Some New Tables for Upper Probability Points of the Largest Root of a Determinantal Equation with Seven and Eight Roots

William W. Chen, Internal Revenue Service

We revisit the Fisher-Girshick-Hsu-Roy distribution (1939), which has interested statisticians for more than six decades. Instead of using K.C.S. Pillai's method of neglecting higher order terms of the cumulative distribution function (C.D.F.) of the largest root to approximate the percentage points, we simply keep the whole C.D.F. and apply its natural non-decreasing property to calculate the exact probabilities. At the duplicated percentage points, we found our computed percentage points to be consistent with existing tables. However, our tabulations have greatly extended the existing tables.

In 2002 [1], we were concerned with the distribution of the largest characteristic roots in multivariate analysis when there are two to six roots. Now, we will extend the size to seven and eight roots.Fisher-Girshick-Hsu-Roy(1939) discuss this in detail and present the joint probability density function in general. This well-known distribution depends on the number of characteristic roots and two parameters m and n, which are defined differently for various situations, as described by Pillai (1955). The upper percentage points of the distribution are commonly used in three different multivariate hypothesis tests: tests of equality of the variance-covariance matrices of two p-variate normal populations, tests of equality of the p-dimensional mean vectors for k p-variate normal populations, and tests of independence between a p-set and a q-set of variates in a (p+q)-variate normal population. When the null hypotheses are true, these three proposed tests depend only on the characteristic roots of matrices using observed samples. The problem can be stated as follows: using a random sample from the multivariate normal population, we will compute the characteristic roots from a sum of product matrices of this sample. We will then compare the largest characteristic root of the matrices with the percentage points tabulated in this paper to determine whether or not the null hypothesis is rejected at a certain probability confidence.

There are already many published tables that focus on upper percentage point tabulations or chart the various sizes of roots. The best-known contributor in this area is Pillai, who gave general rules for finding the C.D.F. of the largest root and tabulated upper percentage points of 95% and 99% for various sizes of roots. Other contributors, including Nanda (1948, 1951), Foster and Rees (1957, 1958), and Heck (1960) will be discussed in more detail in section 2. Section 3 contains the joint distribution of s non-null characteristic roots of a matrix in general form and the C.D.F. of the seven and eight largest characteristic roots. The algorithm used to create the tables in this paper is the same as in reference [2], and we will not repeat it. Also, we will ignore the discussion of precision of the results.

Cumulative Function and Historical Work

The joint distribution of s non-null characteristic roots of a matrix in multivariate distribution was first given by Fisher-Girshick-Hsu-Roy (1939) and can be expressed in the form of (3.1). We further extended the distribution of the largest characteristic root to seven and eight roots. Even though the form of the joint density function is known, it is not easy to write out the C.D.F. of the largest characteristic root to seven roots. To solve this problem, two methods can be used to find the C.D.F. more easily. Pillai (1965) suggests that the C.D.F. of the largest characteristic root could be presented in determinant form of incomplete beta functions. Since the numerical integration of each of the s factorial multiple integrals when the determinant is expanded is difficult, he suggests an alternative reduction formula that gives exact expressions for the C.D.F. of the largest root in terms of incomplete beta functions or functions of incomplete beta functions for various values of s. An alternative method suggested by Nanda (1948) yields the same results. He started with the Vandermonde determinant and expanded it in minors of a row, then repeated applied integration by part to find the C.D.F. of the largest characteristic root. In this paper, we use the Pillai notation and present the case with seven roots in equation (3.2). Following this C.D.F. and the algorithm previously used, we tabulate the upper percentage points.

Here, it is useful to review some of the published tables and reasons to extend the tables. Pillai (1956a, 1959) published tables that focus only on two percentage points: 95% and 99% for s =2,6, m = 0(1)4, and n varying from 5 to 1000. Foster and Rees (1957) tabulated the upper percentage points 80%, 85%, 90%, 95%, and 99% of the largest root for s=2, m=-0.5, 0(1)9, n=1(1)19 (5)49,59,79. Foster(1957, 1958) further extended these tables for values of s=3 and 4. Heck(1960) has given some charts of upper 95%, 97.5%, and 99% points for s=2(1)5, m=-0.5, 0(1)10, and n greater than 4. These table values can be applied to our statistical analysis with some standard textbooks as references. For example, recently, Rencher included the percentage point 0.950 in three textbooks [18],[19]).

Without a modern computer, it is difficult and tedious to compute the whole C.D.F.(3.2) at each percentage point. Therefore, deleting higher order terms and retaining a few lower order terms to approximate the roots is a reasonable solution. However, this approach involves intolerable error at lower percentage points, such as 80%,82.5%,85%,87.5%, 90%, or 92.5%. These percentage points are usually ignored due to the difficulty of their computation, and not due to their lack of use. Traditional methods treat intermediate percentage points by interpolation, but without, for example, 85% or 90% percentage points, it is difficult to interpolate 87.5%. In recent years, computers have gradually matured in memory, speed, and flexibility in usage, which has greatly changed the methods by which we study statistics. In this study, we use one of the most basic properties of C.D.F. and revisit this most important distribution. As many percentage points as are needed in one computer run are included: these are 0.80, 0.825, 0.850, 0.875, 0.890, 0.900, 0.910(0.005), 0.995. Different authors have selected different m and n parameter values, but we selected these parameters such that all existing table values are included. For the parameters m=0(1)10 and n=3(1)20(2)30(5)80(10),150,200 (100)1000, our table provides the percentage points and probabilities while avoiding the interpolation problem.

The Distribution Function of Seven Characteristic Roots

Suppose $x = \{x_{ij}\}$ and $x^* = \{x^*_{ij}\}$ are two p-variate random matrices with n_1 and n_2 the degree of freedom, respectively. Assume the two multivariate populations have the same covariance matrix: for example, $S_1 = xx^T / n_1$ and $S_2 = x^* x^{*T} / n_2$. When the null hypothesis is true, both S_1 and S_2 are independent estimators of the unknown but equal covariance matrices. The joint distribution of the roots of the determinantal equation $\left| A - \theta (A + B) \right| = 0$ where $A = n_1 S_1$ and $B = n_2 S_2$ has been given by Fisher-Girshick-Hsu-Roy(1939) and can be written as :

$$f(\theta_1, \ldots, \theta_s) = C(s, m, n) \prod_{i=1}^{s} \theta_i^m (1 - \theta_i)^n \prod_{i>j} (\theta_i - \theta_j)$$

$$(0 < \theta_1 \le \ldots \le \theta_s \le 1),$$

where

$$C(s, m, n) = \frac{\pi^{s/2} \prod_{i=1}^{s} \Gamma(\frac{2m + 2n + s + i + 2}{2})}{\prod_{i=1}^{s} \Gamma(\frac{2m + i + 1}{2}) \Gamma(\frac{2n + i + 1}{2}) \Gamma(\frac{i}{2})}. \quad (3.1)$$

and the parameters m and n are defined differently for various situations as described by Pillai (1955, pp. 118). Following Pillai's method, the cumulative distribution function of the largest characteristic root for seven and eight is given below:

When $s = 7$, the C.D.F. of the largest characteristic root is :

$$\Pr(\theta_7 \le x) = \frac{C(7, m, n)}{m + n + 7} [-IO(x, m + 6, n + 1) * v\_0654321\ x(x, m, n)$$
$$- 2I(x, 2m + 6, 2n + 1) * v\_054321\ x(x, m + 1, n) + 2I(x, 2m + 7, 2n + 1)$$
$$* v\_05432x(x, m, n) - 2I(x, 2m + 8, 2n + 1) * v\_05431x(x, m, n)$$
$$+ 2I(x, 2m + 9, 2n + 1) * v\_05421x(x, m, n) - 2I(x, 2m + 10, 2n + 1)$$
$$* v\_05321x(x, m, n) + 2I(x, 2m + 11, 2n + 1) * v\_054321x(x, m, n)]$$

$$\Pr(\theta_8 \le x) = \frac{C(8, m, n)}{m + n + 8} [-IO(x, m + 7, n + 1) * v\_07654321\ x(x, m, n$$
$$+ 2I(x, 2m + 7, 2n + 1) * v\_0654321\ x(x, m + 1, n) - 2I(x, 2m + 8, 2n$$
$$* v\_065432x(x, m, n) + 2I(x, 2m + 9, 2n + 1) * v\_065431x(x, m, n)$$
$$- 2I(x, 2m + 10, 2n + 1) * v\_065421x(x, m, n) + 2I(x, 2m + 11, 2n + 1)$$
$$* v\_065321x(x, m, n) - 2I(x, 2m + 12, 2n + 1) * v\_064321x(x, m, n)$$
$$+ 2I(x, 2m + 13, 2n + 1) * v\_0654321x(x, m, n)]$$

References

[1] Chen, W.W.S.(2002), Some new tables of the largest root of a matrix in multivariate analysis : a computer approach from 2 to 6, Proceeding of Statistical Computing Section, pp. 503-508.

[2] Chen, W.W.S.(2003), Table for upper percentage points of the largest root of a determinantal equation with five roots, February,http://InterStat.stat.vt.edu/InterStat/

[3] Fisher, R.A.(1939), The sampling distribution of some statistics obtained from non-linear equations, Ann. Eugenics, Vol. 9, pp. 238-249.

[4] Foster, F.G. and Rees, D.H.(1957), Upper percentage points of the generalized beta distribution, I, Biometrika, Vol. 44, pp. 237-247.

[5] Foster, F.G.(1957), Upper percentage points of the generalized beta distribution, II, Biometrika, Vol. 44, pp. 441-453.

[6] Foster, F.G.(1958), Upper percentage points of the generalized beta distribution, III, Biometrika, Vol. 45, pp. 492-503.

[7] Girshick, M.A.(1939), On the sampling theory of the roots of determinantal equations, Ann. Math. Stat., Vol. 10, pp. 203-224.

[8] Heck, D.L.(1960), Charts of some upper percentage points of the

distribution of the largest characteristic root, Ann. Math. Statist., 31, pp. 625-642.

[9] Hsu, P.L. (1939), On the distribution of roots of certain determinantal equations, Ann. Eugenics, Vol. 3, pp. 250-258.

[10] Nanda, D.N.(1948), Distribution of a root of a determinantal equation, Ann. Math. Stat., Vol. 19, pp. 47-57.

[11] Nanda, D.N.(1951), Probability distribution tables of the largest root of a determinantal equation with two roots, J. Indian Soc. Of Agricultural Stat., Vol. 3,, pp. 175-177.

[12] Pillai, K.C.S.(1955), Some new test criteria in multivariate analysis. Ann. Math. Statist. 26, pp. 117-121.

[13] Pillai, K.C.S.(1956a), On the distribution of the largest or smallest root of a matrix in multivariate analysis, Biometrika, Vol. 43, pp. 122-127.

[14] Pillai, K.C.S.(1956b), Some results useful in multivariate analysis, Ann. Math. Statist. Vol. 27, pp. 1106-1114.

[15] Pillai, K.C.S. and Bantegui, C.G.(1959), On the distribution of the largest of Six roots of a matrix in multivariate analysis, Biometrika, Vol. 46, pp. 237-240.

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

[17] Roy, S.N.(1939), p-statistics or some generalizations in analysis of variance appropriate to multivariate problems, Sankhya, Vol. 4, pp. 381-396.

[18]Rencher,A.C.(1998), Multivariate statistical inference and application, John Wiley & Sons, Inc.

[19] Rencher A.C.(2002), Methods of multivariate analysis, John Wiley & Sons, Inc.

$$C(7,m,n)=\frac{\Gamma(2m+2n+14)*\Gamma(2m+2n+12)*\Gamma(2m+2n+10)*\Gamma(2m+2n+8)}{46080*\Gamma(2m+6)*\Gamma(2n+6)*\Gamma(2m+4)*\Gamma(2n+4)*\Gamma(2m+2)*\Gamma(2n+2)*\Gamma(m+4)*\Gamma(n+4)}$$

$$C(8,m,n)=\frac{\Gamma(2m+2n+17)*\Gamma(2m+2n+15)*\Gamma(2m+2n+13)*\Gamma(2m+2n+11)}{8847360*\Gamma(2m+8)*\Gamma(2n+8)*\Gamma(2m+6)*\Gamma(2n+6)*\Gamma(2m+4)*\Gamma(2n+4)*\Gamma(2m+2)*\Gamma(2n+2)}$$

Upper percentage points of .900 of theta(p,m,n),
the largest eigenvalue of |B-theta(W+B)|=0,when s=7

n	m=0	1	2	3	4	5	6	7	8	9	10
3	.9040	.9188	.9295	.9378	.9442	.9495	.9538	.9576	.9608	.9630	.9644
4	.8650	.8842	.8986	.9097	.9186	.9259	.9320	.9371	.9415	.9453	.9490
5	.8266	.8497	.8671	.8809	.8920	.9012	.9090	.9156	.9212	.9261	.9307
6	.7899	.8160	.8362	.8522	.8653	.8763	.8855	.8935	.9004	.9064	.9116
7	.7552	.7838	.8062	.8242	.8391	.8515	.8622	.8714	.8794	.8865	.8927
8	.7226	.7533	.7774	.7971	.8135	.8273	.8392	.8495	.8586	.8665	.8737
9	.6923	.7244	.7501	.7711	.7888	.8038	.8168	.8281	.8380	.8469	.8548
10	.6639	.6973	.7241	.7463	.7650	.7811	.7950	.8072	.8180	.8276	.8363
11	.6376	.6717	.6995	.7226	.7423	.7592	.7740	.7869	.7985	.8088	.8181
12	.6130	.6478	.6763	.7002	.7206	.7382	.7537	.7674	.7796	.7905	.8004
13	.5901	.6253	.6543	.6788	.6999	.7182	.7342	.7485	.7613	.7727	.7832
14	.5687	.6042	.6336	.6586	.6801	.6989	.7155	.7303	.7436	.7556	.7664
15	.5487	.5843	.6140	.6394	.6613	.6806	.6976	.7128	.7266	.7390	.7503
16	.5300	.5656	.5955	.6211	.6434	.6630	.6804	.6960	.7101	.7229	.7346
17	.5124	.5480	.5780	.6038	.6263	.6462	.6640	.6799	.6943	.7074	.7194
18	.4959	.5314	.5614	.5873	.6100	.6302	.6482	.6644	.6791	.6925	.7048
19	.4805	.5157	.5457	.5717	.5945	.6148	.6330	.6495	.6644	.6781	.6906
20	.4659	.5009	.5308	.5568	.5797	.6001	.6185	.6351	.6503	.6642	.6769
22	.4391	.4736	.5032	.5291	.5520	.5726	.5912	.6081	.6236	.6378	.6509
24	.4152	.4490	.4782	.5039	.5268	.5474	.5661	.5832	.5988	.6133	.6267
26	.3937	.4267	.4554	.4809	.5036	.5242	.5429	.5600	.5758	.5904	.6040
28	.3743	.4065	.4347	.4598	.4823	.5027	.5214	.5386	.5544	.5691	.5828
30	.3567	.3881	.4158	.4404	.4627	.4829	.5015	.5186	.5344	.5492	.5629
35	.3190	.3486	.3748	.3984	.4198	.4394	.4576	.4744	.4901	.5047	.5184
40	.2885	.3162	.3410	.3635	.3840	.4030	.4205	.4369	.4523	.4667	.4802
45	.2632	.2894	.3128	.3342	.3538	.3720	.3889	.4048	.4197	.4338	.4471
50	.2420	.2666	.2888	.3092	.3279	.3454	.3617	.3770	.3915	.4052	.4181
55	.2240	.2472	.2683	.2876	.3055	.3223	.3380	.3528	.3668	.3800	.3926
60	.2084	.2304	.2504	.2688	.2860	.3020	.3171	.3314	.3449	.3578	.3700
65	.1949	.2157	.2348	.2524	.2688	.2842	.2987	.3124	.3255	.3379	.3498
70	.1830	.2028	.2210	.2378	.2535	.2683	.2822	.2955	.3081	.3202	.3317
75	.1725	.1914	.2087	.2248	.2399	.2541	.2675	.2803	.2925	.3042	.3154
80	.1631	.1811	.1977	.2131	.2276	.2413	.2542	.2666	.2784	.2897	.3005
90	.1470	.1636	.1788	.1931	.2065	.2192	.2313	.2428	.2538	.2645	.2747
100	.1339	.1492	.1633	.1765	.1889	.2008	.2121	.2229	.2333	.2432	.2529
110	.1229	.1371	.1502	.1625	.1741	.1852	.1958	.2060	.2158	.2252	.2343
120	.1136	.1268	.1390	.1506	.1615	.1719	.1819	.1915	.2007	.2096	.2182
130	.1056	.1179	.1294	.1403	.1506	.1604	.1698	.1788	.1876	.1960	.2042
140	.0986	.1102	.1211	.1313	.1410	.1503	.1592	.1678	.1761	.1841	.1919
150	.0925	.1035	.1137	.1234	.1326	.1414	.1498	.1580	.1659	.1735	.1810
200	.0706	.0792	.0872	.0948	.1021	.1091	.1158	.1223	.1287	.1348	.1408
300	.0480	.0539	.0595	.0648	.0699	.0749	.0796	.0843	.0888	.0932	.0975
400	.0363	.0409	.0451	.0492	.0532	.0570	.0607	.0643	.0678	.0712	.0746
500	.0292	.0329	.0364	.0397	.0429	.0460	.0490	.0519	.0548	.0576	.0604
600	.0244	.0275	.0305	.0332	.0359	.0386	.0411	.0436	.0460	.0484	.0507
700	.0210	.0237	.0262	.0286	.0309	.0332	.0354	.0375	.0396	.0417	.0437
800	.0184	.0208	.0230	.0251	.0271	.0291	.0311	.0330	.0348	.0367	.0384
900	.0164	.0185	.0205	.0224	.0242	.0260	.0277	.0294	.0311	.0327	.0343
1000	.0148	.0167	.0184	.0202	.0218	.0234	.0250	.0265	.0280	.0295	.0309

Upper percentage points of 0.900 of theta(p,m,n),
the largest eigenvalue of |B-theta(W+B)|=0,when s=8

m

n	0	1	2	3	4	5	6	7	8	9
5	0.8517	0.8702	0.8845	0.8959	0.9052	0.9124	0.9225	0.9563	0.9570	0.9882
6	0.8183	0.8396	0.8563	0.8698	0.8808	0.8897	0.8978	0.9014	0.9096	0.9727
7	0.7862	0.8099	0.8286	0.8439	0.8566	0.8667	0.8781	0.8788	0.9019	0.9532
8	0.7558	0.7814	0.8019	0.8187	0.8328	0.8446	0.8559	0.8763	0.8994	0.9279
9	0.7270	0.7542	0.7761	0.7943	0.8096	0.8226	0.8347	0.8428	0.8664	0.8882
10	0.7000	0.7283	0.7515	0.7708	0.7872	0.8012	0.8139	0.8237	0.8385	0.8695
11	0.6744	0.7038	0.7280	0.7482	0.7656	0.7806	0.7937	0.8070	0.8163	0.8390
12	0.6505	0.6807	0.7056	0.7267	0.7448	0.7606	0.7746	0.7874	0.8013	0.8057
13	0.6280	0.6587	0.6843	0.7061	0.7248	0.7412	0.7555	0.7688	0.7826	0.7865
14	0.6069	0.6381	0.6641	0.6864	0.7057	0.7227	0.7376	0.7513	0.7665	0.7798
15	0.5870	0.6185	0.6449	0.6676	0.6874	0.7048	0.7202	0.7351	0.7468	0.7731
16	0.5683	0.5999	0.6267	0.6498	0.6699	0.6877	0.7036	0.7182	0.7341	0.7421
17	0.5507	0.5824	0.6094	0.6327	0.6532	0.6713	0.6874	0.7028	0.7153	0.7347
18	0.5340	0.5658	0.5929	0.6164	0.6371	0.6555	0.6720	0.6866	0.6981	0.7066
19	0.5183	0.5500	0.5772	0.6009	0.6218	0.6404	0.6572	0.6719	0.6841	0.6935
20	0.5035	0.5351	0.5623	0.5861	0.6071	0.6259	0.6428	0.6579	0.6713	0.6799
22	0.4761	0.5074	0.5345	0.5583	0.5796	0.5986	0.6159	0.6313	0.6450	0.6560
24	0.4514	0.4823	0.5092	0.5330	0.5542	0.5734	0.5908	0.6067	0.6205	0.6319
26	0.4291	0.4595	0.4861	0.5097	0.5309	0.5501	0.5676	0.5837	0.5982	0.6101
28	0.4089	0.4387	0.4649	0.4883	0.5094	0.5286	0.5461	0.5622	0.5768	0.5897
30	0.3904	0.4196	0.4454	0.4686	0.4895	0.5085	0.5260	0.5422	0.5569	0.5699
35	0.3507	0.3784	0.4031	0.4254	0.4457	0.4643	0.4816	0.4975	0.5123	0.5258
40	0.3182	0.3444	0.3679	0.3893	0.4089	0.4270	0.4438	0.4595	0.4741	0.4874
45	0.2912	0.3160	0.3383	0.3588	0.3776	0.3951	0.4114	0.4267	0.4410	0.4542
50	0.2684	0.2918	0.3131	0.3327	0.3507	0.3676	0.3833	0.3982	0.4121	0.4250
55	0.2488	0.2711	0.2913	0.3100	0.3274	0.3436	0.3588	0.3731	0.3866	0.3994
60	0.2319	0.2531	0.2724	0.2903	0.3069	0.3225	0.3371	0.3510	0.3641	0.3766
65	0.2172	0.2373	0.2557	0.2728	0.2888	0.3038	0.3179	0.3313	0.3441	0.3562
70	0.2042	0.2234	0.2410	0.2574	0.2727	0.2871	0.3008	0.3137	0.3261	0.3378
75	0.1926	0.2110	0.2278	0.2435	0.2583	0.2722	0.2854	0.2979	0.3098	0.3212
80	0.1824	0.1999	0.2160	0.2311	0.2453	0.2587	0.2715	0.2836	0.2952	0.3062
90	0.1647	0.1808	0.1958	0.2097	0.2229	0.2354	0.2473	0.2587	0.2696	0.2800
100	0.1502	0.1651	0.1790	0.1920	0.2043	0.2159	0.2271	0.2378	0.2480	0.2578
110	0.1380	0.1519	0.1648	0.1770	0.1885	0.1994	0.2099	0.2200	0.2296	0.2389
120	0.1277	0.1406	0.1527	0.1641	0.1749	0.1852	0.1951	0.2046	0.2138	0.2226
130	0.1188	0.1309	0.1423	0.1530	0.1632	0.1730	0.1823	0.1913	0.2000	0.2084
140	0.1110	0.1225	0.1332	0.1433	0.1530	0.1622	0.1711	0.1796	0.1879	0.1958
150	0.1042	0.1150	0.1252	0.1348	0.1439	0.1527	0.1611	0.1692	0.1771	0.1847
200	0.0798	0.0882	0.0962	0.1038	0.1111	0.1181	0.1248	0.1313	0.1377	0.1438
300	0.0543	0.0602	0.0658	0.0711	0.0762	0.0812	0.0860	0.0907	0.0952	0.0997
400	0.0412	0.0457	0.0500	0.0541	0.0580	0.0619	0.0656	0.0692	0.0728	0.0762
500	0.0331	0.0368	0.0403	0.0436	0.0469	0.0500	0.0530	0.0560	0.0589	0.0618
600	0.0277	0.0308	0.0338	0.0366	0.0393	0.0419	0.0445	0.0470	0.0495	0.0519
700	0.0238	0.0265	0.0290	0.0315	0.0338	0.0361	0.0383	0.0405	0.0426	0.0447
800	0.0209	0.0232	0.0255	0.0276	0.0297	0.0317	0.0337	0.0356	0.0375	0.0393
900	0.0186	0.0207	0.0227	0.0246	0.0265	0.0283	0.0300	0.0317	0.0334	0.0351
1000	0.0168	0.0187	0.0205	0.0222	0.0239	0.0255	0.0271	0.0286	0.0301	0.0316