The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1664-1670

ON THE MARGINAL DISTRIBUTIONS OF THE LATENT ROOTS OF THE MULTIVARIATE BETA MATRIX[1]

By A. W. Davis

C.S.I.R.O., Adelaide, and University of North Carolina

The marginal distributions of the latent roots of the multivariate beta matrix are shown to constitute a complete system of solutions of an ordinary differential equation (d.e.), which is related to the author's d.e.'s for Hotelling's generalized T_0^2 and Pillai's $V^{(m)}$ statistics. Results may be derived for the latent roots of the multivariate F and Wishart matrices ($\Sigma = I$). Pillai's approximations to the distributions of the largest and smallest roots are interpreted as exact solutions, the contributions of higher order solutions being neglected.

1. Introduction. Let $S(m \times m)$ and $T(m \times m)$ have independent Wishart distributions $W(q, \Sigma)$ and $W(n, \Sigma)$, respectively, where Σ is the population covariance matrix and $q, n \geqq m$. The latent roots $l_1 > \cdots > l_m > 0$ of the multivariate beta matrix $B = S(S+T)^{-1}$ are well known to have the joint density function

$$\phi_{m;q,n}(\boldsymbol{l}) = K(m; q, n) \prod_{i=1}^m l_i^{\frac{1}{2}(q-m-1)} \prod_{i=1}^m (1-l_i)^{\frac{1}{2}(n-m-1)} \prod_{i<j} (l_i - l_j)\,, \tag{1}$$

where $\boldsymbol{l} = (l_1, \cdots, l_m)'$ and

$$K(m; q, n) = \pi^{\frac{1}{2}m} \prod_{i=0}^{m-1} [\Gamma(\tfrac{1}{2}(q+n-i))/\Gamma(\tfrac{1}{2}(m-i))\Gamma(\tfrac{1}{2}(q-i))\Gamma(\tfrac{1}{2}(n-i))]\,. \tag{2}$$

The marginal distributions of the individual l_i have been investigated by Roy [14], [15], who showed that the largest root l_1 is of basic importance in testing hypotheses and constructing confidence regions in multivariate analysis of variance; also by Pillai [10], Khatri [8], Sugiyama and Fukutomi [17], Sugiyama [16], and Al-Ani [1]. Pillai [11] gave very accurate approximations to the upper and lower tails of the distributions of l_1 and l_m, respectively, and l_1 has been extensively tabulated by Heck [7] for $m \leqq 5$, using Pillai's approximation, and Pillai ([11], [12], etc.) for $m \leqq 20$. Studies of the noncentral distributions have been made by Khatri [9] and Pillai and Dotson [13].

As $n \to \infty$, $nB \to W_I$, say, having the distribution $W(q, I)$ where $I(m \times m)$ is the unit matrix. Hanumara and Thompson [6] have tabulated the largest and smallest roots of W_I using limiting forms of Pillai's approximations, and discussed their application.

The present author [3], [5] has shown that the null distributions for tr B (Pillai's $V^{(m)}$) and tr F (Hotelling's generalized T_0^2), where $F = ST^{-1}$, satisfy certain ordinary linear differential equations (d.e.'s) of order m which are related by a simple transformation. In Section 2 it is shown that the marginal

Received July 10, 1970; revised May 17, 1971.

[1] This research was supported by the National Institutes of Health Grant No. GM-12868.

distributions of the l_i form a complete system of solutions of a similar d.e. Thus, the power-series of Sugiyama and Fukutomi are solutions at the regular singularities 0 and 1. Pillai's approximations are also shown in Section 4 to be exact solutions of the d.e., but approximations to the distributions insofar as contributions from higher-order solutions are neglected. Corresponding results for the latent roots of F and W_I are readily deduced.

2. The differential equation. Let $D^r(s, l) = \{0 < x_r < \cdots < x_s < l < x_{s-1} < \cdots < x_1 < 1\} \subset R^r$, where R is the real line. The marginal density function $f_s(l)$ of l_s is given by

$$f_s(l) = \int_{D^{m-1}(s,l)} \phi_{m;q,n}(x_1, \cdots, x_{s-1}, l, x_s, \cdots x_{m-1})\, d\mathbf{x} \tag{3}$$

where $d\mathbf{x} = \prod_{i=1}^{m-1} dx_i$; it is proportional to

$$l^{\frac{1}{2}(q-m-1)}(1 - l)^{\frac{1}{2}(n-m-1)} \int_{D^{m-1}(s,l)} \Phi(\mathbf{x}) \prod_{i=1}^{m-1} (l - x_i)\, d\mathbf{x}\,, \tag{4}$$

in which Φ denotes $\phi_{m-1;q-1,n-1}$. Define

$$\Psi_r(l; x) = \Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots (l - x_{\alpha(m-1-r)})\,, \quad (r = 0, 1, \cdots, m-1)\,, \tag{5}$$

the summation being extended over the $\binom{m-1}{r}$ selections of integers $\alpha(1) < \cdots < \alpha(m - 1 - r)$ from the set $1, 2, \cdots, m - 1$. When $r = m - 1$, the sum is taken to be unity. We now introduce the m functions

$$L_{s,r}(l) = \int_{D^{m-1}(s,l)} \Psi_r(l; \mathbf{x})\, d\mathbf{x}\,, \qquad (r = 0, 1, \cdots, m-1)\,, \tag{6}$$

noting that $f_s(l)$ is proportional to $l^{\frac{1}{2}(q-m-1)}(1 - l)^{\frac{1}{2}(n-m-1)}L_{s,0}$. Our object is to show that, for each s, the $L_{s,r}$ are related by a system of first-order differential equations which are independent of s. Differentiating (6),

$$L'_{s,r}(l) = -Z^{(1)}_{s,r} + Z^{(2)}_{s,r} + (r+1)L_{s,r+1}\,, \tag{7}$$

where

$$\begin{aligned} Z^{(1)}_{s,r} &= \int_{D^{m-2}(s-1,l)} \Psi_r(l; x_1, \cdots, x_{s-2}, l, x_{s-1}, \cdots, x_{m-2})\, d\mathbf{x}\,, \\ Z^{(2)}_{s,r} &= \int_{D^{m-2}(s,l)} \Psi_r(l; x_1, \cdots, x_{s-1}, l, x_s, \cdots, x_{m-2})\, d\mathbf{x}\,. \end{aligned} \tag{8}$$

Now let $\beta(1), \cdots, \beta(r)$ denote the set of subscripts complementary to $\alpha(1), \cdots, \alpha(m - 1 - r)$. We have

$$\begin{aligned} rlL_{s,r} &= \int_{D^{m-1}(s,l)} \Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots (l - x_{\alpha(m-1-r)}) \\ &\qquad \times [(l - x_{\beta(1)}) + \cdots + (l - x_{\beta(r)}) + (x_{\beta(1)} + \cdots + x_{\beta(r)})]\, d\mathbf{x} \\ &= (m - r)L_{s,r-1} + \Theta_{s,r}\,, \end{aligned} \tag{9}$$

say. Integration by parts with respect to the $x_{\beta(i)}$ yields

$$\begin{aligned} \tfrac{1}{2}(q + n - 2m + 2)\Theta_{s,r} &= l(1 - l)[Z^{(1)}_{s,r} - Z^{(2)}_{s,r}] \\ &\qquad + \tfrac{1}{2}r(q - m + r)L_{s,r} + \Psi_{s,r}\,, \end{aligned} \tag{10}$$

where

$$\begin{aligned} \Psi_{s,r} = \int_{D^{m-1}(s,l)} &\Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots \\ &(l - x_{\alpha(m-1-r)}) \sum_{j=1}^{r} x_{\beta(j)} (1 - x_{\beta(j)}) \sum_{k \neq \beta(j)} (x_{\beta(j)} - x_k)^{-1}\, d\mathbf{x}\,. \end{aligned} \tag{11}$$

A term similar to (11) occurred in the derivation of a d.e. for Hotelling's generalized T_0^2 ([3] (2.13)), and the same approach yields

$$(12)\qquad \Psi_{s,r} = \tfrac{1}{2}(m-r)(m+r-3)L_{s,r-1} + \tfrac{1}{2}r(r-1)(1-2l)L_{s,r} + \tfrac{1}{2}r(r+1)l(1-l)L_{s,r+1}\,.$$

Finally, eliminating the Z's, Θ's and Ψ's from (7), (9), (10) and (12), we find that

$$(13)\qquad \begin{aligned} l(1-l)L'_{s,r} = {} & \tfrac{1}{2}(m-r)(q+n-m+r-1)L_{s,r-1} \\ & + \tfrac{1}{2}r[(1-l)(q-m+r) - l(n-m+r)]L_{s,r} \\ & + \tfrac{1}{2}(r+1)(r+2)l(1-l)L_{s,r+1}\,, \\ & \qquad (r = 0, 1, \cdots, m-1)\,, \end{aligned}$$

where $L_{s,-1} \equiv L_{s,m} \equiv 0$.

We observe that the system (13) is independent of s, and in principle one could successively eliminate $L_{s,1}, \cdots, L_{s,m-1}$, arriving at a homogeneous linear d.e. of order m having each $L_{s,0}$ as a solution. Clearly the f_s will be solutions of a similar d.e.; furthermore, they will constitute a linearly independent and hence complete system of solutions, since as $l \to 0^+$

$$(14)\qquad f_s(l)/l^{\frac{1}{2}(q-m-1)} \sim k_s(m; q, n)l^{\frac{1}{2}(m-s)(q-s+2)}\,, \qquad (s = 1, \cdots, m)\,,$$

where

$$(15)\qquad k_s(m; q, n) = K(m; q, n)/[K(s-1; q+m-s+1, n-m+s-1) \times K(m-s; q-s, m-s+3)]\,.$$

This is easily proved by writing $x_j = lw_j$ $(j = s, \cdots, m-1)$ in (4) and letting $l \to 0$. We note in addition that (13) is invariant under

$$(16)\qquad q \to n\,, \qquad n \to q\,, \qquad l \to 1-l\,,$$

provided that we also replace $L_{s,r}$ by $(-1)^r L_{s,r}$; this reflects the obvious result that (16) transforms $f_s(l)$ into $f_{m+1-s}(l)$.

3. Solutions of the d.e. It is convenient to introduce $H_r = (1-l)^r L_{s,r}$ $(r = 0, 1, \cdots, m-1)$ and to express (13) as a matrix d.e. for $H = (H_0, \cdots, H_{m-1})'$:

$$(17)\qquad dH/dl = [l^{-1}A + (1-l)^{-1}C]H\,,$$

where

$$(18)\qquad A = \begin{bmatrix} a_0 & & & 0 \\ b_1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & b_{m-1} & a_{m-1} \end{bmatrix}, \qquad C = \begin{bmatrix} c_0 & d_0 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & d_{m-2} \\ 0 & & & c_{m-1} \end{bmatrix}$$

$$a_r = \tfrac{1}{2}r(q-m+r)\,, \qquad b_r = \tfrac{1}{2}(m-r)(q+n-m+r-1)$$
$$c_r = -\tfrac{1}{2}r(n-m+r+2)\,, \qquad d_r = \tfrac{1}{2}(r+1)(r+2)\,.$$

The d.e. (17) is of Fuchsian type, with regular singularities at $l = 0$, 1 and infinity, and we refer to [2], Chapter 4, for the general theory of such d.e.'s.

Assuming a series solution $H = \sum_{r=0}^{\infty} h_r l^{\rho+r}$ in $|l| < 1$, we obtain $Ah_0 = \rho h_0$, so that ρ must be one of the latent roots $a_0, \cdots, a_{m-1}$ of A, and h_0 the corresponding latent vector. To relate this fundamental set of solutions to the $f_s(l)$, we first obtain a non-singular transformation $H = PM$, where $P(m \times m)$ is independent of l, such that $P^{-1}AP = \text{diag}(a_r)$. A suitable choice is

$$P = \{p_{ij}\}, \quad p_{ij} = (-1)^{i-j}\binom{m-j-1}{i-j} \prod_{r=j}^{i-1} (q+n-m+r)/(q-m+j+r+1), \tag{19}$$

with inverse

$$P^{-1} = \{p_{ij}^*\}, \qquad p_{ij}^* = \binom{m-j-1}{i-j} \prod_{r=j}^{i-1} (q+n-m+r)/(q-m+i+r). \tag{20}$$

Both P and P^{-1} are lower triangular, and $M_0 = H_0$. It may be shown that

$$P^{-1}CP = G = \begin{bmatrix} \mu_0, & \nu_0 & & & 0 \\ \lambda_1, & \mu_1 & \nu_1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \nu_{m-2} \\ 0 & & & \lambda_{m-1} & \mu_{m-1} \end{bmatrix} \tag{21}$$

where

$$\begin{aligned} \lambda_i &= (m-i)(n-i)(q+i-1)(q-m+i-1)(q-m+i-2) \\ &\quad \times (q+n-m+i-1) \\ &\quad \div [2(q-m+2i-2)(q-m+2i-1)^2(q-m+2i)], \\ \mu_i &= i^2 - \tfrac{1}{2}i(m+n-3) - m + 1 \\ &\quad + \tfrac{1}{2}i(i+1)(m-i)(n-i)/(q-m+2i-1) \\ &\quad - \tfrac{1}{2}(i+1)(i+2)(m-i-1)(n-i-1)/(q-m+2i+1), \\ \nu_i &= \tfrac{1}{2}(i+1)(i+2). \end{aligned} \tag{22}$$

The d.e. (17) now takes the form

$$dM/dl = [l^{-1}\text{diag}(a_r) + (1-l)^{-1}G]M, \tag{23}$$

and assuming a solution $M = \sum_{r=0}^{\infty} \eta_r l^{a_p+r}$ corresponding to the latent root a_p of A, we obtain the following recurrence relations for the components $(\eta_{0,r}, \cdots, \eta_{m-1,r})$ of the η_r:

$$\begin{aligned} &\eta_{p,0} = 1, \qquad \eta_{i,0} = 0 \qquad (i \neq p), \\ &(r - a_i + a_p)\eta_{i,r} \\ &\quad = \lambda_i \eta_{i-1,r-1} + [\mu_i + (r-1) - a_i + a_p]\eta_{i,r-1} + \nu_i \eta_{i+1,r-1}, \\ &\qquad (i = 0, \cdots, m-1; \quad r = 1, 2, \cdots). \end{aligned} \tag{24}$$

This form of solution unfortunately breaks down if $a_i - a_p$ is a positive integer for some i. In fact, $a_{p+1} - a_p = \frac{1}{2}(q-m+1) + p$ $(p \leqq m-2)$, which is an integer if $q - m$ is odd, while $a_{p+2} - a_p = q - m + 2(p+1)$ $(p \leqq m-3)$ is always an integer. Generally in such situations the solution must be obtained by limiting procedures which may produce logarithmic terms. However, it may be seen from (6) that the $L_{s,r}$ are in fact representable by power series, and it appears that if $a_i - a_p$ is a positive integer for $i > p$, then the right-hand side of the ith equation in (24) vanishes identically when $r = a_i - a_p$. Thus $\eta_{i,r}$ is an undetermined constant introducing the a_i-solution at this stage, and the power series form is preserved.

We also see from (24) that the $\eta_{0,r}$ are zero for $r < p$ in the a_p-solution, while

(25) $$\eta_{0,p} = (p+1)!/\prod_{i=1}^{p}(q-m+p+i+1) = \xi_p\,, \quad \text{say.}$$

Hence $M_0(l) = O(l^{a_p+p})$ as $l \to 0^+$, and since $a_{m-s} + (m-s) = \frac{1}{2}(m-s)(q-s+2)$, it follows from (14) that $L_{0,s}$ must be some linear combination of the $a_{m-s}, a_{m-s+1}, \cdots,$ and a_{m-1}-solutions. The coefficient of the a_{m-s}-solution is clearly $k_s(m; q, n)/\xi_{m-s}$, but the remaining coefficients have not been determined for general s. However, the density function $f_1(l)$ of the largest root corresponds to the largest root a_{m-1} of A, and is thus completely specified by (24). The resulting power series coincides with the result of Sugiyama and Fukutomi [17].

4. Pillai's approximations. A particular solution corresponding to the smallest root $a_0 = 0$ of A may be given explicitly as an $(m-1)$th degree polynomial. Writing $(z)_i = z(z+1)\cdots(z+i-1)$, $(z)_{-i} = z(z-1)\cdots(z-i+1)$, it may be shown that

(26) $$\eta_{i,r} = (-1)^r\binom{m-1}{r}\binom{r}{i}(q)_i(n-1)_{-i}(q+n-m)_r \div [(q-m+1)_{i+r}(q-m+i)_i]\,.$$

Thus we obtain the following approximation to the lower tail of $f_m(l)$ for large q by neglecting the $a_1, \cdots, a_{m-1}$ solutions (i.e., terms of order l^{q-m+1} at least):

(27) $$f_m(l) \approx k(m; q, n)l^{\frac{1}{2}(q-m-1)}(1-l)^{\frac{1}{2}(n-m-1)} \times \sum_{r=0}^{m-1}\binom{m-1}{r}(q+n-m)_r(q-1)_{-(m-1-r)}(-l)^r\,,$$

where

(28) $$k(m; q, n) = \pi^{\frac{1}{2}}\Gamma(\tfrac{1}{2}(q+n-m+1))/[2^{m-1}\Gamma(\tfrac{1}{2}m)\Gamma(\tfrac{1}{2}q)\Gamma(\tfrac{1}{2}n)]\,.$$

Using (16), a corresponding approximation to the upper tail of $f_1(l)$ for large n (near the regular singularity $l = 1$) is obtained. The result is found to be simply the right-hand side of (27) multiplied by $(-1)^{m-1}$.

The integrated form of the approximation was arrived at by Pillai [11] using a different approach, and used in a series of tabulations of the upper 5% and 1% points of l_1. Its accuracy to essentially five places of decimals when $n_2 \geqq m + 11$ was demonstrated at least for $m \leqq 10$ by substituting in explicit expressions for the distribution function [10]. In order to investigate the usefulness of the d.e. (23), some percentage points were calculated by following the a_{m-1}-solution out from the origin, using the same computation procedure as in [4]. The method appeared to be effective at least up to $m = 7$, since on comparing the 1% points, i.e., the less accurate results of the d.e. and the more accurate results of the approximation, these were generally found to differ by no more than a unit in the fifth decimal place. On the other hand, the 5% points obtained from the d.e. tended to exceed Pillai's by about three units in the fifth decimal place. The d.e. approach should be more accurate at lower significance levels, and a tabulation of upper 10% points has been made.

5. Some remarks. The success of the Pillai approximation suggests a similar approach to the other roots, approximating the lower tail of f_s by the a_{m-s}-solution for large q, and deducing a corresponding result for the upper tail of f_{m-s+1} when n is large using (16). No general results corresponding to (27) have been obtained, but it has been found, for instance that when $m = 3$ the distribution of l_2, the median root, is closely approximated by a beta density with parameters $q - 1$ and $n - 1$. Upper 5% and 1% points based on this approximation are identical to five decimal places with those published by Pillai and Dotson [13], except where the latter have employed interpolation.

Differential equations for the latent roots of F and W_I (defined in Section 1) are readily deduced from (13). The approximation used by Hanumara and Thompson [6] corresponds of course to an exact solution of the d.e. in the Wishart case. The Wishart d.e. is in fact closely related to the author's d.e. [5] for the moment generating function of Pillai's $V^{(m)} = \operatorname{tr} B$. If we write $\lambda_{m;q,n} = E\exp(-sV^{(m)})$, then it is easily seen that the density of the largest root of W_I is proportional to $e^{-\frac{1}{2}u}u^{\frac{1}{2}mq-1}\lambda_{m-1;q-1,m+2}(\frac{1}{2}u)$.

Acknowledgment. The author's thanks are due to Mrs. Darien Wei and Dr. R. W. Helms of the Department of Biostatistics, School of Public Health, University of North Carolina, for their computing assistance.

REFERENCES

[1] AL-ANI, S. (1968). On the distribution of the ith latent root under null hypotheses concerning complex multivariate normal populations. Mimeo Series 145, Department of Statistics, Purdue University.

[2] CODDINGTON, E. A. and LEVINSON, N. (1955). *Theory of Ordinary Differential Equations.* McGraw-Hill, New York.

[3] DAVIS, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39** 815–832.

[4] DAVIS, A. W. (1970a). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* **57** 187–191.

[5] DAVIS, A. W. (1970b). On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. Math. Statist.* **41** 1557–1562.

[6] HANUMARA, R. C. and THOMPSON, W. A. (1968). Percentage points of the extreme roots of a Wishart matrix. *Biometrika* **55** 505–512.

[7] HECK, D. L. (1960). Charts of some upper percentage points of the distribution of the largest characteristic root. *Ann. Math. Statist.* **31** 625–642.

[8] KHATRI, C. G. (1964). Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations. *Ann. Math. Statist.* **35** 1807–1810.

[9] KHATRI, C. G. (1969). Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations. *Ann. Inst. Statist. Math.* **21** 23–32.

[10] PILLAI, K. C. S. (1954). On some distribution problems in multivariate analysis. Mimeo Series 88, Institute of Statistics, University of North Carolina.

[11] PILLAI, K. C. S. (1965). On the distribution of the largest characteristic root of a matrix in multivariate analysis. *Biometrika* **52** 405-414.

[12] PILLAI, K. C. S. (1967). Upper percentage points of the largest root of a matrix in multivariate analysis. *Biometrika* **54** 189-194.

[13] PILLAI, K. C. S. and DOTSON, C. O. (1969). Power comparisons of tests of two multivariate hypotheses based on individual characteristic roots. *Ann. Inst. Statist. Math.* **21** 49-66.

[14] ROY, S. N. (1945). The individual sampling distribution of the maximum, the minimum and any intermediate of the p-statistics on the null hypothesis. *Sankhyā* **7** 133-158.

[15] ROY, S. N. (1957). *Some Aspects of Multivariate Analysis*. Wiley, New York.

[16] SUGIYAMA, T. (1967). Distribution of the largest latent root and the smallest latent root of the generalized B statistic and F statistic in multivariate analysis. *Ann. Math. Statist.* **38** 1152-1159.

[17] SUGIYAMA, T. and FUKUTOMI, K. (1967). On the distribution of the extreme characteristic roots of the matrices in multivariate analysis. *Rep. Statist. Appl. Res. Un. Japan. Sci. Eng.* **14** 156-160.

C.S.I.R.O.
DIVISION OF MATHEMATICAL STATISTICS
GLEN OSMOND
SOUTH AUSTRALIA 5064