THE EXACT DISTRIBUTION OF THE WALTZ STATISTIC:

THE NON CENTRAL CASE

P. C. B. Phillips and S. Ouliaris

November 27, 1984
THE EXACT DISTRIBUTION OF THE WALD STATISTIC:
THE NON CENTRAL CASE

P. C. B. Phillips* and S. Ouliaris
Cowles Foundation for Research in Economics
Yale University

0. ABSTRACT

This paper extends earlier results, which were reported in [7], to include non null distributions. As in [7], attention is concentrated on the Wald statistic for testing general linear restrictions on the coefficients in the multivariate linear model. The results of the present paper encompass the null distributions derived in [7] and generalize all previously known results for such statistics as the standard regression $F$ test and Hotelling's $T^2$ test.

*Our thanks go to Glena Ames for her skill and effort in typing the manuscript of the paper. Phillips' research was supported by the NSF under Grant No. SES 8218792.
1. INTRODUCTION

In an earlier paper [7] Phillips gave a general formula for the distribution of the Wald statistic for testing general linear restrictions in the multivariate linear model. This formula was sufficiently comprehensive to include all previously known null distributions for special cases of the Wald statistic such as the regression $F$ and Hotelling's $T^2$ and generalized $T^0_n$ tests.

The present paper extends the analysis and the results of [7] to include the general case of non null distributions. The formulae presented herein therefore cover all earlier results including those of [7]. Specializations of our formulae to the commonly occurring cases of the non-central regression $F$ and Hotelling's $T^2$ are given in detail.

Conventional classical assumptions of normally distributed errors and nonrandom exogenous variables are employed.

2. THE MODEL AND NOTATION

We use the same notation as [7] for the multivariate linear model:

(1) \[ y_t = Ax_t + u_t , \quad t = 1, \ldots, T \]

$y_t$ is a vector of $n$ dependent variables, $A$ is an $n \times m$ matrix of parameters, $x_t$ is a vector of nonrandom independent variables and the $u_t$ are i.i. $N(0,\Omega)$ errors with $\Omega$ positive definite. The hypothesis under consideration takes the general form

(2) \[ H_0 : \text{D vec } A = d , \quad H_1 : \text{D vec } A - d = b \neq 0 \]
where $D$ is a $q \times nm$ matrix of known constants of rank $q$, $d$ is a known vector and $\text{vec}(A)$ stacks the rows of $A$.

From least squares estimation of (1) we have:

$$ (3) \quad A^* = Y'X(X'X)^{-1}, \quad \Omega^* = Y'(I - P_X)Y/N $$

where $Y' = [y_1, \ldots, y_T]$, $X' = [x_1, \ldots, x_T]$, $P_X = X(X'X)^{-1}X'$ and $N = T - m$. We take $X$ to be a matrix of full rank $m \leq T$ and define $M = (X'X)^{-1}$.

The Wald statistic for testing the hypothesis (2) is

$$ (4) \quad w = (D \text{ vec } A^* - d)'\{D(\Omega^* \otimes M)D'\}^{-1}(D \text{ vec } A^* - d) $$

$$ = N\ell' B \ell $$

where $\ell = D \text{ vec } A^* - d$ is $N(b, V)$ under $H_1$ with $V = D(\Omega \otimes M)D'$, and $B = \{D(C \otimes M)D'\}^{-1}$, where $C = Y'(I - P_X)Y$ is central Wishart with covariance matrix $\Omega$ and $N$ degrees of freedom.

We define $y = \ell'B\ell$ and write $y$ in canonical form as

$$ (5) \quad y = g'Gg $$

where $g = V^{-1/2}\ell$ is $N(m, I_q)$, $m = V^{-1/2}b$ and $G = V^{-1/2}(D(C \otimes M)D')V^{-1/2} = D(C \otimes M)D'$, say.

### 3. The Noncentral Distribution of $W$

We start with the canonical variate $y$ as given in (5). The conditional distribution of $y$ given $C$ is that of a noncentral positive definite quadratic form in the normally distributed random vector $g$.

We define $z = G^{1/2}g$ and then $y = z'z$ and
(6) \( \text{pdf}(z|C) = (2\pi)^{-q/2}(\det G)^{-1/2} \exp(-m'm/2) \exp(-z'G^{-1}z/2) \exp(G^{-1/2}zm') \). 

Note that \( y \) is invariant under \( z \to zk \) where \( k \in O(1) \) (i.e., \( k^2 = 1 \)). Making this substitution in (6) and integrating over the (normalized) orthogonal group \( O(1) \) we have:

(7) \( (2\pi)^{-q/2}(\det G)^{-1/2} \exp(-m'm/2) \exp(-z'G^{-1}z/2) 0_F^1 \left( \frac{1}{2}, \frac{1}{4} z'G^{-1/2}mm'G^{-1/2}z \right) \).

We now transform \( z \to (h,y) \) according to the decomposition \( z = hy^{1/2} \) where \( y = z'z \) and \( h \in V_{1,q} \) (that is, the unit sphere \( h'h = 1 \)). The measure changes according to

\[ dz = \frac{1}{2} y^{q/2 - 1} dy(dh) \]

where \((dh)\) denotes the (unnormalized) Haar measure on the Stiefel manifold \( V_{1,q} \). It follows from this decomposition that the required density of \( y \) conditional on \( C \) is:

\[ \text{pdf}(y|C) = 2^{-q/2 - 1} \pi^{-q/2} \exp(-m'm/2)y^{q/2 - 1} \]

\[ \cdot (\det G)^{-1/2} \int_{V_{1,q}} \text{etr}(-yG^{-1}hh'/2) 0_F^1 \left( \frac{1}{2}, \frac{1}{4} y'h'G^{-1/2}mm'G^{-1/2}h \right) (dh) \]

(8) \[ = 2^{-q/2} [\gamma(q/2)]^{-1} \exp(-m'm/2)y^{q/2 - 1}(\det G)^{-1/2} \]

\[ \cdot \int_{V_{1,q}} \text{etr}(-yG^{-1}hh'/2) 0_F^1 \left( \frac{1}{2}, \frac{1}{4} y'h'G^{-1/2}mm'G^{-1/2}h \right) (dh) \]

where \((dh)\) denotes the normalized measure on \( V_{1,q} \) (that is, \( \int_{V_{1,q}} (dh) = 1 \)).

Series representations of the factors in the integrand of (8) are as
follows:

\[ (9) \quad \text{etr}(-yG^{-1}hh'/2) = \sum_{j=0}^{\infty} \frac{(-y/2)^j}{j!} C^{(j)}(G^{-1}hh') \]

\[ (10) \quad 0 \frac{1}{4} h'G^{-1/2}hh'G^{-1/2}h = \sum_{k=0}^{\infty} \frac{(y/4)^k}{k!(1/2)_k} C^{(k)}(G^{-1/2}mm'G^{-1/2}hh') \]

in terms of top order zonal polynomials \( C^{(j)}() \) where \( (j) \) denotes the partition \( (j, 0, \ldots, 0) \) of \( j \) with only one nonzero part. Formulae for \( C^{(j)}() \) are given in [4].

We substitute (9) and (10) into (8) and integrate term by term, which is permissible in view of the absolute and uniform convergence of the series. The integral

\[ \int_{V_1, q} C^{(j)}(G^{-1}hh')C^{(k)}(G^{-1/2}mm'G^{-1/2}hh')(dh) \]

\[ = \int_{0(q)} C^{(j)}(G^{-1}HE_{11}H')C^{(k)}(G^{-1/2}mm'G^{-1/2}HE_{11}H')(dH) \]

\[ = \sum_{\varphi \in (j), (k)} C^{(j)}(G^{-1}, G^{-1/2}mm'G^{-1/2})C^{(k)}(E_{11}, E_{11})/C_{\varphi}(I_q) \]

\[ (11) = C^{(j)(k)}(G^{-1}, G^{-1/2}mm'G^{-1/2})/C^{(f)}(I_q) . \]

In the last two expressions \( C^{(j), (k)} \) denotes an invariant polynomial in the elements of its two argument matrices. These polynomials were introduced by Davis [2, 3] to extend the zonal polynomials and the reader is referred to his articles for a detailed presentation of their properties. \( \varphi \) is a partition of the integer \( f = j+k \) into \( \leq q \) parts and the notation \( \varphi \in (j)\cdot(k) \) which is defined in [2] relates the two sets of partitions
that appear in the summation. In the present case \( E_{11} = e_1 e_1^T \) where \( e_1 \) is the first unit vector and only top order partitions appear in the summation. The final expression above follows because \( C_{(f)}^{(j),(k)}(E_{11}, E_{11}) = 1 \). To simplify notation we will use \( C_f^{j,k}(\ , \ ) \) in place of \( C_{(f)}^{(j),(k)}(\ , \ ) \) in what follows.

From (8)-(11) we deduce that

\[
(12) \quad \text{pdf}(y|C) = 2^{-q/2} \frac{1}{\Gamma(q/2)} y^{q/2-1} (\text{det } G)^{-1/2} \sum_{j,k} \frac{(-1)^j (1/2)^j (1/4)^k}{j! k!(1/2)_k} C_f^{j,k}(G^{-1}, G^{-1/2} y^{1/2} G^{-1/2}) C_f (I_q).
\]

Note that when \( m = 0 \) only terms for which \( k = 0 \) are nonzero, leading to

\[
2^{-q/2} \frac{1}{\Gamma(q/2)} y^{q/2-1} (\text{det } G)^{-1/2} \sum_{j} \frac{(-y/2)^j}{j!} C(j)(G^{-1}) C(j)(I_q)
\]

\[
= 2^{q/2} \frac{1}{\Gamma(q/2)} y^{q/2-1} (\text{det } G)^{-1/2} {}_0F_0 \left( -\frac{1}{2} G^{-1}, y \right)
\]

as in [7].

Since

\[
\text{pdf}(C) = \frac{\text{etr}(-G^{-1}C/2)(\text{det } C)^{(N-n-1)/2}}{2^{nN/2} \Gamma(n/2)(\text{det } \Omega)^{n/2}}
\]

we find that
\[
\text{pdf}(y) = \frac{2^{-q/2} \Gamma(q/2)^{-1} \exp(-m\Gamma/2)}{2^n \Gamma(n/2) (\det \Omega)^{n/2}} \sum_{j,k} \frac{(-1)^j (1/2)^j (1/4)^k y^{j-1}}{j!k!(1/2)_k} C_f(j, k, I_q) J_{C > 0}
\]

\[
\cdot \text{etr}(-\Omega^{-1} C/2)
\]

\[
\cdot (\det C)^{(N-n-1)/2} \det (C \otimes M)^{1/2} C_f^j k (C \otimes M)^{1/2} (C \otimes M)^{1/2} (\det C)^{(N-n-1)/2} dC
\]

\[
= \frac{\exp(-m\Gamma/2) y^{q/2-1}}{2^{(q+nN)/2} \Gamma(q/2) \Gamma(n/2) (\det \Omega)^{N/2}} \sum_{j,k} \frac{(-1)^j (1/4)^k y^{j-1}}{j!k!(1/2)_k} C_f(j, k, I_q) J_{C > 0}
\]

\[
\cdot \int_{C > 0} \left[ \det (C \otimes M)^{1/2} C_f^j k (C \otimes M)^{1/2} (C \otimes M)^{1/2} (\det C)^{(N-n-1)/2} \right]_{Z=0} (\det C)^{(N-n-1)/2} dC
\]

where \( Z \) is an \( n \times n \) matrix of auxiliary variables and \( \Im Z = \Im \Im Z \). The integral over \( C \) in the expression above is absolutely and uniformly convergent for all \( Z \) satisfying \( \Re(Z) \leq \epsilon I \) where \( \epsilon \) is any positive quantity less than the smallest latent root of \( \Omega^{-1/2} \). As in [7], we may therefore take both the operator involving \( \Im Z \) and the evaluation at \( Z = 0 \) outside the integration, yielding:

\[
\text{pdf}(y) = \frac{\exp(-m\Gamma/2) y^{q/2-1}}{2^{(q+nN)/2} \Gamma(q/2) (\det \Omega)^{N/2}} \sum_{j,k} \frac{(-1)^j (1/4)^k y^{j-1}}{j!k!(1/2)_k} C_f(j, k, I_q) J_{C > 0}
\]

\[
\cdot \left[ \det (C \otimes M)^{1/2} C_f^j k (C \otimes M)^{1/2} (C \otimes M)^{1/2} (\det C)^{(N-n-1)/2} \right]_{Z=0}
\]
\[
\frac{\exp(-m'm/2)y^{q/2-1}}{2^{q/2} \Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y_f^j}{j!k!(1/2)^k c(\xi^j)} (1_q^j)
\]

\[\cdot \left[ \det(\bar{D}(\bar{\alpha}\otimes M)\bar{D}')^{1/2} \right]^{1/2} \cdot \frac{1}{2^{q/2} \Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y_f^j}{j!k!(1/2)^k c(\xi^j)} (1_q^j)
\]

\[\cdot \left[ \det(\bar{D}(\bar{\alpha}\otimes M)\bar{D}')^{1/2} \right]^{1/2} \cdot \left[ \det(I - 2\Omega)^{-N/2} \right]_{z=0}
\]

Transforming \( Z + 2\Omega^{1/2}z\Omega^{1/2} = x \) and using the rule (see [7]) that
\[\bar{\alpha}_z = 2\Omega^{1/2}z\Omega^{1/2}\]
we obtain:

\[\text{(13)} \quad \text{pdf}(y) = \frac{\exp(-m'm/2)y^{q/2-1}}{\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y_f^j}{j!k!(1/2)^k c(\xi^j)} (1_q^j)
\]

\[\cdot \left[ \det(L(\bar{\alpha}X\otimes I)L')^{1/2} \right]^{1/2} \cdot \left[ \det(I-x)^{-N/2} \right]_{x=0}
\]

where

\[L = V^{-1/2}(\Omega^{1/2} \otimes M^{1/2}) = (D(\bar{\alpha} \otimes M)D')^{-1/2}D(\bar{\alpha}^{1/2} \otimes M^{1/2}) .
\]

Since \( w = Ny \) we have:

\[\text{(14)} \quad \text{pdf}(w) = \frac{\exp(-m'm/2)w^{q/2-1}}{N^{q/2} \Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k (w/N)^j}{j!k!(1/2)^k c(\xi^j)} (1_q^j)
\]

\[\cdot \left[ \det(L(\bar{\alpha}X\otimes I)L')^{1/2} \right]^{1/2} \cdot \left[ \det(I-x)^{-N/2} \right]_{x=0}
\]

4. SPECIALIZATIONS

4.1. The Regression F Statistic

When \( n = 1 \), the model reduces to the general linear model \( \Omega = \sigma^2 \),
say, the hypothesis (2) becomes \( H_0: D\alpha = d \) and \( \bar{\alpha}X \) becomes the scalar
operator \( \bar{\alpha}x = d/dx \). Since \( LL' = I_q \) we find that the density (13) re-
duces to:
\[ p(y) = \exp\left(-\frac{m'm}{2}\right)y^{q/2-1} \sum_{j,k} (-1)^j \frac{j!(1/2)^k}{j!(1/2)^k} C_f(I_q) \]

Note that by the rules of fractional differentiation developed in [6]:

\[ ax^\mu (1-x)^{-\beta} = \frac{\Gamma(\beta+\mu)}{\Gamma(\beta)} (1-x)^{-\beta-\mu} ; \quad \text{Re}(\beta) > 0 , \quad \text{Re}(\beta+\mu) > 0 \]

and from (2):

\[ C_f^{j,k}(I_q, m') = C_f(I_q) (m'm)^k / C(k)(I_q) . \]

Now using the fact that

\[ \left(\frac{1}{2}\right)_k C(k)(I_q) = \left(\frac{q}{2}\right)_k \]

we deduce from (15), (16) and (17) that:

\[ p(y) = \exp\left(-\frac{m'm}{2}\right) \frac{\Gamma((N+q)/2)}{\Gamma(N/2)} y^{q/2-1} \sum_{j,k} (-y)^j \frac{(m'm/2)^k y^k ((N+q)/2)_f}{j!(1/2)^k} \]

\[ = \frac{\exp\left(-\frac{m'm}{2}\right)}{B(q/2, N/2)} y^{q/2-1} \sum_{k} \frac{(m'm/2)^k y^k ((N+q)/2)_k}{k!(q/2)_k} \sum_{j} \frac{(-y)^j ((N+q)/2)_j}{j!} \]

\[ \quad = \frac{\exp\left(-\frac{m'm}{2}\right)}{B(q/2, N/2)} \frac{y^{q/2-1}}{(1+y)^{(N+q)/2}} \, _1F_1\left(\frac{N+q}{2} \cdot \frac{q}{2}; \frac{m'm}{2} \left(\frac{y}{1+y}\right)\right) . \]

It follows that

\[ F = N y/q \equiv F_{q, N}(\sigma^2) \]
as in standard regression theory. The non centrality parameter is 
\[ \delta^2 = m'm = b'V^{-1}b = (Da-d)'(DMD')^{-1}(Da-d)/\sigma^2. \]

4.2. Hotelling's $T^2$

In this case the null hypothesis takes the form $H_0 : FA = d$ so that $D = F \otimes g'$ for some $m$-vector $g$ and $q \times n$ matrix $F$ of full rank $q < n$. Setting $E = (F_{nF})^{-1/2}F_{nF}^{-1/2}$ we find that (13) is:

\[
\text{pdf}(y) = \frac{\exp(-m'm/2)q^{q/2-1}}{\Gamma(q/2)} L_{j,k} \frac{(-1)^j(1/2)^j \kappa_f}{j!k!(1/2)_k C_f(I_q)} \times \left[ \det(E \otimes E')^{1/2} c_{j,k}^f (E \otimes E')^{1/2} (E \otimes E')^{1/2} \right]_{X=0} \cdot \det(I-X)^{-N/2}.
\]

As in Section 5 of [7] we construct an $n \times n$ orthogonal matrix $P' = [E' : K']$ and transforming $X \rightarrow PXP' = Z$ we find

\[
(19) \quad \text{pdf}(y) = \frac{\exp(-m'm/2)q^{q/2-1}}{\Gamma(q/2)} L_{j,k} \frac{(-1)^j(1/2)^j \kappa_f}{j!k!(1/2)_k C_f(I_q)} \times \left[ \det(z_{11})^{1/2} c_{j,k}^f (z_{11}, z_{11}^{1/2} (z_{11}^{-1/2}) \det(I-Z_{11})^{-N/2} \right]_{Z_{11}=0}
\]

where $Z_{11}$ is the leading $q \times q$ sub matrix of $Z$. Now

\[
(20) \quad (\det z_{11})^{1/2} \det(I-Z_{11})^{-N/2} = \frac{\Gamma_q((N+1)/2)}{\Gamma_q(N/2)} \det(I-Z_{11})^{-(N+1)/2}
\]

and

\[
c^j_k (z_{11}, z_{11}^{1/2} (z_{11}^{-1/2}) \det(I-Z_{11})^{-(N+1)/2}
\]

\[
= \frac{\Gamma_q((N+1)/2)}{\Gamma_q((N+1)/2)} \int_{S>0} \text{etr}(-S)(\det S)^{-N/2-(q+1)/2} c^j_k (S, Smm') dS
\]

\[
(21) = \frac{\Gamma_q((N+1)/2, f)}{\Gamma_q((N+1)/2)} c^j_k (I, mm')
\]
where the final expression follows from one of the Laplace transforms given in [2] and $\Gamma_q((N+1)/2,f)$ is the constant introduced by Constantine [1].

In the present case

$$\frac{\Gamma_q((N+1)/2,f)}{\Gamma_q((N+1)/2)} = ((N+1)/2)_f .$$

From (19)-(21) and (17) we deduce that:

$$\text{pdf}(y) = \frac{\exp(-m'm/2)\Gamma_q((N+1)/2)y^{q/2-1}}{\Gamma(q/2)\Gamma((N-q+1)/2)} \sum_{k} \frac{(-y)^j (m'm/2)^k y^{k((N+1)/2)_f}}{j! k!(q/2)_k}$$

$$= \frac{\exp(-m'm/2)\Gamma_q((N+1)/2)y^{q/2-1}}{\Gamma(q/2)\Gamma((N-q+1)/2)} \sum_{j} \frac{(m'm/2)^k y^{k((N+1)/2)_f}}{k!(q/2)_k} \left(-y\right)^j \left((N+1)/2+k\right)_j$$

$$= \frac{\exp(-m'm/2)y^{q/2-1}}{B(q/2, (N-q+1)/2)(1+y)^{(N+1)/2}} \text{F}(\frac{1}{2}, \frac{q}{2}, \frac{m'm}{2(1+y)}) .$$

Thus

$$F = \frac{N-q+1}{q} \equiv F_{q,N-q+1}(\delta^2) ;$$

where the non centrality parameter is $\delta^2 = m'm = b'V^{-1}b$

$$= (F\delta_d)'(F\delta_P')^{-1}(F\delta_d)/g'Mg .$$

4.3. Asymptotic Theory

As a first approximation to the exact density (14) in the general case we make the replacement

$$\det(I-X)^{-N/2} \sim \text{str}(NX/2)$$

which is appropriate in the neighborhood of $X = 0$. Formula (14) simplifies under this approximation to:
\[ \text{pdf}(w) = \frac{\exp(-m' m/2) w^{q/2-1}}{\Gamma(q/2) 2^{q/2}} \sum_{j,k} (-1)^j \frac{j!}{(1/2)_j} \frac{(w/N)^j}{\Gamma((1/2)_j) C(j)(1q)} \frac{f^j_k}{\Gamma((1/2)_k)} \left[ \frac{N/2}{q/2} \right] \frac{q/2 + f^j_k}{f^j_k} C(j)(1q) \frac{m' m'}{q/2} \]

\[ = \frac{\exp(-m' m/2) w^{q/2-1}}{2^{q/2} \Gamma(q/2)} \sum_{j=0}^\infty \frac{(-w/2)^j j!}{j!} \sum_{k=0}^\infty \frac{(m' m/2)^k (w/2)^k k!}{k! (q/2)_k} \]

\[ = \frac{\exp(-m' m/2) w^{q/2-1} e^{-w/2}}{2^{q/2} \Gamma(q/2)} \text{F}_1 \left( q/2; \frac{m' m}{2} \right) \]

\[ = \chi_q^2(\delta^2) \]

where the noncentrality parameter is

\[ \delta^2 = m' m = b' V^{-1} b = (D \text{ vec } A-d)' (D(\Omega \otimes I) D')^{-1} (D \text{ vec } A-d) \]

4.4. The Null Distribution

When \( m = 0 \) only terms for which \( k = 0 \) in (13) are nonzero. Since \( C(j,0)(A,B) = C_{(j)}(A) \), (13) becomes:

\[ \text{pdf}(y) = \frac{\gamma^{q/2-1}}{\Gamma(q/2)} \sum_{j=1}^\infty \frac{(-y)^j j!}{j! C(j)(1q)} \left[ \text{det}(L(\partial X \otimes I)L')^{1/2} C_{(j)}(L(\partial X \otimes I)L') \text{det}(I-X)^{-N/2} \right]_{X=0} \]

\[ = \frac{\gamma^{q/2-1}}{\Gamma(q/2)} \text{det}(L(\partial X \otimes I)L')^{1/2} \text{F}_0 \left( -L(\partial X \otimes I)L', y \right) \text{det}(I-X)^{-N/2} \right]_{X=0} \]

as established in [7], equation (14).

5. SUMMARY AND CONCLUSION

This paper extends the distribution theory for the Wald statistic given in [7] to the noncentral case. Well known formulae for the noncentral distributions of the regression \( F \) statistic and Hotelling's \( T^2 \) are derived as special cases of our theory. Also of interest is the fact that
our general expression (14) for the probability density of the Wald statistic yields after a simple manipulation the asymptotic noncentral $\chi^2$ distribution that holds for local alternatives.

The methods used here and in the earlier articles [5, 6, 7] seem likely to be useful in many other problems of distribution theory. As we have seen in Section 4, they have the interesting property of allowing asymptotic distributions to be deduced quite simply as specializations of finite sample results. In this respect they differ from conventional methods that have been used in this field which often lead to expressions from which it is very difficult to deduce asymptotic results.

6. REFERENCES


