On the Non-Central Distribution of the Sum of Two Characteristic Roots and the Monotonicity of the Power Function of Associated Tests

by

K. C. Sreedharan Pillai

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series Number 8

July 1963

On the Non-Central Distribution of the Sum of Two Characteristic Roots and the Monotonicity of the Power Function of Associated Tests[1]

By K. C. Sreedharan Pillai
Purdue University

1. Introduction and Summary. Let A_1 and A_2 be two positive definite matrices of order p, A_1 having a Wishart distribution [3,7] with f_1 degrees of freedom and A_2 having an independent non-central Wishart distribution with f_2 degrees of freedom, corresponding to the linear case [1,2]. Now transform

$$A_1 = C\,L\,C'$$

where C is a lower triangular matrix such that

$$A_1 + A_2 = C\,C' \, .$$

It has been shown [4] that the density function of L is given by

$$(1.1)\quad f(L) = Ke^{-\lambda^2/2}\, {}_1F_1\left\{\tfrac{1}{2}(f_1+f_2), \tfrac{1}{2}f_2, \tfrac{1}{2}\lambda^2(1-\ell_{11})\right\} |L|^{(f_1-p-1)/2} |I-L|^{(f_2-p-1)/2}$$

where

$$K = \pi^{-p(p-1)/4} \prod_{i=1}^{p} \Gamma\left[\tfrac{1}{2}(f_1+f_2+1-i)\right] \Big/ \left\{ \Gamma\left[\tfrac{1}{2}(f_1+1-i)\right] \Gamma\left[\tfrac{1}{2}(f_2+1-i)\right]\right\} ,$$

[1]Research supported by XL Grant from the Purdue Research Foundation.

λ^2 is the single non-centrality parameter in the linear case, ℓ_{11} is the element in the top left corner of the L matrix, and ${}_1F_1$ denotes the confluent hypergeometric function.

In this paper, the first four moments of the trace of L have been obtained for $p = 2$, and approximation to the distribution of the trace of L and that of I-L have been suggested. In addition, the latter approximate distribution is used to show the monotonic character of the power of the test using the trace criterion (of I-L) where the null hypothesis equates λ to zero and the upper tail end of the distribution is taken as the critical region. Similar results are obtained for the case of a single non-null population cononical correlation.

2. Moments of the trace of L when $p = 2$. Let

$$L = T\,T'$$

where T is a lower triangular matrix $[t_{ij}]$. It has been shown [4] that then the diagonal elements t_{ii} are independently distributed and that $t_{ii}^2 (i=2,3,\ldots,p)$ follows the distribution

$$f(t_{ii}^2) = \frac{(t_{ii}^2)^{\frac{1}{2}(f_1+1-i)-1}(1-t_{ii}^2)^{\frac{1}{2}f_2-1}}{\beta\left\{\frac{1}{2}(f_1+1-i),\ \frac{1}{2}f_2\right\}} \qquad (2.1)$$

$$(0 \le t_{ii}^2 \le 1)\ ,$$

while t_{11}^2 is distributed as

$$(2.2)\quad f(t_{11}^2) = \frac{e^{-\lambda^2/2}(t_{11}^2)^{\frac{f_1}{2}-1}(1-t_{11}^2)^{\frac{f_2}{2}-1}\,{}_1F_1\left\{\frac{1}{2}(f_1+f_2), \frac{1}{2}f_2, \frac{1}{2}\lambda^2 \ (1-t_{11}^2)\right\}}{\beta(\frac{1}{2}f_1, \frac{1}{2}f_2)}$$

$$(0 \leq t_{11}^2 \leq 1) .$$

Now, if $p = 2$, it can be shown that

$$(2.3)\qquad f(\ell_{11}, \ell_{22}, \ell_{21}) = f(u_1)\, f(u_2)\, f(u_3)$$

where

$$(2.4)\qquad u_1 = t_{11}^2,\ u_2 = t_{22}^2 \text{ and } u_3 = t_{21}^2 / \left\{(1-t_{11}^2)(1-t_{22}^2)\right\},$$

$f(u_1)$ is given by (2.2), $f(u_2)$ by (2.1) with $i = 2$,
and

$$(2.5)\qquad f(u_3) = \frac{u_3^{\frac{1}{2}-1}(1-u_3)^{\frac{1}{2}(f_2-1)-1}}{\beta\left\{\frac{1}{2}, \frac{1}{2}(f_2-1)\right\}} \qquad (0 \leq u_3 \leq 1)$$

Thus, from (2.3) it may be seen that u_1, u_2 and u_3 are independently distributed. Now

$$\ell_{11} + \ell_{22} = t_{11}^2 + t_{22}^2 + t_{21}^2$$

and, therefore, using (2.4)

(2.6) $$\ell_{11} + \ell_{22} = u_1 + u_2 + u_3(1-u_1)(1-u_2) .$$

Thus the moments of the trace of L when $p = 2$ can be obtained using the relation (2.6) and the density functions $f(u_1)$, $f(u_2)$ and $f(u_3)$ given above, remembering that u_1, u_2 and u_3 are independently distributed.

Now denote $\ell_{11} + \ell_{22}$ by $W^{(2)}$. The first four moments of $W^{(2)}$ will be given by

(2.7) $$\mu_1'(W^{(2)}) = \{2f_1 e^{-\lambda^2/2}/(\nu-1)\} \sum_{i=o}^{\infty} a_i(\tfrac{1}{2}\lambda^2)^i/i!$$

where

(2.8) $$a_i = (\nu+i-1)/g_i$$

$$\nu = (f_1+f_2) \quad \text{and} \quad g_i = \nu + 2i .$$

(2.9) $$\mu_2'(W^{(2)}) = \{4f_1 e^{-\lambda^2/2}/(\nu^2-1)\} \sum_{i=o}^{\infty} b_i(\tfrac{1}{2}\lambda^2)^i/i!$$

where

(2.10) $$b_i = \{f_1\nu^2 + 2(i+1)f_1^2 + (i^2+3i-1)f_1 + (2i+3)f_1f_2 + f_2^2 + (2i-1)f_2 + 2(i^2-1)\}/e_0$$

where $e_0 = g_i(g_i + 2)$.

(2.11) $$\mu_3'(W^{(2)}) = \left[8f_1 e^{-\lambda^2/2} / \left\{(\nu^2-1)(\nu+3)\right\}\right] \sum_{i=0}^{\infty} c_i (\tfrac{1}{2}\lambda^2)^i / i!$$

where

(2.12) $$c_i = e_1/e_2$$

and where

$$\begin{aligned} e_1 = {} & f_1^2 \nu^3 + (3i+9)f_1^4 + (6i+21)f_1^3 f_2 + (3i+15)f_1^2 f_2^2 + (3i^2+21i+25)f_1^3 \\ & + (3i^2+30i+41)f_1^2 f_2 + (i^3+18i^2+44i+15)f_1^2 + (9i+18)f_1 f_2^2 \\ & + 3f_1 f_2^3 + 2f_2^3 + (12i^2+39i+9)f_1 f_2 + 6if_2^2 \\ & + (6i^3+30i^2+18i-26)f_1 + (12i^2+6i-26)f_2 + 8i^3 + 12i^2 - 20i - 24 \end{aligned}$$

and

$$e_2 = g_i(g_i+2)(g_i+4) .$$

(2.13) $$\mu_4'(W^{(2)}) = \left[f_1 e^{-\lambda^2/2} / \left\{(\nu^2-1)(\nu+3)(\nu+5)\right\}\right] \sum_{i=0}^{\infty} d_i (\tfrac{1}{2}\lambda^2)^i / i!$$

where

(2.14) $$d_i = e_3/e_4$$

and where

$$e_3 = (\nu+5)[(f_1+2)(f_1+4)(\nu+1)(\nu+3)(\nu f_1+4f_1g_i+23f_1+6\nu-4g_i-30)$$

$$+ 4(f_1+2)(\nu+3)h_i(f_1\nu+3f_1g_i+19f_1+4\nu-3g_i-14)$$

$$+ 2(f_1^2-1)(g_i+4)(g_i+6)(2f_1g_i+3f_1\nu+13f_1+6\nu+6g_i+30)$$

$$+12(f_1-1)(g_i+6)\quad(f_1g_i+4f_1+g_i+3h_i+10)$$

$$+ 6h_i(h_i+2)(3f_1\nu+9f_1+6\nu+10h_i+58)]$$

$$+ (g_i+5)[(f_1-1)(g_i+6)\{(f_1(g_i+5)+h_i+4)(f_1(g_i+5)+12h_i+6)+45h_i(h_i+2)\}$$

$$+ 105h_i(h_i+2)(h_i+4)]$$

and

$$e_4 = g_i(g_i+2)(g_i+4)(g_i+6)$$

and where

$$h_i = f_2 + 2i \ .$$

It may be observed that, if we denote the trace of I-L for $p = 2$ by $V^{(2)}$, the moments of $V^{(2)}$ can be obtained from those of $W^{(2)}$ using the relation $V^{(2)} = 2-W^{(2)}$, which again in terms of the u's equals $(1-u_1) + (1-u_2) - u_3(1-u_1)(1-u_2)$.

3. Approximations to the distributions of $W^{(2)}$ and $V^{(2)}$. On the basis of the moments presented in the preceding section, the following approximate distribution for $W^{(2)}$ is suggested.

$$(3.1)\quad f(W^{(2)}) = \frac{(W^{(2)})^{\frac{1}{2}f_1-1}(1-W^{(2)}/2)^{\frac{1}{2}f_2-1}\left[1+e^{-\frac{1}{2}\lambda^2}\,{}_1F_1\left\{\frac{1}{2}\nu\,,\ \frac{1}{2}f_2, \frac{1}{2}\lambda^2(1-W^{(2)}/2)\right\}\right]}{\beta(\frac{1}{2}f_1,\frac{1}{2}f_2)2^{\frac{1}{2}(f_1+2)}}$$

$$(0 \le W^{(2)} \le 2)\ .$$

It may be pointed out that the approximation has been obtained such that the difference between the coefficients of $(\frac{1}{2}\lambda^2)^i/i$ in the series for the exact and approximate first moments is

$$(3.2)\qquad 2\,i\,f_1/\{(\nu-1)\nu\,(\nu+2i)\}$$

while the exact coefficient itself can be written in the form

$$\{2f_1/(\nu+2i)\} + 2if_1/\{(\nu-1)(\nu+2i)\}$$

which shows that the difference (3.2) is negligible when f_1 or f_2 is large. In fact, the coefficient of $e^{-\frac{1}{2}\lambda^2}(\frac{1}{2}\lambda^2)^i/i!$ in the series for the exact and approximate moments can be shown to tend to

$$(2f_1)^j/\{g_i(g_i+2)\ldots(g_i+2(j-1))\}$$

for the j^{th} $(j=1,2,3,4)$ moment for large values of f_1 or f_2.

Now transform $V^{(2)} = 2-W^{(2)}$, we get from (3.1)

$$(3.3)\quad f(V^{(2)}) = \frac{(V^{(2)})^{\frac{1}{2}f_2-1}(1-V^{(2)}/2)^{\frac{1}{2}f_1-1}\left[1+e^{-\frac{1}{2}\lambda^2}{}_1F_1\left\{\frac{1}{2}\nu\ ,\ \frac{1}{2}f_2,\ \frac{1}{2}\lambda^2\ \frac{1}{2}V^{(2)}\right\}\right]}{\beta(\frac{1}{2}f_2,\ \frac{1}{2}f_1)\ 2^{\frac{1}{2}(f_2+2)}}$$

$$(0 \leq V^{(2)} \leq 2)\ .$$

4. Monotonicity of the power of a test based on $V^{(2)}$. Now define a test as follows:

$$H_0:\ \lambda = 0$$
$$H_1:\ \lambda > 0,$$

reject H_0 if $V^{(2)} \geq V_0$, where V_0 is such that $\int_{V_0}^{2} f(V^{(2)}/\lambda=0)dV^{(2)} = \alpha$,
accept H_0 if $V^{(2)} < V_0$.

To show that the power function of the test is monotonic increasing, differentiate with respect to λ

$$I(\lambda) = \int_{V_0}^{2} f(V^{(2)}/\lambda)dV^{(2)}$$

or since λ is positive, with respect to $\gamma = \frac{1}{2}\lambda^2$. Now put $x = V^{(2)}/2$, then

$$(4.1)\quad \frac{dI(\gamma)}{d\gamma} = \frac{d}{d\gamma}\int_{V_0/2}^{1} f(x/\gamma)dx$$

$$= (\tfrac{1}{2}e^{-\gamma})\left[-\int_{V_0/2}^{1}\frac{x^{\frac{1}{2}f_2-1}(1-x)^{\frac{1}{2}f_1-1}{}_1F_1\left\{\frac{1}{2}\nu\ ,\ \frac{1}{2}f_2,\ \gamma x\right\}}{\beta(\frac{1}{2}f_2,\ \frac{1}{2}f_1)}\,dx\ \cdots\right.$$

$$\left.+\int_{V_0/2}^{1}\frac{x^{\frac{1}{2}(f_2+2)-1}(1-x)^{\frac{1}{2}f_1-1}{}_1F_1\left\{\frac{1}{2}(\nu+2),\ \frac{1}{2}(f_2+2),\gamma x\right\}}{\beta(\frac{1}{2}(f_2+2),\ \frac{1}{2}f_1)}\,dx.\right]$$

Now it can be shown (see Appendix) that

$$(4.2) \qquad \int_k^1 \frac{y^p(1-y)^{q-1}}{\beta(p+1, q)}\, dy - \int_k^1 \frac{y^{p-1}(1-y)^{q-1}}{\beta(p,q)}\, dy > 0$$

unless $k=0$ or 1 in which case obviously the left side of (4.2) equals zero.

Now consider the coefficient of $\frac{e^{-\gamma}}{2} \gamma^i/i!$ on the right side of (4.1). An application of (4.2) to this coefficient will show immediately that it is positive and hence $\frac{dI(\gamma)}{d\gamma} > 0$. This proves the monotonic character of the power of the test defined above.

5. Canonical correlation. In the case of relation between a p-set of variates, $x' = (x_1,\ldots,x_p)$, and a q-set, $y' = (y_1,\ldots,y_q)$, from a (p+q)-variate normal population, where there is only one non-null population canonical correlation coefficient, ρ_1, and $p \leqq q$, $(p+q) < n$ where n is the sample size,

$$(5.1.) \qquad \lambda^2 = \rho_1^2 \sum_{t=1}^{\nu} y_{1t}^2/(1-\rho_1^2)$$

where $y_{1t}(t=1,\ldots,\nu)$ are related to the sample observations of y_1, and y, here, is considered fixed [4]. In this situation, the approximation to the distribution of the sum of squares of two sample canonical correlation coefficients is given by (3.3), where $f_2 = q$ and $f_1 = k-q-1$ such that $f_1+f_2=\nu$.

If, however, y is not fixed, then $\sum_{t=1}^{\nu} y_{1t}^2$ in λ^2 of (5.1) is a chi-square with ν degrees of freedom and, therefore, the density function of u_1 will be given by [4]

$$(5.2)\qquad f(u_1) = \frac{(1-\rho_1^2)^{\frac{1}{2}\nu}\, u_1^{\frac{1}{2}f_1-1}(1-u_1)^{\frac{1}{2}f_2-1}\, {}_2F_1\left\{\frac{1}{2}\nu,\ \frac{1}{2}\nu,\ \frac{1}{2}f_2,\quad \rho_1^2(1-u_1)\right\}}{\beta(\frac{1}{2}f_1,\ \frac{1}{2}f_2)} .$$

$$(0 \le u_1 \le 1)$$

where ${}_2F_1$ is the hypergeometric function.

In this case, the first moment of $W^{(2)}$ can be obtained from the right side of (2.7) by replacing $e^{-\frac{1}{2}\lambda^2}$ by $(1-\rho_1^2)$, $(\lambda^2)^i$ by $(\rho_1^2)^i$ and a_i by $\nu \cdots\cdot(\nu+2(i-1))a_i$.

Similarly the second moment of $W^{(2)}$ in this case is obtained from (2.9) by the same changes as above except that instead of a_i the appropriate coefficient, i.e. b_i should be substituted. The third and fourth moments are obtained in a similar manner from (2.11) and (2.13) respectively.

Further, an approximation to the distribution of $W^{(2)}$ will be given by

$$(5.3)\qquad f(W^{(2)}) = \frac{(W^{(2)})^{\frac{1}{2}f_1-1}(1-W^{(2)}/2)^{\frac{1}{2}f_2-1}\left[1+(1-\rho_1^2)^{\frac{1}{2}\nu}\, {}_2F_1\left\{\frac{1}{2}\nu,\ \frac{1}{2}\nu,\frac{1}{2}f_2,\rho_1^2(1-W^{(2)}/2)\right\}\right]}{\beta(\frac{1}{2}f_1,\ \frac{1}{2}f_2)2^{\frac{1}{2}(f_1+2)}}$$

$$(0 \le W^{(2)} \le 2) .$$

Again, transformation in (5.3) of $V^{(2)} = 2 - W^{(2)}$ gives

$$(5.4)\quad f(V^{(2)}) = \frac{(V^{(2)})^{\frac{1}{2}f_2-1}\,(1-V^{(2)}/2)^{\frac{1}{2}f_1-1}\left[1-(1-\rho_1^2)^{\frac{1}{2}\nu}\,{}_2F_1\left\{\tfrac{1}{2}\nu,\ \tfrac{1}{2}\nu,\ \tfrac{1}{2}f_2,\ \rho_1^2 V^{(2)}/2\right\}\right]}{\beta(\frac{1}{2}f_2,\ \frac{1}{2}f_1)\ 2^{\frac{1}{2}(f_2+2)}}$$

$$(0 \leq V^{(2)} \leq 2)\ .$$

In the null case (i.e., when $(\lambda=0$ or $\rho_1=0))$, (3.1),(3.3),(5.3) and (5.4) reduce to beta distributions. However, the approximate distributions suggested by Pillai under the null hypotheses [5,6] in the respective cases have degrees of freedom $2f_1$ and $2f_2$ in place of f_1 and f_2 respectively obtained here. In the null case the approximations seem to be better with $2f_1$ and $2f_2$ degrees of freedom.

Now consider the test of section 4 replacing λ by ρ_1. In order to show that the power function is monotonic increasing, follow the method used in the preceding section with (5.4) instead of (3.3) and we get

$$(5.5)\quad \frac{dI(\gamma)}{d\gamma} = \frac{1}{4}\nu(1-\gamma)^{\frac{1}{2}\nu-1}\Big[-\int_{V_0/2}^{1}\frac{x^{\frac{1}{2}f_2-1}(1-x)^{\frac{1}{2}f_1-1}\,{}_2F_1\{\frac{1}{2}\nu,\frac{1}{2}\nu,\frac{1}{2}f_2,\gamma x\}}{\beta(\frac{1}{2}f_2,\ \frac{1}{2}f_1)}\,dx$$

$$+(1-\gamma)\int_{V_0/2}^{1}\frac{x^{\frac{1}{2}(f_2+2)-1}(1-x)^{\frac{1}{2}f_1-1}\,{}_2F_1\{\frac{1}{2}(\nu+2),\frac{1}{2}(\nu+2),\frac{1}{2}(f_2+2),\gamma x\}}{\beta(\frac{1}{2}(f_2+2),\ \frac{1}{2}f_1)}\,dx$$

Here γ equals ρ_1^2 and x, as before, equals $V^{(2)}/2$. Now, the coefficient of γ^i within the brackets in (5.5) is the integral with respect to x between the limits $V_0/2$ to 1, of

$$(5.6)\quad \left\{\nu(\nu+2)\ \ldots\ (\nu+2(i-1))/2^i\right\}\left\{\frac{(\nu+2i)x^{\frac{1}{2}(f_2+2(i+1))-1}(1-x)^{\frac{1}{2}f_1-1}}{\beta\left\{\frac{1}{2}(f_2+2(i+1)),\ \frac{1}{2}f_1\right\}} - \frac{(\nu+2)x^{\frac{1}{2}(f_2+2i)-1}(1-x)^{\frac{1}{2}f_1-1}}{\beta\left\{\frac{1}{2}(f_2+2i),\frac{1}{2}f_1\right\}}\right\}$$

$$(i = 1,\ 2,\ \ldots\ldots)$$

and when $i = 0$, of

$$\frac{x^{\frac{1}{2}(f_2+2)-1}(1-x)^{\frac{1}{2}f_1-1}}{\beta(\frac{1}{2}(f_2+2),\ \frac{1}{2}f_1)} - \frac{x^{\frac{1}{2}f_2-1}(1-x)^{\frac{1}{2}f_1-1}}{\beta(\frac{1}{2}f_2,\ \frac{1}{2}f_1)}. \tag{5.7}$$

By virtue of (4.2), expresssions in (5.6) and (5.7) can at once be shown to be positive thus proving the monotonicity of the power function of the test.

The method used above for showing the monotonicity of the power of the test can be applied to similar tests with multiple correlation coefficient.

REFERENCES

[1] T. W. Anderson and M. A. Girschick, "Some extensions of the Wishart distribution," Ann. Math. Stat., Vol. 15(1944) pp. 345-357.

[2] T. W. Anderson, "The non-central Wishart distribution and certain problems of multivariate statistics," Ann. Math. Stat., Vol. 17(1946) pp. 409-431.

[3] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York, 1958.

[4] A. M. Kshirsagar, " The non-central multivariate beta distribution," Ann. Math. Stat., Vol. 32(1961) pp. 104-111.

[5] K. C. Sreedharan Pillai, "Some new test criteria in multivariate analysis," Ann. Math. Stat., Vol. 26, (1955) pp. 117-121.

[6] K. C. Sreedharan Pillai, Statistical Tables for Tests of Multivariate Hypotheses, The Statistical Center, Manila, 1960.

[7] S. N. Roy, Some Aspects of Multivariate Analysis, John Wiley and Sons, New York, 1957.

Appendix

Two results on incomplete beta functions are obtained below:

1. $$\int_k^1 \frac{x^p(1-x)^{q-1}}{\beta(p+1,q)}\,dx - \int_k^1 \frac{x^{p-1}(1-x)^{q-1}}{\beta(p,q)}\,dx > 0 .$$

except when $k = 0$ or 1, in which case the left side equals zero.

To show this, expand $(1-x)^{q-1}$ in the above integrals and integrate term by term. We get the difference of the two integrals to be

$$(A.1)\qquad \begin{aligned} &- \frac{k^p}{\beta(p+1,q)}\left[\frac{k}{p+1} - \frac{(q-1)k^2}{p+2} + \frac{(q-1)(q-2)}{2!}\,\frac{k^3}{p+3} - \dots\right] \\ &+ \frac{k^p}{\beta(p,q)}\left[\frac{1}{p} - \frac{(q-1)k}{p+1} + \frac{(q-1)(q-2)}{2!(p+2)}\,k^2 - \dots\right] . \end{aligned}$$

Now combining coefficients of like powers of k and taking the sum

(A.1) reduces to $\dfrac{k^p(1-k)^q}{p\cdot\beta(p,q)} > 0$ unless $k = 0$ or 1.

2. $$\int_k^1 \frac{x^{p-1}(1-x)^{q-1}}{\beta(p,q)}\,dx - \int_k^1 \frac{x^{p-1}(1-x)^{q}}{\beta(p,q+1)}\,dx > 0$$

except when $k = 0$ or 1., in which case the left side equals zero. As before, we can show that the difference between the two integrals above equals

$\dfrac{k^p(1-k)^q}{q\cdot\beta(p,q)} > 0$ unless $k = 0$ or 1.