Cowles Foundation Paper #654

Econometrica, Vol. 54, No. 4 (July, 1986), 881-895

THE EXACT DISTRIBUTION OF THE WALD STATISTIC

BY P. C. B. PHILLIPS[1]

This paper derives the exact finite sample distribution of the Wald statistic for testing general linear restrictions on the coefficients in the multivariate linear model. This generalizes all previously known results, including those for the standard F statistic in linear regression, for Hotelling's T^2 test, and for Hotelling's generalized T_0^2 test. The results presented here encompass both the null and the non-null distributions. They also yield in a simple and elegant way the asymptotic distribution theory and related higher order asymptotic expansions. Various specializations of our general result are presented, including a computable formula for the null distribution in the case of a test of single restriction. Conventional classical assumptions of normally distributed errors and nonrandom exogenous variables are employed.

KEYWORDS: Asymptotics, exact distribution, fractional calculus, linear model, Wald statistic.

1. INTRODUCTION

ONE FIELD IN WHICH ECONOMETRIC DISTRIBUTION THEORY would appear to be in a particularly well developed state is the linear model with normally distributed errors. In this model we have an established battery of exact statistical tests for practitioners. The most well known of these are the commonly used t and F ratio test statistics. Also well known, but less commonly used, are Hotelling's T^2 statistic and Hotelling's generalized T_0^2 test. Other test statistics for which distributional results have been obtained in the linear model are the likelihood ratio statistic, Pillai's statistic, and Roy's largest latent root test. Muirhead (1982) contains a detailed review of existing analytical results in this field including both exact formulae and asymptotic expansions. All the above mentioned statistics may be used for testing linear hypotheses about the coefficients in the multivariate linear model.

In spite of the extensive research in this field (exemplified by Chapter 10 of Muirhead (1982)) there are still major unsolved problems. One of these, which is particularly interesting to econometricians, is the distribution of the Wald statistic for testing general linear restrictions on the coefficients. In special cases where the coefficient matrix in the restrictions takes on a Kronecker product form (corresponding to the GMANOVA and classical MANOVA models in multivariate analysis) the Wald statistic reduces to the generalized T_0^2 statistic and known results for the distribution of this statistic apply. However, when the coefficient

[1] This paper is dedicated to Professor J. D. Sargan. Some of its results were first presented by the author in an invited address to the SSRC Conference in honor of Professor Sargan at Oxford University during March, 1984. The present paper combines, at the suggestion of a Co-Editor, and extends the results contained in two earlier papers (Phillips, 1984b; Phillips and Ouliaris, 1984). My co-author in Phillips and Ouliaris (1984) graciously waived his right of authorship to some of the material in this combined paper. I am grateful to six anonymous referees of *Econometrica* for supportive reviews and some helpful comments.

My thanks, as always, go to Glena Ames for her skill and effort in typing the manuscript of this paper and to the NSF for research support under Grant No. SES 8218792. The work reported here was completed during the author's tenure of a Guggenheim Fellowship in 1984/85.

matrix is not of the Kronecker product form, none of the presently available results are applicable. Important practical examples arise in applied demand analysis in econometrics where the Slutsky symmetry condition leads to simple linear across equation restrictions which cannot be formulated in a Kronecker product representation.

The purpose of the present paper is to provide a unified and comprehensive general solution to the distribution of the Wald statistic in this multivariate linear model setting. More specifically, the paper derives the exact probability density function (pdf) of the Wald statistic for testing the null hypothesis of quite general linear restrictions on the coefficients. The most general result we present applies both to the null distribution of the statistic and to the non-null distribution which obtains under a generally specified alternative hypothesis. Various specializations of our general result are considered in detail. These include the regression F statistic, Hotelling's T^2 statistic, and the important special case in which only a single restriction is under test. For the latter case, a formula for the null distribution which can be readily computed is also presented. The analysis that leads to the non-null distribution of the Wald statistic involves the derivation of the density of a noncentral positive definite quadratic form in normal variates. This density should be useful in other work as well.

We also provide a simple and novel derivation of conventional asymptotic theory as a specialization of our general finite sample results. This approach may be extended to generate the formulae for higher order asymptotic expansions. Thus, the results of the paper provide a meaningful unification of conventional asymptotics, higher order asymptotic expansions, and exact finite sample distribution theory in this context.

The analytical methods employed here rely on the fractional matrix calculus developed by the author in other recent work (Phillips, 1985). The reader is referred to Phillips (1985) for an introduction to these techniques and for another application of them in econometric distribution theory.

2. THE MODEL AND NOTATION

We write the multivariate linear model in the form:

$$y_t = Ax_t + u_t \qquad (t = 1, \ldots, T). \tag{1}$$

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

$$H_0\colon D \operatorname{vec} A = d, \qquad H_1\colon D \operatorname{vec} A - d = b \neq 0, \tag{2}$$

where D is a $q \times np$ matrix of known constants of rank q, d is a known vector, and vec (A) stacks the rows of A.

From least squares estimation of (1) we have

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

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

The Wald statistic for testing the hypothesis (2) is

$$(4)\qquad \begin{aligned} w &= (D\operatorname{vec}A^*-d)'\{D(\Omega^*\otimes M)D'\}^{-1}(D\operatorname{vec}A^*-d)\\ &= Nl'Bl, \end{aligned}$$

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

We define $y=l'Bl$ and write y in canonical form as

$$(5)\qquad y=g'Gg,$$

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

3. THE GENERAL 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. Our first task is to derive this conditional distribution.

We define $z=G^{1/2}g$. Then $y=z'z$ and

$$(6)\qquad \begin{aligned} \operatorname{pdf}(z\mid C) &= (2\pi)^{-q/2}(\det G)^{-1/2}\exp(-m'm/2)\\ &\quad\cdot\exp(-z'G^{-1}z/2)\exp(G^{-1/2}zm'). \end{aligned}$$

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:

$$(7)\qquad \begin{aligned} &(2\pi)^{-q/2}(\det G)^{-1/2}\exp(-m'm/2)\\ &\quad\cdot\exp(-z'G^{-1}z/2)\,{}_0F_1(\tfrac{1}{2},\tfrac{1}{4}z'G^{-1/2}mm'G^{-1/2}z). \end{aligned}$$

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

$$dz=(1/2)y^{q/2-1}\,dy(dh)$$

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

$$(8)\qquad \begin{aligned} \operatorname{pdf}(y\mid C) &= 2^{-q/2-1}\pi^{-q/2}\exp(-m'm/2)y^{q/2-1}(\det G)^{-1/2}\\ &\quad\cdot\int_{V_{1,q}}\operatorname{etr}(-yG^{-1}hh'/2)\,{}_0F_1(\tfrac{1}{2},\tfrac{1}{4}yh'G^{-1/2}mm'G^{-1/2}h)(dh)\\ &= 2^{-q/2}[\Gamma(q/2)]^{-1}\exp(-m'm/2)y^{q/2-1}(\det G)^{-1/2}\\ &\quad\cdot\int_{V_{1,q}}\operatorname{etr}(-yG^{-1}hh'/2)\,{}_0F_1(\tfrac{1}{2},\tfrac{1}{4}yh'G^{-1/2}mm'G^{-1/2}h)(\boldsymbol{dh}) \end{aligned}$$

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

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

$$\text{(9)} \qquad \operatorname{etr}(-yG^{-1}hh'/2) = \sum_{j=0}^{\infty} \frac{(-y/2)^j}{j!} C_{(j)}(G^{-1}hh'),$$

$$\text{(10)} \qquad {}_0F_1(\tfrac{1}{2}, \tfrac{1}{4}yh'G^{-1/2}mm'G^{-1/2}h) = \sum_{k=0}^{\infty} \frac{(y/4)^k}{k!(1/2)_k} C_{(k)}(G^{-1/2}mm'G^{-1/2}hh')$$

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 James (1964).

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

$$\text{(11a)} \qquad \int_{V_{1,q}} C_{(j)}(G^{-1}hh')C_{(k)}(G^{-1/2}mm'G^{-1/2}hh')(dh)$$

$$= \int_{O(q)} C_{(j)}(G^{-1}HE_{11}H')C_{(k)}(G^{-1/2}mm'G^{-1/2}HE_{11}H')(dH)$$

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

$$\text{(11b)} \qquad = C_{(f)}^{(j),(k)}(G^{-1}, G^{-1/2}mm'G^{-1/2})/C_{(f)}(I_q)$$

$$\text{(11c)} \qquad = C_{(f)}^{(j),(k)}(G^{-1}, G^{-1}mm')/C_{(f)}(I_q).$$

In the above expressions $C_{\varphi}^{(j),(k)}$ denotes an invariant polynomial in the elements of its two argument matrices. These polynomials were introduced by Davis (1979, 1980) to extend the zonal polynomials and the reader is referred to his articles for a detailed presentation of their properties. φ 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 (Davis, 1979), relates the two sets of partitions that appear in the summation. In the present case only top order partitions appear in the summation since, from (11a), $E_{11} = e_1 e_1'$ where e_1 is the first unit vector. Moreover, $C_{(f)}^{(j),(k)}(E_{11}, E_{11}) = 1$, leading to (11b). Finally, for any two integers a and b

$$\text{(12)} \qquad \operatorname{tr}\{(G^{-1})^a(G^{-1/2}mm'G^{-1/2})^b\} = \operatorname{tr}\{(G^{-1})^a(G^{-1}mm')^b\}$$

and (11c) follows because distinct products of powers of traces such as (12) form a basis for the invariant polynomials (Davis, 1979). To simplify notation in what follows we shall use $C_f^{j,k}(\ ,\ )$ in place of $C_{(f)}^{(j),(k)}(\ ,\ )$.

From (8)-(11) we deduce that

$$\text{(13)} \qquad \operatorname{pdf}(y \mid C) = 2^{-q/2}[\Gamma(q/2)]^{-1} \exp(-m'm/2) y^{q/2-1} (\det G)^{-1/2}$$

$$\cdot \sum_{j,k} \frac{(-1)^j (1/2)^j (1/4)^k y^f}{j!k!(1/2)_k} C_f^{j,k}(G^{-1}, G^{-1}mm')/C_{(f)}(I_q).$$

It is worthwhile to remark that (13) provides a general expression for the pdf of a noncentral positive definite quadratic form in normal variates. Interestingly, this density does not seem to have been derived before in the statistical literature. Note that when $m=0$ only terms for which $k=0$ are nonzero, leading to

$$(14)\qquad 2^{-q/2}[\Gamma(q/2)]^{-1}y^{q/2-1}(\det G)^{-1/2}\sum_j \frac{(-y/2)^j}{j!}C_{(j)}(G^{-1})/C_{(j)}(I_q)$$
$$=2^{q/2}[\Gamma(q/2)]^{-1}y^{q/2-1}(\det G)^{-1/2}{}_0F_0(-\tfrac{1}{2}G^{-1},y)$$

which is the density of a central positive definite form in normal variates, as given in James (1964).

In the present case (13) and (14) are conditional densities given G or, equivalently, C the random Wishart matrix upon which G depends. In fact, $G^{-1}=\bar{D}(C\otimes M)\bar{D}'$ is linear in the elements of C. Our second task in finding the unconditional density of y is to average the conditional distribution weighted by the density of C.

Since

$$\operatorname{pdf}(C)=\frac{\operatorname{etr}(-\Omega^{-1}C/2)(\det C)^{(N-n-1)/2}}{2^{nN/2}\Gamma_n(N/2)(\det\Omega)^{N/2}},$$

we find that

$$\operatorname{pdf}(y)=\int_{C>0}\operatorname{pdf}(y\,|\,C)\operatorname{pdf}(C)\,dC$$
$$=\frac{2^{-q/2}[\Gamma(q/2)]^{-1}\exp(-m'm/2)y^{q/2-1}}{2^{nN/2}\Gamma_n(N/2)(\det\Omega)^{N/2}}$$
$$\cdot\sum_{j,k}\frac{(-1)^j(1/2)^j(1/4)^k y^f}{j!k!(1/2)_k C_{(f)}(I_q)}\int_{C>0}\operatorname{etr}(-\Omega^{-1}C/2)$$
$$\cdot(\det C)^{(N-n-1)/2}\det(\bar{D}(C\otimes M)\bar{D}')^{1/2}C_f^{j,k}$$
$$\cdot(\bar{D}(C\otimes M)\bar{D}',(\bar{D}(C\otimes M)\bar{D}')mm')\,dC.$$

Term by term integration of the series is justified by a theorem of Hardy (see Titchmarsh, 1939, p. 47). The series may be shown by majorization to be absolutely convergent for $0\leq y<1$.

Using the theory of matrix fractional calculus developed in Phillips (1985) we now write:

$$(16)\qquad \det(\bar{D}(C\otimes M)\bar{D}')^{1/2}C_f^{j,k}(\bar{D}(C\otimes M)\bar{D}',\bar{D}(C\otimes M)\bar{D}'mm')$$
$$=[\det(\bar{D}(\partial Z\otimes M)\bar{D}')^{1/2}C_f^{j,k}(\bar{D}(\partial Z\otimes M)\bar{D}',\bar{D}(\partial Z\otimes M)\bar{D}'mm')$$
$$\cdot\operatorname{etr}(CZ)]_{Z=0}$$

where Z is an auxiliary matrix of dimension $n\times n$ and ∂Z denotes the matrix differential operator $\partial/\partial Z$. In (16)$C_f^{j,k}$ is a polynomial in the elements of the matrix operator ∂Z and is well defined in the sense of conventional calculus. The operator $\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{1/2}$ in (16) is a fractional matrix operator which is to be understood in the sense of the definition provided in Phillips (1985). Thus, if $f(Z)$ is an analytic function of the matrix of complex variables Z and if μ is any complex constant, we define:

(17)
$$\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{\mu} f(Z)$$
$$= (\Gamma_q(\alpha))^{-1} \int_{S>0} [\operatorname{etr}\{-\bar{D}(\partial Z \otimes M)\bar{D}'S\} \det(\bar{D}(\partial Z \otimes M)\bar{D}')^{m} f(Z)] \cdot (\det S)^{\alpha-(q+1)/2}\, dS$$

provided the integral is absolutely convergent. In (17) α is a complex quantity for which $Re(\alpha) > (q-1)/2$ and m is a positive integer; they are selected so that $\mu = m - \alpha$. When $f(Z) = \operatorname{etr}(CZ)$ we find by a simple evaluation that $\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{1/2} \operatorname{etr}(CZ) = \operatorname{etr}(CZ) \det(\bar{D}(C \otimes M)\bar{D}')^{1/2}$ as required for the validity of (16). The reader is referred to Phillips (1985) for a detailed development of the theory and the rules for the manipulation of fractional (and possibly complex) matrix operators such as (17). Note that, with this interpretation, (16) provides a linear pseudodifferential operator representation of the function on the left side of the equation.

Using (16) we write (15) in alternate form as:

$$\operatorname{pdf}(y) = \frac{\exp(-m'm/2)y^{q/2-1}}{2^{(q+nN)/2}\Gamma(q/2)\Gamma_n(N/2)(\det \Omega)^{N/2}} \sum_{j,k} \frac{(-1/2)^j (1/4)^k y^f}{j!k!(1/2)_k C_{(f)}(I_q)}$$
$$\cdot \int_{C>0} [\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{1/2}$$
$$\cdot C_f^{j,k}(\bar{D}(\partial Z \otimes M)\bar{D}', (\bar{D}(\partial Z \otimes M)\bar{D}')mm')$$
$$\cdot \operatorname{etr}\{-((1/2)\Omega^{-1} - Z)C\}]_{Z=0} (\det C)^{(N-n-1)/2}\, dC.$$

The integral over C in the expression above is absolutely and uniformly convergent for all Z satisfying $Re(Z) \leq \varepsilon I$, where ε is any positive quantity less than the smallest latent root of $\Omega^{-1}/2$. We may therefore take both the operator involving ∂Z and the evaluation at $Z = 0$ outside the integration, yielding:

$$\operatorname{pdf}(y) = \frac{\exp(-m'm/2)y^{q/2-1}}{2^{(q+nN)}\Gamma(q/2)(\det \Omega)^{N/2}} \sum_{j,k} \frac{(-1/2)^j (1/4)^k y^f}{j!k!(1/2)_k} C_{(f)}(I_q)$$
$$\cdot [\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{1/2}$$
$$\cdot C_f^{j,k}(\bar{D}(\partial Z \otimes M)\bar{D}', (\bar{D}(\partial Z \otimes M)\bar{D}')mm')$$
$$\cdot \det(\Omega^{-1}/2 - Z)^{-N/2}]_{Z=0}$$
$$= \frac{\exp(-m'm/2)y^{q/2-1}}{2^{q/2}\Gamma(q/2)} \sum_{j,k} \frac{(-1/2)^j (1/4)^k y^f}{j!k!(1/2)_k C_{(f)}(I_q)}$$
$$\cdot [\det(\bar{D}(\partial Z \otimes M)\bar{D}')^{1/2} C_f^{j,k}$$
$$\cdot (\bar{D}(\partial Z \otimes M)\bar{D}', (\bar{D}(\partial Z \otimes M)\bar{D}')mm')$$
$$\cdot \det(I - 2\Omega Z)^{-N/2}]_{Z=0}.$$

Transforming $Z \to 2\Omega^{1/2} Z \Omega^{1/2} = X$ and using the easily established rule of transformation that $\partial Z = 2\Omega^{1/2} \partial X \Omega^{1/2}$, we obtain:

$$\text{(18)} \qquad \mathrm{pdf}(y) = \frac{\exp(-m'm/2) y^{q/2-1}}{\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y^f}{j!k!(1/2)_k C_{(f)}(I_q)} \cdot [\det(L(\partial X \otimes I)L')^{1/2} \cdot C_f^{j,k}(L(\partial X \otimes I)L', (L(\partial X \otimes I)L')mm') \det(I-X)^{-N/2}]_{X=0}$$

where

$$L = V^{-1/2} D(\Omega^{1/2} \otimes M^{1/2}) = (D(\Omega \otimes M)D')^{-1/2} D(\Omega^{1/2} \otimes M^{1/2}).$$

The series (18) is certainly convergent for $0 \leq y < 1$. Moreover, as we shall see from several examples in the next section, the domain of convergence may often be extended by analytic continuation to the entire interval $0 \leq y < \infty$ after a simple manipulation of the series.

Since $W = Ny$ we deduce from (18) the following expression for the density of $W (0 \leq W < N)$:

$$\text{(19)} \qquad \mathrm{pdf}(w) = \frac{\exp(-m'm/2) w^{q/2-1}}{N^{q/2}\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k (w/N)^f}{j!k!(1/2)_k C_{(f)}(I_q)} \cdot [\det(L(\partial X \otimes I)L')^{1/2} \cdot C_f^{j,k}(L(\partial X \otimes I)L', (L(\partial X \otimes I)L')mm') \det(I-X)^{-N/2}]_{X=0}.$$

Alternative formulae for these densities which are everywhere convergent may be obtained by proceeding directly from (8) rather than (13). The derivations are omitted and we give only the final expression here:

$$\text{(20)} \qquad \mathrm{pdf}(y) = [\Gamma(q/2)]^{-1} \exp(-m'm/2) y^{q/2-1} \Bigg[\det(L(\partial X \otimes I)L')^{1/2} \cdot \int_{V_{1,q}} \exp\{-yh'L(\partial X \otimes I)L'h\} \cdot {}_0F_1(\tfrac{1}{2}, \tfrac{1}{2} yh'L(\partial X \otimes I)L'mm'h)(dh) \det(I-X)^{-N/2} \Bigg]_{X=0}.$$

The corresponding formula for the density of $W = Ny$ follows by transformation.

4. SPECIALIZATIONS

4.1. *The Regression F Statistic*

When $n = 1$, the model reduces to the general linear model, Ω is a scalar parameter (σ^2, say), the hypothesis (2) becomes H_0: $Da = d$ and ∂X becomes the scalar operator $\partial x = d/dx$. Since $LL' = I_q$ we find that the density (18) reduces to:

$$\text{(21)} \qquad \mathrm{pdf}(y) = \frac{\exp(-m'm/2) y^{q/2-1}}{\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y^f}{j!k!(1/2)_k C_{(f)}(I_q)} \cdot C_f^{j,k}(I_q, mm')[\partial x^{q/2+f}(1-x)^{-N/2}]_{x=0}.$$

Note that by the rules of fractional differentiation developed in Phillips (1985):

$$\partial x^{\mu}(1-x)^{-\beta} = \frac{\Gamma(\beta+\mu)}{\Gamma(\beta)}(1-x)^{-\beta-\mu}, \quad Re\,(\beta)>0, \quad Re\,(\beta+\mu)>0, \tag{22}$$

and from Davis (1979):

$$C_f^{j,k}(I_q, mm') = C_{(f)}(I_q)(m'm)^k / C_{(k)}(I_q). \tag{23}$$

Now using the fact that

$$\left(\frac{1}{2}\right)_k C_{(k)}(I_q) = \left(\frac{q}{2}\right)_k,$$

we deduce from (21), (22), and (23) that

$$\begin{aligned}
\text{pdf}\,(y) &= \frac{\exp(-m'm/2)\Gamma((N+q)/2)y^{q/2-1}}{\Gamma(q/2)\Gamma(N/2)} \\
&\quad \cdot \sum_{j,k} \frac{(-y)^j (m'm/2)^k y^f ((N+q)/2)_f}{j!k!(q/2)_k} \\
&= \frac{\exp(-m'm/2)y^{q/2-1}}{B(q/2, N/2)} \\
&\quad \cdot \sum_k \frac{(m'm/2)^k y^k ((N+q)/2)_k}{k!(q/2)_k} \sum_j \frac{(-y)^j ((N+q)/2+k)_j}{j!} \\
&= \frac{\exp(-m'm/2)}{B(q/2, N/2)} \frac{y^{q/2-1}}{(1+y)^{(N+q)/2}} {}_1F_1\left(\frac{N+q}{2}, \frac{q}{2}; \frac{m'm}{2}\left(\frac{y}{1+y}\right)\right).
\end{aligned}$$

The final line of this derivation uses the series representation of $(1+y)^{-(N+q)/2-k}$, which is convergent for $0 \leq y < 1$. However, by analytic continuation this condition on y may be relaxed and the stated formula represents the density over the entire interval $[0, \infty)$.

It follows by inspection that:

$$F = Ny/q \equiv F_{q,N}(\delta^2)$$

as in standard regression theory. The noncentrality 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: $D_1 A d_2 = d$ so that $D = D_1 \otimes d_2'$ for some p-vector d_2 and $q \times n$ matrix D_1 of full rank $q \leq n$. Setting $E = (D_1 \Omega D_1')^{-1/2} D_1 \Omega^{1/2}$ we find that (18) is

$$\begin{aligned}
\text{pdf}\,(y) &= \frac{\exp(-m'm/2)y^{q/2-1}}{\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j (1/2)^k y^f}{j!k!(1/2)_k C_f(I_q)} \\
&\quad \cdot [\det(E\,\partial X E')^{1/2} C_f^{j,k}(E\,\partial X E', E\,\partial X E' mm') \\
&\quad \cdot \det(I-X)^{-N/2}]_{X=0}.
\end{aligned}$$

We construct an $n\times n$ orthogonal matrix $P'=[E' \vdots K']$ and, transforming $X\to PXP'=Z$, we find

$$\text{(24)}\qquad \operatorname{pdf}(y)=\frac{\exp(-m'm/2)y^{q/2-1}}{\Gamma(q/2)}\sum_{j,k}\frac{(-1)^j(1/2)^k y^f}{j!k!(1/2)_k C_f(I_q)} \cdot[(\det\partial Z_{11})^{1/2}C_f^{j,k}(\partial Z_{11},\partial Z_{11}\,mm')\det(I-Z_{11})^{-N/2}]_{Z_{11}=0}$$

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

$$\text{(25)}\qquad (\det\partial Z_{11})^{1/2}\det(I-Z_{11})^{-N/2}=\frac{\Gamma_q((N+1)/2)}{\Gamma_q(N/2)}\det(I-Z_{11})^{-(N+1)/2}$$

and

$$\text{(26)}\qquad [C_f^{j,k}(\partial Z_{11},\partial Z_{11}mm')\det(I-Z_{11})^{-(N+1)/2}]_{Z_{11}=0}$$
$$=[\Gamma_q((N+1)/2)]^{-1}\int_{S>0}\operatorname{etr}(-S)\cdot(\det S)^{(N+1)/2-(q+1)/2}C_f^{j,k}(S,Smm')\,dS$$
$$=\frac{\Gamma_q((N+1)/2,f)}{\Gamma_q((N+1)/2)}C_f^{j,k}(I,mm')$$

where the final expression follows from one of the Laplace transforms given in Davis (1979) and $\Gamma_q((N+1)/2,f)$ is the constant introduced by Constantine (1963). In the present case

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

From (24)-(26) and (22) we deduce that:

$$\operatorname{pdf}(y)=\frac{\exp(-m'm/2)\Gamma_q((N+1)/2)y^{q/2-1}}{\Gamma(q/2)\Gamma_q(N/2)}\cdot\sum_{j,k}\frac{(-y)^j(m'm/2)^k y^k((N+1)/2)_f}{j!k!(q/2)_k}$$
$$=\frac{\exp(-m'm/2)\Gamma((N+1)/2)y^{q/2-1}}{\Gamma(q/2)\Gamma((N-q+1)/2)}\cdot\sum_k\frac{(m'm/2)^k y^k((N+1)/2)_k}{k!(q/2)_k}\sum_j\frac{(-y)^j((N+1)/2+k)_j}{j!}$$
$$=\frac{\exp(-m'm/2)y^{q/2-1}}{B(q/2,(N-q+1)/2)(1+y)^{(N+1)/2}}\cdot{}_1F_1\left(\frac{N+1}{2},\frac{q}{2};\frac{m'm}{2}\left(\frac{y}{1+y}\right)\right).$$

Thus, using analytic continuation as before, we find that

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

where the noncentrality parameter is

$$\delta^2 = m'm = b'V^{-1}b = (FAg-d)'(F\Omega F')^{-1}(FAg-d)/g'Mg.$$

4.3. *The Case of a Single Restriction*

If $q=1$ we have a test of a single restriction H_0: $d'\text{vec}\,A = e$, say, and $L = l' = (d'(\Omega\otimes M)d)^{-1/2}d'(\Omega^{1/2}\otimes M^{1/2})$. It will be convenient in what follows to set $\bar{l} = Kl$, where K is the commutation matrix[2] of order np, and to partition this vector as $\bar{l}' = (\bar{l}_1', \ldots, \bar{l}_p')$ where $\bar{l}_i$ is $n\times 1$. Define $Q = \sum_1^p \bar{l}_i\bar{l}_i'$. The density (19) is in this case:

$$\text{(27)}\qquad \text{pdf}(w) = \frac{e^{-m^2/2}w^{-1/2}}{N^{1/2}\Gamma(\frac{1}{2})}\sum_{j,k}\frac{(-1)^j(1/2)^k(w/N)^f}{j!k!(1/2)_k}m^{2k} \cdot[(\text{tr}\,(\partial XQ))^{f+1/2}\det(I-X)^{-N/2}]_{X=0}.$$

Now

$$(\text{tr}\,(\partial XQ))^{f+1}\det(I-X)^{-N/2} = \sum_{\varphi'} C_{\varphi'}(\partial XQ)\det(I-X)^{-N/2} = \sum_{\varphi'}\left(\frac{N}{2}\right)_{\varphi'} C_{\varphi'}((I-X)^{-1}Q)\cdot\det(I-X)^{-N/2}$$

where the summation is over all partitions φ' of the integer $f+1$ into $\leq n$ parts. Moreover

$$\begin{aligned}\text{(28)}\qquad &(\text{tr}\,(\partial XQ))^{f+1/2}\det(I-X)^{-N/2}\\ &= (\text{tr}\,(\partial XQ))^{-1/2}(\text{tr}\,(\partial XQ))^{f+1}\det(I-X)^{-N/2}\\ &= \sum_{\varphi'}\left(\frac{N}{2}\right)_{\varphi'}(\text{tr}\,(\partial XQ))^{-1/2}C_{\varphi'}((I-X)^{-1}Q)\det(I-X)^{-N/2}\\ &= \sum_{\varphi'}\left(\frac{N}{2}\right)_{\varphi'}\Gamma\left(\frac{1}{2}\right)^{-1}\int_0^\infty \text{etr}\,(-t\partial XQ)C_{\varphi'}((I-X)^{-1}Q)\\ &\quad\cdot\det(I-X)^{-N/2}t^{-1/2}\,dt\\ &= \sum_{\varphi'}\left(\frac{N}{2}\right)_{\varphi'}\Gamma\left(\frac{1}{2}\right)^{-1}\int_0^\infty t^{-1/2}C_{\varphi'}((I+tQ-X)^{-1}Q)\\ &\quad\cdot\det(I+tQ-X)^{-N/2}\,dt.\end{aligned}$$

After substitution of (28) in (27) we obtain:

$$\text{(29)}\qquad \text{pdf}(w) = \frac{e^{-m^2/2}w^{-1/2}}{N^{1/2}\Gamma(\frac{1}{2})^2}\sum_{j,k}\frac{(-1)^j(m^2/2)^k(w/N)^f}{j!k!(1/2)_k}\cdot\sum_{\varphi'}\left(\frac{N}{2}\right)_{\varphi'}\int_0^\infty t^{-1/2}C_{\varphi'}((I+tQ)^{-1}Q)\det(I+tQ)^{-N/2}\,dt.$$

[2] Specifically, $K = \sum_{i=1}^n\sum_{j=1}^p H_{ij}\otimes H_{ij}'$ where H_{ij} is the $n\times p$ matrix with 1 in the ijth position and zeros elsewhere.

A somewhat simpler expression for this density may be obtained by working from (20) rather than (19). The details of the derivation are omitted and we give only the final formula:

$$(30)\qquad \operatorname{pdf}(w) = \frac{e^{-m^2/2} w^{-1/2}}{N^{1/2}\Gamma(\frac{1}{2})^2} \sum_k \frac{(m^2/2)^k (w/N)^k}{k!(1/2)_k} \cdot \sum_{\kappa'} \left(\frac{N}{2}\right)_{\kappa'} \int_0^\infty t^{-1/2} C_{\kappa'}(Q[I+(w/n)Q+tQ]^{-1} \cdot \det[I+(w/N)Q+tQ]^{-N/2}\, dt$$

where the summation $\sum_{\kappa'}$ is over all partitions κ' of the integer $k+1$ into $\leq n$ parts.

4.4. *The Null Distribution in the General Case*

When $m=0$ only terms for which $k=0$ in (18) are nonzero. Since $C_j^{j,0}(A, B) = C_{(j)}(A)$, (18) becomes:

$$(31)\qquad \begin{aligned} \operatorname{pdf}(y) &= \frac{y^{q/2-1}}{\Gamma(q/2)} \sum_{j=1}^{\infty} \frac{(-y)^j}{j! C_{(j)}(I_q)} [\det(L(\partial X \otimes I)L')^{1/2} C_{(j)} \\ &\quad \cdot (L(\partial X \otimes I)L') \det(I-X)^{-N/2}]_{X=0} \\ &= \frac{y^{q/2-1}}{\Gamma(q/2)} [\det(L(\partial X \otimes I)L')^{1/2} \\ &\quad \cdot {}_0F_0(-L(\partial X \otimes I)L', y) \det(I-X)^{-N/2}]_{X=0} \end{aligned}$$

as found in Phillips (1984b) by direct methods. Note that (31) may also be deduced immediately from (20) and the definition of the ${}_0F_0$ function.

Thus, the null density of the Wald statistic W is:

$$(32)\qquad \operatorname{pdf}(w) = \frac{w^{q/2-1}}{N^{q/2}\Gamma(q/2)} [\det(L(\partial X \otimes I)L')^{1/2} \cdot {}_0F_0(-L(\partial X \otimes I)L', w/N) \det(I-X)^{-N/2}]_{X=0}.$$

This is a simple and general expression for the exact density which is very useful in analytic work. All presently known null distributions for the Wald statistic may be deduced quite simply from (32), including the complicated formulae obtained in Constantine (1966) for Hotelling's generalized T_0^2 statistic. The reader is referred to Phillips (1984b and 1984c) for a complete discussion and for algebraic details.

4.5. *The Null Distribution When* $q=1$

Setting $m=0$ in (30) we immediately obtain the null density:

$$(33)\qquad \operatorname{pdf}(w) = \frac{w^{-1/2} N^{1/2}}{2\Gamma(\frac{1}{2})^2} \int_0^\infty t^{-1/2} \cdot \det[I+(t+(w/N))Q]^{-N/2} \operatorname{tr}\{Q[I+(t+(w/N))Q]^{-1}\}\, dt.$$

This expression may also be found from the general formula (32) for the null density and from the alternate formula (30) given earlier for the non-null distribution when $q=1$.

Formula (33) may be used to compute the exact density by means of a simple unidimensional numerical integration. Recall that the matrix $Q=\sum_1^p \bar{l}_i\bar{l}_i'$ and the elements of $\bar{l}=Kl$ depend on the restriction vector d, the sample matrix $M=(X'X)^{-1}$, and the nuisance parameter Ω. Thus, in general, (33) depends on these parameters also.

The most important exception occurs when $n=1$ and (33) reduces to:

$$\begin{aligned}\text{pdf}(w) &= \frac{w^{-1/2}N^{1/2}}{2\Gamma(\frac{1}{2})^2(1+w/N)^{N/2+1}}\int_0^\infty t^{-1/2}\left(1+\frac{t}{1+w/N}\right)^{-(N/2)1} dt \\ &= \frac{N^{N/2}w^{1/2-1}}{B(1/2, N/2)(N+w)^{(N+1)/2}} \\ &\equiv F_{1,N}\end{aligned}$$

as otherwise expected.

4.6. *Asymptotic Theory*

As a first approximation to the exact density (19) in the general case we make the replacement

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

which is appropriate in the neighborhood of $X=0$. Formula (19) simplifies immediately under this approximation to:

$$\begin{aligned}\text{pdf}(w) &= \frac{\exp(-m'm/2)w^{q/2-1}}{N^{q/2}\Gamma(q/2)}\sum_{j,k}\frac{(-1)^j(1/2)^k(w/N)^f}{j!k!(1/2)_k C_f(I_q)} \\ &\quad \cdot [(N/2)^{q/2+f}C_f^{j,k}(I_q, mm')] \\ &= \frac{\exp(-m'm/2)w^{q/2-1}}{2^{q/2}\Gamma(q/2)}\sum_{j=0}^\infty \frac{(-w/2)^j}{j!}\sum_{k=0}^\infty \frac{(m'm/2)^k(w/2)^k}{k!(q/2)_k} \\ &= \frac{\exp(-m'm/2)w^{q/2-1}e^{-w/2}}{2^{q/2}\Gamma(q/2)}\,{}_0F_1\left(\frac{q}{2}; \frac{m'm}{2}\left(\frac{w}{2}\right)\right) \\ &\equiv \chi_q^2(\delta^2)\end{aligned}$$

where the noncentrality parameter is

$$\delta^2 = m'm = b'V^{-1}b = (D\,\text{vec}\,A-d)'(D(\Omega\otimes M)D')^{-1}(D\,\text{vec}\,A-d).$$

This provides us with a simple and elegant method for deducing the conventional asymptotic distribution of the Wald statistic from the general formula (19).

The approach may be made rigorous by noting that in (19) X is a dummy variable and we may transform $X \to NX = Z$ giving

$$\text{pdf}(w) = \frac{\exp(-m'm/2)w^{q/2-1}}{\Gamma(q/2)} \sum_{j,k} \frac{(-1)^j(1/2)^k w^f}{j!k!(1/2)_k C_{(f)}(I_q)}$$
$$\cdot [\det(L(\partial Z \otimes I)L')^{1/2}$$
$$\cdot C_f^{j,k}(L(\partial Z \otimes I)L', L(\partial Z \otimes I)L'mm') \det(I - Z/N)^{-N/2}]_{Z=0}.$$

Taking the first order term in the expansion of $\det(I - Z/N)^{-N/2}$ as $N \uparrow \infty$ we then obtain the asymptotic $\chi_q^2(\delta^2)$ distribution as shown above.

We remark that higher order asymptotics may be obtained by an obvious extension of this approach. The details will be reported elsewhere.

5. CONCLUSION

This paper provides a comprehensive and unified treatment of the distribution of the Wald statistic for testing general linear hypotheses in the multivariate linear model. All presently known special case distributions of this statistic may be obtained from our general formulae (19) and (20). This includes the well known formulae for the noncentral distributions of the regression F statistic and Hotelling's T^2. Other specializations, such as that of Hotelling's generalized T_0^2 statistic (Constantine, 1966), are discussed in detail in two related papers by the author (Phillips, 1984b and 1984c).

Also of interest is the fact that our general expression (19) yields in a simple and elegant way the asymptotic noncentral χ^2 approximation to the distribution of W. This approach to the asymptotic distribution is quite rigorous. Moreover, it may be extended to provide a simple algorithm for the development of higher order asymptotic approximations. Thus, our results may be said to provide a unification of asymptotic theory, asymptotic expansions and exact finite sample theory for the Wald statistic. In this respect, the results and the methods of this article are quite different from those of conventional statistical distribution theory, wherein the various branches of asymptotic analysis and sampling distribution theory are usually quite distinct and involve quite different mathematical techniques.

We remark that the null distribution of the Wald statistic given by (32) depends, in general, on the elements of the matrix L and, thereby, on the sample precision matrix $M = (X'X)^{-1}$, the restriction coefficient matrix D, and the error covariance matrix Ω. Only in very special cases, such as when D has the Kronecker structure $D = D_1 \otimes D_2$, will the density be independent of these parameters. Since M and D are known while Ω is not, it is the latter matrix of nuisance parameters which presents difficulties for the use of (32) in empirical work. This is a manifestation of a widely occurring problem in econometric tests. Of course, the use of conventional first order asymptotics in empirical work bypasses the problem entirely because, under conditions which are assumed for the application of this theory, $\Omega^* \rightarrow_p \Omega$ as $T \uparrow \infty$ and the null distribution of the statistic W is asymptotically χ_q^2 and independent of Ω. However, the use of first order asymptotics brings with it another set of problems (see Phillips (1983) for a detailed discussion). One of the most important of these is that, in neglecting the effect of the use of an estimated value of Ω in the computed value of the statistic W, inferences based upon asymptotic critical values are often seriously biased towards rejection of the null. Some idea of the extent of the bias involved in the use of asymptotic formulae in the present context is given in the experimental results of Meisner (1979) and recent numerical computations by the author (Phillips, 1984a). These studies reveal that the asymptotic bias appears to be much larger than was earlier thought and can lead to serious distortions, particularly when the ratio q/T is not small.

An alternative procedure, which allows for the presence of nuisance parameters, is to employ Edgeworth corrected critical values in statistical testing. Formulae for these corrections are now known (see Phillips (1984a) and Rothenberg (1984)) and, from the computations reported in Phillips (1984a), they do appear to yield improvements over first order asymptotics. Practical implementation of these corrections, however, requires the use of consistent estimates of the nuisance parameters. The results of the present paper should be useful in calibrating the accuracy of such Edgeworth corrections. The most important special case is that of a test of a single restriction, and here, numerical computation of the exact distribution can be readily achieved using (33).

Dept. of Economics, Yale University, New Haven, CT 06520, U.S.A.

Manuscript received August, 1984, final revision received August, 1985

REFERENCES

CONSTANTINE, A G. (1963): "Some Noncentral Distribution Problems in Multivariate Analysis," *Annals of Mathematical Statistics*, 34, 1270-1285.

——— (1966): "The Distribution of Hotelling's Generalized T_0^2," *Annals of Mathematical Statistics*, 37, 215-225.

DAVIS, A. W. (1979): "Invariant Polynomials with Two Matrix Arguments Extending to Zonal Polynomials. Applications to Multivariate Distribution Theory," *Annals of the Institute of Statistical Mathematics*, 31, 465-485.

——— (1980). "Invariant Polynomials With Two Matrix Arguments Extending the Zonal Polynomials," in P. R. Krishnaiah, Ed., *Multivariate Analysis*. Amsterdam: North-Holland, pp. 287-299.

JAMES, A. T. (1964): "Distributions of Matrix Variates and Latent Roots Derived from Normal Samples," *Annals of Mathematical Statistics*, 35, 475-501.

MEISNER, J. F. (1979): "The Sad Fate of the Asymptotic Slutsky Symmetry Test for Large Systems," *Economics Letters*, 2, 231-233.

MUIRHEAD, R. (1982): *Aspects of Multivariate Statistical Theory*. New York: Wiley.

PHILLIPS, P. C. B. (1983): "Exact Small Sample Theory in the Simultaneous Equations Model," in *Handbook of Econometrics*, ed by M. D. Intriligator and Z. Griliches. Amsterdam: North-Holland.

——— (1984a): "Finite Sample Econometrics Using ERA's," *Journal of the Japan Statistical Society*, 14, 107-124.

——— (1984b): "The Exact Distribution of the Wald Statistic," Cowles Foundation Discussion Paper No. 722, Yale University.

——— (1984c). "An Everywhere Convergent Series Representation of the Distribution of Hotelling's Generalized T_0^2," Cowles Foundation Discussion Paper No 723, Yale University.

——— (1985): "The Exact Distribution of the SUR Estimator," *Econometrica*, 53, 745-756.

PHILLIPS, P. C B., AND S. OULIARIS (1984) "The Exact Distribution of the Wald Statistic: The Noncentral Case," Cowles Foundation Discussion Paper No. 731, Yale University.

ROTHENBERG, T. J. (1984): "Hypothesis Testing in Linear Models When the Error Covariance Matrix is Nonscalar," *Econometrica*, 52, 827-842.

TITCHMARSH, E. C. (1939): *The Theory of Functions*. London: Oxford University Press.