Econometrica, Vol. 54, No. 5 (September, 1986), 1243-1248

NOTES AND COMMENTS

WALD CRITERIA FOR JOINTLY TESTING EQUALITY AND INEQUALITY RESTRICTIONS

BY DAVID A. KODDE AND FRANZ C. PALM[1]

1. INTRODUCTION

HYPOTHESES IN ECONOMICS are usually formulated in terms of constraints on the parameters of a model. They take the form of equality and/or inequality restrictions which are then to be jointly tested. For instance, the homogeneity of degree zero of a demand equation implies that the price and income elasticities add up to zero, whereas the negativity of the substitution matrix in consumer demand theory requires that all latent roots of the substitution matrix be nonpositive. Procedures for testing a set of inequality constraints have been studied by several authors. We refer the interested reader to Gouriéroux et al. (1982) and Perlman (1969) and the references therein. Joint testing of equality and inequality restrictions received little attention in the literature.

In this paper, we propose a large sample Wald test for sets of equality and inequality constraints on the parameters of a model. The null or the alternative hypothesis may be subject to inequality constraints. The computation of the test statistic is relatively simple. As it is difficult in the presence of inequality constraints to derive the (asymptotic) distribution of the test statistic, we give bounds for the critical value (under the null hypothesis). When these bounds are sufficient, the test can be straightforwardly applied.

The paper is organized as follows. In Section 2, we show how the Wald test applies to sets of equality and inequality restrictions, and we present the large sample distribution of the test under the null hypothesis. Upper and lower bound critical values for the joint test of equality and inequality restrictions are given in Section 3. Section 4 contains some concluding remarks. In the Appendix, we derive the large sample distribution of the test statistic under the null hypothesis.

2. THE WALD TEST FOR EQUALITY AND INEQUALITY RESTRICTIONS

We assume that the restrictions on a vector of parameters of interest θ are formulated in terms of p independent continuous functions $h(\theta)$, which are differentiable in some open neighborhood of the true parameters θ_0. The hypothesis to be tested is of the form

$$H_0\colon h_1(\theta)=0,\ h_2(\theta)=0 \quad \text{against} \quad H_1\colon h_1(\theta)\neq 0,\ h_2(\theta)\geqslant 0 \quad \text{(case 1)}, \tag{2.1}$$

or

$$H_0\colon h_1(\theta)=0,\ h_2(\theta)\geqslant 0 \quad \text{against} \quad H_1\colon h_1(\theta)\neq 0,\ h_2(\theta)\ngeqslant 0 \quad \text{(case 2)}. \tag{2.2}$$

The dimensions of the partition of $h(\theta)$ into $h_1(\theta)$ and $h_2(\theta)$ are q and $p-q$ respectively. When $q=0$, we assume that under H_0 there is at least one strict inequality in case 1. Under H_1 in case 2, the parameters are completely unrestricted. To present the Wald test for equality and inequality restrictions, we assume that θ can be consistently estimated by $\bar{\theta}$ such that the asymptotic distribution is given by

$$T^{1/2}(\bar{\theta}-\theta_0) \underset{A}{\sim} N(0,\Omega), \tag{2.3}$$

where Ω can be consistently estimated by $\bar{\Omega}$; T denotes the sample size.

We transform the functions of parameters $h(\theta)$ into new parameter vectors $\gamma=(\gamma_1', \gamma_2')'$ and $\bar{\gamma}=(\bar{\gamma}_1', \bar{\gamma}_2')'$ where

$$\gamma_i = T^{1/2}h_i(\theta) \quad \text{and} \quad \bar{\gamma}_i = T^{1/2}h_i(\bar{\theta}). \tag{2.4}$$

[1] The authors thank two anonymous referees for their useful comments.

Applying the mean value theorem we obtain the large sample covariance matrix of $\bar{\gamma}$:

(2.5) $$\Sigma=(\partial h/\partial\theta')\Omega(\partial h'/\partial\theta),$$

where the argument θ has been deleted for the sake of simplicity. The covariance matrix Σ can be consistently estimated by $\bar{\Sigma}$, evaluating expression (2.5) at $\bar{\theta}$ and $\bar{\Omega}$.

Let S_0 and S_1 denote the feasible space for γ under the null and the alternative hypothesis respectively. Under the hypotheses (2.1) and (2.2), the feasible spaces are convex in γ. Define the distance function in the metric of Σ of a vector x from the origin by

(2.6) $$\|x\|=x'\Sigma^{-1}x.$$

Except when explicitly stated otherwise, the distance will be measured in the metric of Σ. Usually the matrix Σ is not known and has to be substituted for by a consistent estimator such as, e.g., $\bar{\Sigma}$.

Let $\tilde{\gamma}$ and $\hat{\gamma}$ be the minimum distance estimators which satisfy the restrictions under H_0 and H_1 respectively;

(2.7) $$D_0=\|\bar{\gamma}-\tilde{\gamma}\|=\min_{\gamma\in S_0}\|\bar{\gamma}-\gamma\|$$

and

(2.8) $$D_1=\|\bar{\gamma}-\hat{\gamma}\|=\min_{\gamma\in S_1}\|\bar{\gamma}-\gamma\|.$$

D_i is the minimum distance from the data (i.e. $\bar{\gamma}$) to the closest feasible point under H_i, $i\in\{0,1\}$. In fact, $\tilde{\gamma}$ and $\hat{\gamma}$ are orthogonal projections of $\bar{\gamma}$ onto S_0 and S_1 respectively. Since S_0 and S_1 are convex, $\tilde{\gamma}$ and $\hat{\gamma}$ are uniquely determined. We define the Wald or distance test as

(2.9) $$D=D_0-D_1.$$

Using the properties of orthogonal projections, we straightforwardly get a useful alternative formulation of the Wald test

(2.10) $$D=\|\hat{\gamma}\|-\|\tilde{\gamma}\|.$$

As $S_0 C S_1$, D will always be nonnegative. If D exceeds the critical value, we reject the null hypothesis.

Now we discuss the two types of composite hypotheses (2.1) and (2.2) and we give the asymptotic distribution of D under H_0.

CASE 1: When H_0: $\gamma=(\gamma_1',\gamma_2')'=0$ and H_1: $\gamma_1\neq 0$, $\gamma_2\geqslant 0$, we have $\tilde{\gamma}=0$ and $\hat{\gamma}$ equals

(2.11) $$\hat{\gamma}_1=\bar{\gamma}_1+\Sigma_{12}\Sigma_{22}^{-1}(\hat{\gamma}_2-\bar{\gamma}_2),$$

where $\hat{\gamma}_2$ solves the program

(2.12) $$\min_{\gamma_2\geqslant 0}(\bar{\gamma}_2-\gamma_2)'\Sigma_{22}^{-1}(\bar{\gamma}_2-\gamma_2).$$

The partitioning of Σ corresponds to that of γ. The Wald test equals

(2.13) $$D=\|\hat{\gamma}\|=(\bar{\gamma}_1-\Sigma_{12}\Sigma_{22}^{-1}\bar{\gamma}_2)'(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(\bar{\gamma}_1-\Sigma_{12}\Sigma_{22}^{-1}\bar{\gamma}_2)+\hat{\gamma}_2'\Sigma_{22}^{-1}\hat{\gamma}_2.$$

The two terms on the right-hand side of (2.13) are asymptotically independent. The first term is $\chi^2(q)$ distributed. The distribution of the second term is a mixture of $(p-q)$ χ^2 distributions, so that the large sample distribution of (2.13) can be written as

(2.14) $$\Pr(D\geqslant c\,|\,\Sigma)=\sum_{i=0}^{p-q}\Pr[\chi^2(q+i)\geqslant c]w(p-q,i,\Sigma_{22}),$$

where $w(p-q,i,\Sigma_{22})$ denotes the probability that i of the $(p-q)$ elements of $\hat{\gamma}_2$ are strictly positive. The result of (2.14) has been obtained by Gouriéroux et al. (1982), Kudô (1963), Perlman (1969) in the case of maximum likelihood estimates and $q=0$. For $q=0$, $\chi^2(q)$ is the unit mass at the origin.

Notice that when $p=q$, D is the commonly used Wald criterion.

CASE 2: For H_0: $\gamma_1=0$, $\gamma_2\geqslant 0$ and H_1: $\gamma_1\neq 0$, $\gamma_2\ngeqslant 0$, we have $\hat{\gamma}=\bar{\gamma}$, $\tilde{\gamma}_1=0$ and $\tilde{\gamma}_2$ is the solution of

$$(2.15)\qquad \min_{\gamma_2\geqslant 0}(\bar{\gamma}_2-\gamma_2-\Sigma_{21}\Sigma_{11}^{-1}\bar{\gamma}_1)'(\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}(\bar{\gamma}_2-\gamma_2-\Sigma_{21}\Sigma_{11}^{-1}\bar{\gamma}_1).$$

The Wald test equals

$$(2.16)\qquad D=\|\bar{\gamma}-\tilde{\gamma}\|=\bar{\gamma}_1'\Sigma_{11}^{-1}\bar{\gamma}_1 + (\bar{\gamma}_2-\tilde{\gamma}_2-\Sigma_{21}\Sigma_{11}^{-1}\bar{\gamma}_1)'(\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}(\bar{\gamma}_2-\tilde{\gamma}_2-\Sigma_{21}\Sigma_{11}^{-1}\bar{\gamma}_1).$$

For the maximum under the null hypothesis the large sample distribution equals (see the Appendix)

$$(2.17)\qquad \sup_{\gamma_2\geqslant 0}\Pr(D\geqslant c\,|\,\Sigma)=\sum_{i=0}^{p-q}\Pr[\chi^2(p-i)\geqslant c]\,w(p-q,i,\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}),$$

with the weights w denoting the probability that i of the $p-q$ elements of $\tilde{\gamma}_2$ are strictly positive. The covariance matrix in w is the conditional covariance matrix of $\bar{\gamma}_2$ given $\bar{\gamma}_1$. After a minor modification the methods described above can be directly applied to test a set of equality or inequality constraints conditionally on equality restrictions.

Consider the following hypotheses H_0: $\gamma_2\geqslant 0$ and H_1: $\gamma_2\ngeqslant 0$, both subject to the restriction $\gamma_1=0$. Then $\tilde{\gamma}_1=\hat{\gamma}_1=0$ and $\tilde{\gamma}_2$ and $\hat{\gamma}_2$ are the solution of (2.15) subject to the restrictions $\gamma_2\geqslant 0$ and γ_2 being unrestricted respectively. With $\tilde{\gamma}_2$ substituted into (2.15), the distance test can be readily computed. Its large sample distribution is given in (2.17) with $p-i$ being replaced by $p-q-i$.

3. UPPER AND LOWER BOUND CRITICAL VALUES

The main difficulty when testing inequality restrictions consists in computing the weights w in the distribution of the test statistic. Kudô (1963) gives an analytical expression for the weights, which is not very tractable. Gouriéroux et al. (1982) propose numerical simulation to determine the weights. However, from Perlman (1969, Theorem 6.2) we can derive lower and upper bound critical values corresponding to a chosen significance level α.

For case 1, the lower and upper bound critical values are obtained by solving

$$(3.1)\qquad \begin{aligned}\alpha&=\inf_{\Sigma>0}\Pr(D\geqslant c\,|\,\Sigma)=\tfrac{1}{2}\Pr[\chi^2(q)\geqslant c]+\tfrac{1}{2}\Pr[\chi^2(q+1)\geqslant c]\quad\text{and}\\ \alpha&=\sup_{\Sigma>0}\Pr(D\geqslant c\,|\,\Sigma)=\tfrac{1}{2}\Pr[\chi^2(p-1)\geqslant c]+\tfrac{1}{2}\Pr[\chi^2(p)\geqslant c],\end{aligned}$$

respectively for c. Similarly for case 2, c is determined such that it corresponds to inf and sup $\Pr(D\geqslant c\,|\,\Sigma,\gamma=0)$ respectively substituted in (3.1). Notice the coincidence of upper and lower bound values when $p=q+1$. The result obtained by Gouriéroux et al. (1982) is a special case of (3.1) with the number of equality restrictions q being zero. When $q=0$ we have $\Pr[\chi^2(0)\geqslant c]=0$, for $c>0$.

Table I can be used to determine upper and lower bound critical values for a joint test of equality and inequality constraints for commonly used significance levels α. A lower bound c for the critical value is obtained by choosing a level α and setting df equal to $q+1$. For the upper bound c of the critical value, df is set equal to the sum of the number of equality and inequality restrictions p.

4. CONCLUDING REMARKS

We presented a large sample Wald test for jointly testing nonlinear equality and inequality constraints either under H_0 or H_1. The kind of hypotheses which we considered

TABLE I

UPPER AND LOWER BOUNDS FOR THE CRITICAL VALUE FOR JOINTLY TESTING EQUALITY AND INEQUALITY RESTRICTIONS[a]

df	α .25	.10	.05	.025	.01	.005	.001
1	0.455	1.642	2.706	3.841	5.412	6.635	9.500
2	2.090	3.808	5.138	6.483	8.273	9.634	12.810
3	3.475	5.528	7.045	8.542	10.501	11.971	15.357
4	4.776	7.094	8.761	10.384	12.483	14.045	17.612
5	6.031	8.574	10.371	12.103	14.325	15.968	19.696
6	7.257	9.998	11.911	13.742	16.074	17.791	21.666
7	8.461	11.383	13.401	15.321	17.755	19.540	23.551
8	9.648	12.737	14.853	16.856	19.384	21.232	25.370
9	10.823	14.067	16.274	18.354	20.972	22.879	27.133
10	11.987	15.377	17.670	19.824	22.525	24.488	28.856
11	13.142	16.670	19.045	21.268	24.049	26.065	30.542
12	14.289	17.949	20.410	22.691	25.549	27.616	32.196
13	15.430	19.216	21.742	24.096	27.026	29.143	33.823
14	16.566	20.472	23.069	25.484	28.485	30.649	35.425
15	17.696	21.718	24.384	26.856	29.927	32.136	37.005
16	18.824	22.956	25.689	28.219	31.353	33.607	38.566
17	19.943	24.186	26.983	29.569	32.766	35.063	40.109
18	21.060	25.409	28.268	30.908	34.167	36.505	41.636
19	22.174	26.625	29.545	32.237	35.556	37.935	43.148
20	23.285	27.835	30.814	33.557	36.935	39.353	44.646
21	24.394	29.040	32.077	34.869	38.304	40.761	46.133
22	25.499	30.240	33.333	36.173	39.664	42.158	47.607
23	26.602	31.436	34.583	37.470	41.016	43.547	49.071
24	27.703	32.627	35.827	38.761	42.360	44.927	50.524
25	28.801	33.813	37.066	40.045	43.696	46.299	51.986
26	29.898	34.996	38.301	41.324	45.026	47.663	53.403
27	30.992	36.176	39.531	42.597	46.349	49.020	54.830
28	32.085	37.352	40.756	43.865	47.667	50.371	56.248
29	33.176	38.524	41.977	45.128	48.978	51.715	57.660
30	34.266	39.694	43.194	46.387	50.284	53.054	59.064
31	35.354	40.861	44.408	47.641	51.585	54.386	60.461
32	36.440	42.025	45.618	48.891	52.881	55.713	61.852
33	37.525	43.186	46.825	50.137	54.172	57.035	63.237
34	38.609	44.345	48.029	51.379	55.459	58.352	64.616
35	39.691	45.501	49.229	52.618	56.742	59.665	65.989
36	40.773	46.655	50.427	53.853	58.020	60.973	67.357
37	41.853	47.808	51.622	55.085	59.295	62.276	68.720
38	42.932	48.957	52.814	56.313	60.566	63.576	70.078
39	44.010	50.105	54.003	57.539	61.833	64.871	71.432
40	45.087	51.251	55.190	58.762	63.097	66.163	72.780

[a] The values in the table are obtained by solving the equation $\alpha = \frac{1}{2}\Pr[\chi^2(df-1) \geq c] + \frac{1}{2}\Pr[\chi^2(df) \geq c]$ for c, given α and df.

often arise in empirical econometric work. The large sample distribution of the test statistic under the null hypothesis is a mixture of χ^2 distributions.

As in the case of testing equality constraints, the Wald test should be very useful when unrestricted asymptotically normal, consistent but not necessarily efficient estimates of the parameters θ can be easily obtained compared with estimation subject to nonlinear equality and inequality restrictions.

To avoid the computational problems involved in obtaining the asymptotic distribution of the Wald test, we derived upper and lower bound critical values, which should be useful in many applications.

To implement the test, the following steps have to be carried out: (i) From $\bar{\theta}$ and $\bar{\Omega}$, obtain $\bar{\gamma}$ and $\bar{\Sigma}$ and compute the test statistic D using quadratic programming techniques. (ii) Choose α and determine the upper and lower bounds from Table I. (iii) Reject H_0 when D exceeds the upper bound value; do not reject H_0 when D is smaller than the lower bound value. When the test is inconclusive, the weights w in the distribution can be determined numerically. Then D can be compared with the critical value corresponding to the selected size α.

Moreover, when $\bar{\gamma}$ is an efficient estimate of γ, the Wald or distance (D), likelihood ratio (LR) and the Kuhn-Tucker (KT) tests for equality and inequality constraints are asymptotically equivalent. Along the lines of the proof in Gouriéroux et al. (1982) a well-known ordering of these test for other types of models and hypotheses, $KT \leq LR \leq D$, is found to hold in the more general cases that we considered here (see Kodde and Palm (1984)). Also, the KT-test can be obtained as the coefficient of determination in a two-step procedure.

Economics Institute, Catholic University, Nijmegen, The Netherlands
and
Dept. of Economics, Free University, Amsterdam, The Netherlands

Manuscript received June, 1984; final revision received July, 1985.

APPENDIX

DERIVATION OF THE NULL DISTRIBUTION OF THE WALD TEST

In this Appendix, we outline the steps in the proof of the asymptotic distribution of D under H_0. We limit ourselves to case 2 with H_0: $\gamma_1 = 0$, $\gamma_2 \geq 0$ against H_1: $\gamma_1 \neq 0$, $\gamma_2 \ngeq 0$ for which $D = \|\bar{\gamma} - \tilde{\gamma}\|$ in (2.16).

The critical level c of the distance test in (2.16) is determined by

$$\text{(A.1)} \qquad \sup_{\gamma_2 \geq 0} \Pr[D \geq c \mid \Sigma] = \alpha$$

with α being the size of the test. As $\gamma_2 \geq 0$ is a convex cone, the left-hand side is maximal when $\gamma_2 = 0$ (see, e.g., Perlman (1969, Theorem 8.3)).

Consider a partition of $\tilde{\gamma}_2$ into subsets of $(p - q - i)$ zero and i strictly positive values denoted by $\tilde{\gamma}_{21}$ and $\tilde{\gamma}_{22}$ respectively. Since $\tilde{\gamma}_2$ minimizes (2.12), we can apply the lemma of Nüesch (1966) to get

$$\text{(A.2)} \qquad \Pr[\tilde{\gamma}_{21} = 0 \text{ and } \tilde{\gamma}_{22} > 0] = \Pr[\mu_2 > 0 \text{ and } \mu_1 \leq 0],$$

where $\mu_1 = B_{11}^{-1}\bar{\gamma}_{21}$ and $\mu_2 = \bar{\gamma}_{22} - B_{21}B_{11}^{-1}\bar{\gamma}_{21}$, a result which follows from the Kuhn-Tucker conditions for a minimum of (2.12). The matrices B_{11}, B_{21} are submatrices of $\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ and correspond to the partition of γ_2 into γ_{21} and γ_{22}. Using this result, the difference between $\bar{\gamma}$ and $\tilde{\gamma}$ equals

$$\bar{\gamma}_1 - \tilde{\gamma}_1 = \bar{\gamma}_1,$$

$$\bar{\gamma}_2 - \tilde{\gamma}_2 = (\bar{\gamma}_{21}', \bar{\gamma}_{21}' B_{11}^{-1} B_{12})',$$

and the value of the distance test conditionally on the event in (A.2) is a quadratic form in $(\bar{\gamma}_1', \bar{\gamma}_{21}')'$, which can be shown to be $\chi^2(p-i)$-distributed. The probability weights are found by summing over all events of the form (A.2) with i positive elements of $\tilde{\gamma}_2$.

For a given significance level α, the critical value c is found by solving

$$\text{(A.3)} \qquad \sum_{i=0}^{p-q} \Pr[\chi^2(p-i) \geq c]\, w(p-q, i, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) = \alpha.$$

REFERENCES

GOURIÉROUX, CH., A. HOLLY, AND A. MONFORT (1982): "Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters," *Econometrica*, 50, 63-80.

KODDE, D. A., AND F. C. PALM (1984): "Testing Equality and Inequality Restrictions—A Distance Test," Catholic University Nijmegen, Research Paper.

KUDÔ, A. (1963): "A Multivariate Analogue of the One-sided-Test," *Biometrika*, 50, 403-418.

NÜESCH, P. E. (1966): "On the Problem of Testing Location in Multivariate Analysis," *Annals of Mathematical Statistics*, 37, 113-119.

PERLMAN, M. D. (1969): "One Sided Problems in Multivariate Analysis," *Annals of Mathematical Statistics*, 40, 549-567.