme latent roots of three matrices occurring in multivariate analysis. The matrices considered are (i) $S_1S_2^{-1}$ where S_1 and S_2 are independent Wishart matrices estimating different covariance matrices, (ii) $S_1S_2^{-1}$ where S_1 and S_2 are independent and estimate the same covariance matrix, with S_2 having the Wishart distribution and S_1 having the noncentral Wishart distribution, and (iii) the noncentral Wishart matrix. The approximations take the form of upper and lower bounds for the distribution functions of the largest and smallest latent roots respectively. For the three matrices considered above these bounds are expressed very simply in terms of products of (i) F, (ii) noncentral F and (iii) noncentral χ^2 probabilities.

1. $S_1S_2^{-1}$; different covariance matrices

Let S_1 and S_2 be the covariance matrices formed from samples of sizes n_1+1 and n_2+1 drawn from two m-variate normal distributions with covariance matrices Σ_1 and Σ_2 respectively; then n_1S_1 and n_2S_2 are independently distributed as Wishart $W_m(n_1, \Sigma_1)$ and $W_m(n_2, \Sigma_2)$ respectively. Let $l_1 \geqq l_2 \geqq \cdots \geqq l_m > 0$ be the latent roots of $S_1S_2^{-1}$. We derive in this section approximations for the distribution functions of l_1 and l_m respectively.

Let A be an $m\times m$ nonsingular matrix such that

$$A\Sigma_2A' = I_m$$

and

$$A\Sigma_1A' = \Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$$

where $\lambda_1, \lambda_2, \cdots, \lambda_m$ are the latent roots of $\Sigma_1\Sigma_2^{-1}$. Putting $S_i^* = AS_iA'$ $(i=1, 2)$ it then follows that $n_1S_1^*$ and $n_2S_2^*$ are independently distributed as $W_m(n_1, \Lambda)$ and $W_m(n_2, I_m)$ respectively, and $l_1, \cdots, l_m$ are the latent roots of $S_1^* S_2^{*-1}$. It is well-known (see e.g. Roy [7]) that

$$l_1 \geqq \frac{\boldsymbol{x}' S_1^* \boldsymbol{x}}{\boldsymbol{x}' S_2^* \boldsymbol{x}} \geqq l_m \,, \qquad \boldsymbol{x}' S_2^* \boldsymbol{x} > 0 \,.$$

Hence, if we let $S_i^* = (s_{kl}^{(i)})$ $(i=1, 2)$ it follows easily that

$$l_1 \geqq \max\left(\frac{s_{11}^{(1)}}{s_{11}^{(2)}}, \cdots, \frac{s_{mm}^{(1)}}{s_{mm}^{(2)}}\right) \tag{1}$$

and

$$l_m \leqq \min\left(\frac{s_{11}^{(1)}}{s_{11}^{(2)}}, \cdots, \frac{s_{mm}^{(1)}}{s_{mm}^{(2)}}\right) . \tag{2}$$

Now, the $s_{ii}^{(1)}$ and the $s_{ii}^{(2)}$ $(i=1, \cdots, m)$ are all independent, with $n_1 s_{ii}^{(1)}/\lambda_i$ and $n_2 s_{ii}^{(2)}$ having $\chi^2_{n_1}$ and $\chi^2_{n_2}$ distributions respectively; hence the $s_{ii}^{(1)}/\lambda_i s_{ii}^{(2)}$ $(i=1, \cdots, m)$ have independent F_{n_1, n_2} distributions. Thus, using (1) and (2), the following result is easily obtained.

THEOREM 1. *An upper bound for the distribution function of l_1 is given by*

$$\mathrm{P}(l_1 \leqq x) \leqq \prod_{i=1}^{m} \mathrm{P}\left(F_{n_1, n_2} \leqq \frac{x}{\lambda_i}\right) , \tag{3}$$

and a lower bound for the distribution function of l_m is given by

Table 1. Comparison of bound (3) with exact probabilities

n_1	n_2	λ_1	λ_2	x	Exact $\mathrm{P}(l_1 \leqq x)$	Upper bound (3)	Difference
5	13	1	1	4.850	.950	.980	.030
5	13	1	1	7.596	.990	.997	.007
5	33	1	1	3.523	.950	.977	.027
5	33	1	1	4.878	.990	.996	.006
5	83	1	1	3.115	.950	.975	.025
5	83	1	1	4.150	.990	.996	.006
7	33	1	1	3.169	.950	.978	.028
7	33	8	1	3.169	.070	.101	.031
7	33	11	1	3.169	.030	.046	.016
7	33	6	6	3.169	.013	.037	.024
7	83	1	1	2.781	.950	.976	.026
7	83	8	1	2.781	.053	.070	.017
7	83	11	1	2.781	.022	.030	.008
7	83	6	6	2.781	.008	.020	.012

$$\mathrm{P}\,(l_m \leqq x) \geqq 1 - \prod_{i=1}^{m} \mathrm{P}\left(F_{n_1,n_2} \geqq \frac{x}{\lambda_i}\right) . \tag{4}$$

The bounds are clearly exact when $m=1$, *and, when* $\Lambda = I_m$ *i.e.* $\Sigma_1 = \Sigma_2$, *they agree with bounds given by Mickey* [2].

In Table 1 values of the upper bound (3) are compared with exact values of $\mathrm{P}\,(l_1 \leqq x)$ calculated for $m=2$ by Pillai [4] and Pillai and Al-Ani [5]. The upper-tail of the distribution of l_1 is normally of interest and, as a quick approximation to the exact probability, the bound (3) appears quite reasonable. The accuracy increases the further one goes out in the tail of the distribution. More detailed numerical comparisons made in the case $m=2$ further revealed that for fixed n_1, λ_1 and λ_2, the accuracy of the approximation generally increases as n_2 increases and for fixed n_1, n_2 and λ_2, the accuracy first tends to decrease and then increases, as λ_1 increases.

2. $S_1S_2^{-1}$; MANOVA situation

Let $X(n_1 \times m)$ and $Y(n_2 \times m)$ be independent matrix variates distributed as $N(M, I_{n_1} \otimes \Sigma)$ and $N(0, I_{n_2} \otimes \Sigma)$ respectively. Then $n_1S_1 = X'X$ and $n_2S_2 = Y'Y$ are independently distributed, with n_2S_2 having the Wishart distribution $W_m(n_2, \Sigma)$ and n_1S_1 having the noncentral Wishart distribution $W_m(n_1, \Sigma, \Omega)$ with noncentrality matrix $\Omega = \Sigma^{-1}M'M$. Let $l_1 \geqq l_2 \geqq \cdots \geqq l_m > 0$ be the latent roots of $S_1S_2^{-1}$. We derive in this section approximations for the distribution functions of l_1 and l_m respectively.

Let A be an $m \times m$ nonsingular matrix such that

$$A\Sigma A' = I_m$$

and

$$AM'MA = \Omega_D = \mathrm{diag}\,(\omega_1, \omega_2, \cdots, \omega_m)$$

where $\omega_1, \omega_2, \cdots, \omega_m$ are the latent roots of $\Sigma^{-1}M'M = \Omega$. Putting $S_i^* = AS_iA'$ $(i=1, 2)$ we then have that $n_1S_1^*$ and $n_2S_2^*$ are independently distributed as $W_m(n_1, I_m, \Omega_D)$ and $W_m(n_2, I_m)$ respectively, and $l_1, \cdots, l_m$ are the latent roots of $S_1^*S_2^{*-1}$. Put $S_i^* = (s_{kl}^{(i)})$ $(i=1, 2)$; it then follows that the $s_{ii}^{(1)}$ and the $s_{ii}^{(2)}$ are all independent, with $n_2s_{ii}^{(2)}$ having the $\chi_{n_2}^2$ distribution and $n_1s_{ii}^{(1)}$ having the noncentral $\chi_{n_1}^2(\omega_i)$ distribution with noncentrality parameter ω_i; hence the $s_{ii}^{(1)}/s_{ii}^{(2)}$ have independent noncentral $F_{n_1,n_2}(\omega_i)$ distributions. This fact, together with (1) and (2) yields the following

THEOREM 2. *An upper bound for the distribution function of* l_1 *is* *given by*

$$\mathrm{P}(l_1 \leqq x) \leqq \prod_{i=1}^{m} (F_{n_1,n_2}(\omega_i) \leqq x) \tag{5}$$

and a lower bound for the distribution function of l_m *is given by*

$$\mathrm{P}(l_m \leqq x) \geqq 1 - \prod_{i=1}^{m} \mathrm{P}(F_{n_1,n_2}(\omega_i) \geqq x) . \tag{6}$$

In (5) *and* (6), $F_{n_1,n_2}(\omega_i)$ *denotes a random variable having the noncentral* F *distribution on* n_1 *and* n_2 *degrees of freedom and noncentrality parameter* ω_i.

The bounds are exact when $m=1$, and when $\Omega_D=0$ they again agree with the bounds given by Mickey [2].

In Table 2 values of the upper bound (5) are compared with exact values of $\mathrm{P}(l_1 \leqq x)$ calculated for $m=2$ in the "linear" case when $\omega_2=0$ by Pillai and Jayachandran [6]. Again, as a quick approximation to the exact probability, the bound (5) appears quite reasonable. More detailed numerical comparisons showed that, for fixed n_1 and n_2, the accuracy tends to decrease as ω_1 increases, while for fixed n_1 and ω_1 it increases as n_2 increases. The accuracy, of couse, would increase the further one goes out in the tail of the distribution, i.e. for larger values of x.

Table 2. Comparison of bound (5) with exact probabilities

n_1	n_2	ω_1	x	Exact $\mathrm{P}(l_1 \leqq x)$	Upper bound (5)	Difference
3	33	.01	4.236	.950	.976	.026
3	33	.05	4.236	.948	.975	.027
3	33	.10	4.236	.947	.974	.027
3	83	.01	3.809	.950	.974	.024
3	83	.05	3.809	.948	.973	.025
3	83	.10	3.809	.947	.972	.025
5	33	.01	3.523	.950	.977	.027
5	33	.05	3.523	.949	.976	.027
5	33	.10	3.523	.948	.976	.028
5	83	.01	3.115	.950	.975	.025
5	83	.05	3.115	.949	.974	.025
5	83	.10	3.115	.948	.974	.026

3. Noncentral Wishart matrix

Let $nS=X'X$ where X is an $n \times m$ matrix variate distributed as

$N(M, I_n \otimes \Sigma)$; then nS has the noncentral Wishart distribution $W_m(n, \Sigma, \Omega)$ with noncentrality matrix $\Omega=\Sigma^{-1}M'M$. We will assume that Σ is known, and let $w_1 \geqq w_2 \geqq \cdots \geqq w_m$ be the latent roots of $\Sigma^{-1}S$. We derive here upper and lower bounds for the distribution functions of w_1 and w_m respectively.

As in Section 3, let A be an $m \times m$ nonsingular matrix such that

$$A\Sigma A'=I_m$$

and

$$AM'MA'=\Omega_D=\text{diag}(\omega_1, \omega_2, \cdots, \omega_m)$$

where $\omega_1, \omega_2, \cdots, \omega_m$ are the latent roots of Ω. Then $nS^*=nASA'$ has the $W_m(n, I_m, \Omega_D)$ distribution and $w_1, \cdots, w_m$ are the latent roots of S^*, or equivalently, of $\Sigma^{-1}S$. Then, in the same manner as in Muirhead [3], the well-known inequalities, due to Bellman [1] (p. 111), and the fact that the ns_{ii} $(i=1, \cdots, m)$ have independent $\chi_n^2(\omega_i)$ distributions, yield the following

THEOREM 3. *An upper bound for the distribution function of w_1 is given by*

$$\text{(7)} \qquad \mathrm{P}(w_1 \leqq x) \leqq \prod_{i=1}^{m} \mathrm{P}(\chi_n^2(\omega_i) \leqq nx)$$

and a lower bound for the distribution function of w_m is given by

$$\text{(8)} \qquad \mathrm{P}(w_m \leqq x) \geqq 1-\prod_{i=1}^{m} \mathrm{P}(\chi_n^2(\omega_i) \geqq nx)\ .$$

The bounds are exact when $m=1$ and, when $\Omega_D=0$, i.e. nS^* is $W_m(n, I_m)$, they agree with bounds given by Muirhead [3]. An approximation to $\mathrm{P}(w_1 \leqq x)$ somewhat similar to (7) but expressed solely in terms of central χ^2 probabilities has been given by Sugiyama [8]; however it should be noted that the approximation in [8] requires that n be large.

YALE UNIVERSITY
RADIATION EFFECTS RESEARCH FOUNDATION, JAPAN

REFERENCES

[1] Bellman, R. (1960). *Introduction to Matrix Analysis*, McGraw-Hill, New York.
[2] Mickey, R. (1959). Some bounds on the distribution functions of the largest and smallest roots of normal determinantal equations, *Ann. Math. Statist.*, **30**, 242-243.
[3] Muirhead, R. J. (1974). Bounds for the distribution functions of the extreme latent roots of a sample covariance matrix, *Biometrika*, **61**, 641-642.
[4] Pillai, K. C. S. (1956). On the distribution of the largest or the smallest root of a matrix in multivariate analysis, *Biometrika*, **43**, 122-127.

[5] Pillai, K. C. S. and Al-Ani, S. (1970). Power comparisons of tests of equality of two covariance matrices based on individual characteristic roots, *J. Amer. Statist. Ass.*, **65**, 438-446.

[6] Pillai, K. C. S. and Jayachandran, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria, *Biometrika*, **54**, 195-210.

[7] Roy, S. N. (1939). p-statistics, or some generalizations on the analysis of variance appropriate to multivariate problems, *Sankhya*, **3**, 341-396.

[8] Sugiyama, T. (1972). Approximation for the distribution function of the largest latent root of a Wishart matrix, *Aust. J. Statist.*, **14**, 17-24.