MONASH

6/94

MONASH
UNIVERSITY

AUSTRALIA

ONE-SIDED HYPOTHESIS TESTING IN ECONOMETRICS: A SURVEY

Ping X. Wu and Maxwell L. King

Working Paper No. 6/94

May 1994

DEPARTMENT OF ECONOMETRICS

42

ISSN 1032-3813

ISBN 0 7326 0395 1

ONE-SIDED HYPOTHESIS TESTING IN ECONOMETRICS: A SURVEY

Ping X. Wu and Maxwell L. King

Working Paper No. 6/94

May 1994

DEPARTMENT OF ECONOMETRICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

ONE-SIDED HYPOTHESIS TESTING IN ECONOMETRICS: A SURVEY

Ping X. Wu

Department of Economics

University of Melbourne

Parkville, Victoria 3052, Australia

and

Maxwell L. King

Department of Econometrics

Monash University

Clayton, Victoria 3168, Australia

Abstract

Any model is accompanied by a set of assumptions. These assumptions are based either on the underlying theory of the phenomena being modelled or on stylized statistical evidence, or more commonly on both. These, along with functional considerations such as variances being positive, often imply that values of some parameters characterizing a model are restricted to one side of a point in the parameter space. This information can be used to improve the power of hypothesis testing procedures. In this paper, we discuss some recent developments on testing against such one-sided alternative hypotheses with particular emphasis on the econometrics literature. The focus is on two main approaches: that based on maximum likelihood estimation and that based on local power optimization facilitated by the generalized Neyman-Pearson lemma. Both single parameter and multi-parameter testing problems are considered.

Key Words

Kuhn-Tucker tests; likelihood ratio tests; linear regression; locally best tests; partially one-sided testing; point optimal tests.

1. Introduction

The practice of econometrics involves an integration of economic theories and mathematical statistical methods. Economic theories usually provide initial models of the phenomena of interest. Statistical methods are then used to estimate and assess the adequacy of these models. One feature of the interaction between economic theories and statistical methods is that many economic concepts are captured by parameters in models and the range of values these parameters may take are therefore typically restricted. For example, the parameter which characterizes the speed of adjustment in a partial adjustment model is believed to be non-negative and less than one.

Inequality restricted parameters can also arise from stylized statistical evidence. For example, in linear regression models of undifferenced macroeconomic time series, a standard phenomenon is that residual series typically exhibit, if any, positive autocorrelation. It is then natural to model error terms in such models with parameters characterizing low lag autocorrelations that are non-negative.

There are also situations where the parameter space is restricted for functional reasons. An obvious and important example is that variance parameters cannot be negative. Variance parameters play key roles both in traditional variance-component models and random-coefficient models. In these situations, one is typically interested in whether variances are zero (so that the corresponding random terms disappear from the model). This leads to testing problems in which the null hypothesis is on the boundary of the parameter space which consists of non-negative parameters.

It is expected that statistical methods that take explicit account of this kind of non-sample information yield more efficient estimates of the parameters. See for example Thompson (1982) in the case of two regressors and non-negative restrictions on the two coefficients in the linear regression model. More detailed discussion, including pre-test procedures, is contained in Judge and Yancey (1986). In hypothesis testing, using non-sample information should result in greater power. For example, in the context of testing for a linear form of heteroscedasticity in the simple time trend regression with 20 observations, King and Evans (1984) observed an increase in power ranging from 39% to 63%

when a one-sided Lagrange multiplier (LM) test is used instead of the familiar two-sided test. Thus, statistical methods that incorporate this information are highly desirable. It is therefore not surprising to see a significant increase in the econometrics and statistics literature on testing against one-sided alternatives. In this paper we review this literature with an emphasis on newly developed methods.

There seem to be two main approaches to the testing problem. One is based on local optimization of power through the use of the generalized Neyman-Pearson lemma. A variation of this approach is to construct a test that is most powerful among all the tests at a chosen point in the alternative parameter space. This is called a point optimal test. An alternative is to construct tests by maximizing power locally at the null hypothesis. These tests are called locally best (LB) tests. The second approach is to construct one-sided versions of conventional likelihood ratio (LR), Wald and LM tests. We shall call these likelihood-based tests as they involve maximization of the likelihood function with respect to parameters subject to the inequality restrictions. Estimation is then a part of the test procedure.

The plan of this paper is as follows. Section 2 reviews LB tests followed by point optimal tests. We observe that the majority of applications of these procedures involve tests of regression disturbances. Section 3 deals with asymptotic likelihood-based tests and begins with discussion of one-sided LR, Wald and LM tests of one parameter. This leads on to multi-parameter one-sided testing and partially one-sided testing where likelihood-based tests have asymptotic null distributions that are probability mixtures of various chi-squared distributions. The derivation of the probability weights in these asymptotic distributions is also discussed in detail. Section 3 includes a brief survey of one-sided testing of linear regression coefficients and closes with a review of the literature on exact and asymptotic locally most mean powerful (LMMP) tests. Concluding remarks are made in the final section.

2. Optimal Tests

For any testing problem we would obviously like to use a uniformly most powerful (UMP) test. Unfortunately such tests rarely exist. Cox and Hinkley (1974, p.102) discuss three alternative approaches to test construction in the absence of a UMP test. They are: (i) using a test which maximizes power at a "somewhat arbitrary 'typical' point" in the alternative parameter space, (ii) removing this arbitrariness by choosing the point to be close to the null hypothesis which leads to the LB test and (iii) choosing the test which maximizes some weighted average of power over the alternative parameter space. In this section, we review the literature on one-sided testing from the point of view of Cox and Hinkley's three alternatives. Historically, option (ii) has been the most popular so we will consider it first. Option (i) is sometimes known as the point optimal solution and is the subject of subsection 2.2. There is a fledgling literature on option (iii) which is not considered here; see Andrews and Ploberger (1994) and Andrews, Lee and Ploberger (1994).

2.1 Locally Best Tests

The LB solution is reasonably well accepted in the testing literature. Suppose we wish to test $H_0 : \theta = \theta_0$ against $H_a^+ : \theta > \theta_0$ based on the $n \times 1$ random vector x which has probability density function $f(x \mid \theta)$ where θ is $p \times 1$. Here $\theta > \theta_0$ denotes $\theta_i \geq \theta_{0i}$ for each i with inequality for at least one i so that $\theta \neq \theta_0$.

When $p = 1$, the LB test can be constructed simply as a direct consequence of the Neyman-Pearson Lemma as follows. The most powerful test of H_0 against $H_a : \theta = \theta_0 + \Delta$ where $\Delta \neq 0$ involves rejecting H_0 for large values of

$$r(x) = f(x \mid \theta_0 + \Delta)/f(x \mid \theta_0)$$

or equivalently

$$\log r(x) = \log f(x \mid \theta_0 + \Delta) - \log f(x \mid \theta_0). \tag{2.1}$$

If we replace $\log f(x \mid \theta_0 + \Delta)$ in (2.1) with its Taylor's series expansion with respect to θ, we get

$$\log r(x) = \Delta \frac{\partial \log f(x \mid \theta)}{\partial \theta}\bigg|_{\theta=\theta_0} + \frac{\Delta^2}{2!}\frac{\partial^2 \log f(x \mid \theta)}{\partial \theta^2}\bigg|_{\theta=\theta_0}$$
$$+ \frac{\Delta^3}{3!}\frac{\partial^3 \log f(x \mid \theta)}{\partial \theta^3}\bigg|_{\theta=\theta_0} + \cdots. \tag{2.2}$$

By letting Δ tend to zero and dropping those terms involving squares and higher powers of Δ, we see that the LB test of H_0 is given by critical regions of the form

$$\left.\frac{\partial \log f(x \mid \theta)}{\partial \theta}\right|_{\theta=\theta_0} > c \tag{2.3}$$

where c is the critical value; i.e. we reject H_0 for large values of the score evaluated at H_0.

A number of other results flow from (2.2). If we wish to test H_0 against $H_a^- : \theta < \theta_0$, then the inequality in (2.3) is reversed. Occasionally, the score evaluated at θ_0 is zero in which case (2.3) no longer constitutes a test. We see immediately from (2.2) that the LB test against $H_a^+ : \theta > \theta_0$ is now based on

$$\left.\frac{\partial^i \log f(x \mid \theta)}{\partial \theta^i}\right|_{\theta=\theta_0} > c \tag{2.4}$$

where i is the smallest positive integer for which this term is nonzero. It is interesting to note that the LB test against $H_a^- : \theta < \theta_0$ only has the inequality sign in (2.4) reversed when i is an odd integer. This means for even values of i, one-sided and two-sided tests coincide.

The test based on (2.3) is equivalent to the one-sided LM test based on the square-root of the standard LM test statistic with the square-root taking the sign of the score evaluated at H_0. This test maximizes the slope of the power curve at H_0 in the one-sided direction within the class of all tests of the same size. It is often also called the locally most powerful (LMP) test because it is constructed to maximize power in the neighbourhood of H_0. For reasons of continuity, all powers at H_0 will equal size, so it is difficult to see that this test is "most powerful" locally at H_0. We prefer the term locally best because we know the test's power curve has the best slope at the null hypothesis.

There are few applications of this test to problems in which there is truly only one parameter. One such example is SenGupta's (1987) test for nonzero values of the equicorrelation coefficient of a standard symmetric multivariate normal distribution. Most typically, applications involve the use of invariance, similarity or marginal likelihood arguments to effectively eliminate nuisance parameters.

The most popular application involves testing $H_0 : \theta = 0$ against $H_a : \theta > 0$ when $p = 1$ in the linear regression model

$$y = X\beta + u, \tag{2.5}$$

in which

$$u \sim N(0, \sigma^2\Omega(\theta)), \tag{2.6}$$

where y is $n \times 1, X$ is an $n \times k$ nonstochastic matrix of rank $k < n, \Omega(\theta)$ is a positive definite matrix for the θ values of interest and the $k \times 1$ vector β and the scalar σ^2 are unknown nuisance parameters. Without loss of generality, we assume that $\Omega(0) = I_n$ because if this is not the case, (2.5) can be transformed by premultiplying by $\Omega(0)^{-1/2}$ to make it true. This testing problem is invariant to transformations of the form

$$y \to \gamma_0 y + X\gamma$$

where γ_0 is a positive scalar and γ is a $k \times 1$ vector. King and Hillier (1985) show that the LB invariant (LBI) test of H_0 against H_a^+ has critical regions of the form

$$z'Az/z'z > c \tag{2.7}$$

where z is the OLS residual vector from (2.5) and

$$A = \left.\frac{\partial\Omega(\theta)}{\partial\theta}\right|_{\theta=0} = -\left.\frac{\partial\Omega(\theta)^{-1}}{\partial\theta}\right|_{\theta=0}.$$

The most noteworthy application of this result is the Durbin-Watson (1950, 1951, 1971) test which is approximately LBI against first-order autoregressive (AR(1)) disturbances. The true LBI test for this problem has been discussed by King (1981a) and also found to be LBI against first-order moving average (MA(1)) disturbances (King (1983b)); the sum of white noise and independent AR(1) disturbances (King (1982)) and particular forms of spatial first-order autocorrelated disturbances (King and Evans (1985)). Other applications include testing random walk disturbances (Sargan and Bhargava (1983) and Dufour and King (1991)); testing simple higher-order autoregressive disturbances (Wallis (1972), Vinod (1973) and King (1984)); testing for heteroscedasticity (Evans and King (1988)); testing for random regression coefficients (King (1987c), Nabeya and Tanaka (1988), Leybourne and McCabe (1989) and Nyblom (1989)) and testing for spatial autocorrelation (King (1981b)). An example of a testing problem in which the score evaluated at H_0 is zero, is testing the null hypothesis of a unit root in MA(1) regression disturbances; see for example Tanaka and Satchell (1989).

The form of (2.7) means that critical values can be found using methods developed for the Durbin-Watson test statistic. These are based on Imhof's (1961) algorithm, see Koerts and Abrahamse (1969), Davies (1980), King (1987b), Shively, Ansley and Kohn (1990) and Ansley, Kohn and Shively (1992). An alternative approach involves using Tanaka's (1990) Fredholm technique of inverting the characteristic function of the test statistic in order to find critical values. Methods of approximating critical values of (2.7) have been investigated by Evans and King (1985a).

Recently, Hillier (1987) observed that there is an exact equivalence between similar tests and invariant tests of $H_0 : \theta = 0$ in the context of (2.5) and (2.6) for any value of p. Ara and King (1993) have shown a similar equivalence between tests based on a maximal invariant (i.e., invariant tests) and tests based on the marginal likelihood.

A major question is how LB tests can be generalized from one parameter ($p = 1$) to many parameters ($p > 1$). For testing H_0 against $H_a : \theta \neq 0$, Neyman and Pearson (1938) suggested the use of type C LB unbiased (LBU) regions when $p = 2$. These tests have constant power in the neighbourhood of $\theta = \theta_0$ along a given family of concentric ellipses. Isaacson's (1951) type D regions maximize the Gaussian curvature of the power function at $\theta = \theta_0$ but are extremely difficult tests to apply in practice. One version of the type D test which admits nuisance parameters is also mentioned by Lehmann (1986) who refers to it as a type E test. As implied by Lehmann, a type E test can be viewed as a type D test applied to the density of a maximal invariant which does not depend on nuisance parameters. Locally most mean powerful unbiased (LMMPU) tests were introduced by SenGupta and Vermeire (1986). They maximize the mean curvature of the power hypersurface in the neighbourhood of $\theta = 0$ within the class of unbiased tests. Their critical regions are of the form

$$\sum_{i=1}^{p} \frac{\partial^2 f(x \mid \theta)}{\partial \theta_i^2}\bigg|_{\theta=\theta_0} > c_0 f(x \mid \theta)\mid_{\theta=\theta_0} + \sum_{i=1}^{p} c_i \frac{\partial f(x \mid \theta)}{\partial \theta_i}\bigg|_{\theta=\theta_0}$$

where the constants $c_0, c_1, \ldots, c_p$ are chosen so that the critical region is locally unbiased and of the required size.

Of course our interest is in testing $H_0 : \theta = \theta_0$ against $H_a^+ : \theta > \theta_0$ when $p > 1$. To understand the complexities involved in constructing a LB test in this case, observe that the equivalent multivariate expansion to (2.2) is

$$\log r(x) = \Delta' \frac{\partial \log f(x \mid \theta)}{\partial \theta}\bigg|_{\theta=\theta_0} + \frac{1}{2!}\Delta' \frac{\partial^2 \log f(x \mid \theta)}{\partial \theta^2}\bigg|_{\theta=\theta_0} \Delta + \cdots \tag{2.8}$$

where Δ is now a $p \times 1$ vector and rejecting for large values of $\log r(x)$ is the most powerful test of $H_0 : \theta = \theta_0$ against $H_a : \theta = \theta_0 + \Delta$. Write $\Delta = \delta w$ where w is a vector of unit length (i.e., such that $w'w = 1$) representing direction from θ_0 and $\delta > 0$ is distance from θ_0. Then as δ tends to zero we find the LB test is based on critical regions of the form

$$w' \frac{\partial \log f(x \mid \theta)}{\partial \theta}\bigg|_{\theta=\theta_0} = \sum_{i=1}^{p} w_i \frac{\partial \log f(x \mid \theta)}{\partial \theta_i}\bigg|_{\theta=\theta_0} > c \tag{2.9}$$

which we can see depends on the direction w.

One possibility is that each of the components

$$\frac{\partial \log f(x \mid \theta)}{\partial \theta_i}\bigg|_{\theta=\theta_0} > c_i,$$

treated as an individual test, results in the same class of critical regions, for $i = 1, \ldots, p$, so that (2.9) corresponds to the same class of critical regions no matter which w value is chosen. This results in a test that is LB in all directions from H_0 under H_a^+. Neyman and Scott (1967) called this property "robustness of optimality" while King and Evans (1988) call such tests uniformly LB (ULB). Like UMP tests, they seem to rarely exist. King and Evans identified that the true LBI test for AR(1) disturbances in (2.5) (a modified version of the Durbin-Watson test) is also ULB invariant against the sum of q independent ARMA(1,1) disturbance processes, certain spatial autocorrelation disturbance processes involving up to four parameters and a stochastic cycle disturbance model which is a special case of an ARMA(2,1) process.

The next question is what to do when a ULB test does not exist. One option is to choose a w value that corresponds to a central direction and then apply (2.9). King and Wu (1993) advocated the application of SenGupta and Vermeire's optimality criteria of maximizing the mean curvature of the power curve at H_0 in all directions under $H_a^+ : \theta > \theta_0$. They showed the LMMP test results in critical regions of the form

$$\sum_{i=1}^{p} \frac{\partial \log f(x \mid \theta)}{\partial \theta_i}\bigg|_{\theta=\theta_0} > c \tag{2.10}$$

which one can readily see from (2.9) is equivalent to the LB test in the direction of $\theta_1 = \theta_2 = \cdots = \theta_p > 0$. Thus maximizing the mean curvature of the power function at H_0 is achieved by maximizing its slope in the most central direction.

The test is based on the sum of scores which makes it easy to apply as we shall see in subsection 3.8.

2.2 Point Optimal Tests

For any testing problem and a given significance level, in theory there is always at least one test procedure that optimizes power at a predetermined point under the alternative hypothesis. Such a test is a point optimal test. It obviously will have excellent relative power around the point at which power is optimized. It can also have good power away from this point. For example, if a UMP test exists, then the point optimal test will be UMP. Examples have also been found of point optimal tests that are approximately UMP (Shively (1988b)) or are UMP over a subset of the alternative parameter space (King and Smith (1986)). On the other hand, in some situations, the relative power performance of point optimal tests can drop away quickly as one moves away from the point at which power is optimized (see the tests for unit roots and explosive disturbance processes investigated by Dufour and King (1991)).

While point optimal tests are not suited to every testing problem, they appear to work best in situations in which the alternative hypothesis is severely restricted which is the case for one-sided testing problems. It also helps if the null hypothesis can be reduced to a simple hypothesis through invariance, similarity or marginal likelihood arguments.

For our problem of testing H_0 against H_a^+ based on the random vector x with probability density function $f(x \mid \theta)$ where θ is $p \times 1$, the Neyman-Pearson Lemma implies that rejecting H_0 for large values of $f(x \mid \theta_1)/f(x \mid \theta_0)$ is most powerful at $\theta = \theta_1 > \theta_0$. This provides a straightforward method of test construction in the case of a simple null hypothesis.

Now consider the more general problem of testing

$$H_0 : x \text{ has density } g(x \mid \omega)$$

where ω is a $q \times 1$ vector of parameters restricted to the set Ω against

$$H_a : x \text{ has density } f(x \mid \theta)$$

where θ is a $p \times 1$ vector of parameters restricted to the set Θ. This general testing problem includes both nested and nonnested problems as special cases. We assume all knowledge about the range of possible parameter values such as one-sided information is explicitly included in the formulation of Ω and Θ. If Ω is restricted to one point, ω_1, then the test which optimizes power at $\theta = \theta_1$ involves rejecting H_0 for

$$r(\omega_1, \theta_1) = f(x \mid \theta_1)/g(x \mid \omega_1) > c^*. \tag{2.11}$$

King (1987a) explored the question of whether a test of the form of (2.11) still optimizes power at $\theta = \theta_1$ when Ω is broadened from a single point to a set of points. He observed that this required c^* to remain the critical value when the null hypothesis changes from $\omega = \omega_1$ to $\omega \in \Omega$. This is the case if ω_1 and c^* can be chosen such that

$$\Pr\left[r(\omega_1, \theta_1) > c^* \mid x \text{ has density } g(x \mid \omega_1)\right] = \alpha$$

and

$$\sup_{\omega \in \Omega} \Pr\left[r(\omega_1, \theta_1) > c^* \mid x \text{ has density } g(x \mid \omega)\right] = \alpha. \tag{2.12}$$

In other words, we need to be able to choose ω_1 such that the maximum size of the test occurs at $\omega = \omega_1$.

In situations in which such an ω_1 value cannot be found, King (1987a) suggests the construction of approximately point optimal tests. This involves searching for that ω_1 value which minimizes

$$\alpha - \Pr\left[r(\omega_1, \theta_1) > c^* \mid x \text{ has density } g(x \mid \omega_1)\right]$$

subject to (2.12) holding.

In order to make a point optimal (or approximately point optimal) test operational, the point at which power is to be optimized must be chosen. Suggestions for choosing this point have included the end-point of the range of θ values (Berenblut and Webb (1973)), mid-points of this range (King (1983a, 1983b, 1985a)) and choosing θ_1 so that the power at θ_1 has a predetermined value such as 0.5 or 0.8 (King (1985b, 1989), Shively (1988b), and Bhatti and King (1990)). If when $p = 1, H_0$ is a simple hypothesis and the power envelope is monotonic increasing as θ gets further from H_0, then the point optimal test with θ_1 chosen using the latter rule gives rise to a beta-optimal test. Davies (1969) defined the beta-optimal test to be that test whose power reaches a predetermined level such as 0.5 or 0.8 most quickly as one moves away from H_0 in the alternative hypothesis parameter space. Bhatti and King (1990) illustrate this in the case of testing for a non-zero equicorrelation coefficient in a standard symmetric multivariate normal distribution.

A variation on this method of choosing θ_1 has been suggested by King (1989) for the problem of testing for simple AR(4) disturbances in (2.5) in the presence of AR(1) disturbances. Invariance allows the problem to be reduced to one in which

$$\Omega = \{\rho_1 : 0 \leq \rho_1 < 1\}$$

and

$$\Theta = \{\rho_1, \rho_4 : 0 \leq \rho_1 < 1, 0 \leq \rho_4 < 1\}$$

where ρ_1 and ρ_4 are the correlation coefficients of the AR(1) and simple AR(4) processes, respectively. If $(\rho_{11}, \rho_{41})'$ denotes the point at which power is to be optimized, King suggested it be chosen to make the minimum power on $\rho_4 = \rho_{41}$ equal to 0.5. Other suggestions for the choice of θ_1 value are reviewed by King (1987a).

The majority of successful applications of the point optimal approach have been in the area of testing θ in the context of the linear regression (2.5) and (2.6). Noteworthy examples include testing for AR(1) disturbances (Berenblut and Webb (1973), Fraser, Guttman and Styan (1976), King (1985a) and Dufour and King (1991)); testing for MA(1) disturbances (King (1983b, 1985b)); testing for heteroscedasticity (Evans and King (1985b, 1988)), testing for random walk disturbances (Sargan and Bhargava (1983)); testing for random coefficients (Franzini and Harvey (1983), Nyblom (1986), Shively (1986, 1988a, 1988b), Brooks and King (1994) and Rahman and King (1994)); testing for moving average unit roots in ARIMA models (Saikkonen and Luukkonen (1993)); testing for block effects in regression disturbances (Bhatti (1992) and Bhatti and King (1993)); and non-nested tests of autoregressive versus moving average disturbances (King (1983a, 1987a), Silvapulle (1991) and Silvapulle and King (1991, 1993)). King and Smith (1986) applied the point optimal approach to joint one-sided testing of regression coefficients. They showed that the one-sided t-test applied to a weighted sum of the associated regressors is UMPI along the ray in the alternative parameter space whose direction is defined by the weights. An alternative formulation of this test is given by Hillier (1986) who suggested the weights be chosen to maximize minimum power at the boundary of the alternative parameter space.

3. Asymptotic Likelihood-Based Tests

Research on inequality restricted testing started in the early 1950's and was mainly in the context of multivariate analysis. Much of the early work was concerned with the testing problem in which the null hypothesis specifies that the mean of a multivariate normal distribution lies on the boundary of a non-negative part of a parameter space, and LR procedures were used. The main results before the 1970's are contained in the book by Barlow, Bartholomew, Bremner and Brunk (1972). A more recent account of the likelihood-based approach is found in Robertson, Wright and Dykstra's (1988) book.

This section is planned as follows. The general testing problem is outlined in subsection 3.1. This is followed by a survey of asymptotic likelihood-based procedures for single parameter testing. Subsection 3.3 outlines some standard results for general multiparameter one-sided LR, Wald and LM tests. The one-sided LM test is also called the Kuhn-Tucker (KT) test. Analogous tests of partially one-sided multivariate hypotheses are reviewed in subsection 3.4. Subsection 3.5 considers the calculation of the probability weights which characterize the asymptotic null distributions of these one-sided and partially one-sided test statistics. Related tests that were specifically designed to avoid these weight calculations are discussed in subsection 3.6 while subsection 3.7 is concerned with applications in the context of the linear regression model. The final subsection outlines asymptotic LMMP tests and reviews the literature on LMMP tests.

3.1 The General Testing Problem

The key issue in testing against inequality restrictions is how to take account of the one-sided nature of the problem. The conventional two-sided Wald and LM test statistics are typically quadratic forms of maximum likelihood (ML) estimators and scores, respectively. Where the number of parameters under test is one, appropriately signed square roots of these statistics can be applied as one-tailed tests.

Consider a general density function $f(x \mid \theta, \gamma)$ of an $n \times 1$ random vector x, where $\gamma \in \Gamma$ is a $q \times 1$ nuisance parameter vector, Γ is a subset of $R^q, \theta \in \Theta$ is a $p \times 1$ parameter vector under test and Θ is a subset of R^p. The problem of interest is to test

$$H_0 : \theta = \theta_0$$

against

$$H_a : \theta > \theta_0. \tag{3.1}$$

It is important to note that θ_0 is assumed to be an interior point of Θ. This assumption is needed because the derivations of the asymptotic distributions of many likelihood-based test statistics rely on the Taylor's expansion of the score function around θ_0, and more fundamentally, on the probability of $\hat{\theta}_i < \theta_{0i}$, where $\hat{\theta}$ is the unrestricted ML estimator of θ. It is then easy to see why testing a null hypothesis that is on the boundary of the parameter space is difficult. For example in the case of testing $\sigma^2 = 0$ against $\sigma^2 > 0$ where σ^2 is a variance, the probability of $\hat{\sigma}^2 < 0$ is meaningless because the density function extended to $\sigma^2 < 0$ is no longer a valid density function. Subsequently, $\hat{\sigma}^2 < 0$ is not a ML estimate. For a fuller discussion of these difficulties see Self and Liang (1987).

At this point it is convenient to introduce the following notation. Define the asymptotic information matrix as

$$I = \lim_{n \to \infty} -\frac{1}{n} E(\frac{\partial^2 L}{\partial \delta \partial \delta'}), \tag{3.2}$$

where $\delta = (\theta', \gamma')'$. Let I be partitioned as

$$I = \begin{bmatrix} I_{\theta\theta} & I_{\theta\gamma} \\ I_{\gamma\theta} & I_{\gamma\gamma} \end{bmatrix},$$

and let $I^{\theta\theta}$ be the block in I^{-1} corresponding to $I_{\theta\theta}$, i.e.

$$I^{\theta\theta} = (I_{\theta\theta} - I_{\theta\gamma} I_{\gamma\gamma}^{-1} I_{\gamma\theta})^{-1}.$$

3.2 Tests of One Parameter

For the case of $p = 1$, the conventional two-sided Wald test is given by

$$s_w = n(\hat{\theta} - \theta_0)^2 / \hat{I}^{\theta\theta}$$

which under H_0 is asymptotically distributed χ_1^2 (chi-squared with one degree of freedom) where ^ indicates unrestricted ML estimation. To take account of the one-sided nature of (3.1), we can apply

$$W = n^{1/2}(\hat{\theta} - \theta_0)/(\hat{I}^{\theta\theta})^{1/2} \tag{3.3}$$

which under H_0 is asymptotically distributed $N(0,1)$ as a test statistic and use the upper tail of the $N(0,1)$ distribution as the critical region. This test, applied to testing the regression coefficients β in (2.5) and (2.6) with Ω known, and with $\hat{I}^{\theta\theta}$ being properly adjusted, becomes the familiar one-sided t-test and therefore is UMPI.

The two-sided LM test when $p = 1$ is based on

$$s_s = \left(\left. \frac{\partial \log f(x \mid \theta, \hat{\gamma}_0)}{\partial \theta} \right|_{\theta=\theta_0} \right)^2 \hat{I}_0^{\theta\theta} / n$$

which under H_0 is asymptotically distributed χ_1^2 where $\hat{}_0$ indicates ML estimation under H_0. It is obvious then that the test based on

$$S = n^{-1/2} \left. \frac{\partial \log f(x \mid \theta, \hat{\gamma}_0)}{\partial \theta} \right|_{\theta=\theta_0} (\hat{I}_0^{\theta\theta})^{1/2} \tag{3.4}$$

and the upper tail of the $N(0,1)$ distribution is a proper asymptotic test for (3.1). Note that this test is a variant of the LB test discussed in subsection 2.1.

For (3.1) with $p = 1$, one can also use a signed square root (denoted by R) of the two-sided LR statistic. Let $L(\theta, \gamma)$ be the log likelihood function induced by $f(x \mid \theta, \gamma)$. Then R is defined by

$$R = \operatorname{sgn}(\hat{\theta} - \theta_0)\left[2(L(\hat{\theta}, \hat{\gamma}) - L(\theta_0, \hat{\gamma}_0))\right]^{1/2} \tag{3.5}$$

which under H_0 is asymptotically distributed $N(0,1)$, where $\operatorname{sgn}(x) = 1$, -1, 0 if $x > 0$, $x < 0$, $x = 0$, respectively. The upper tail of the $N(0,1)$ distribution is used as the critical region.

It is well known that the error of the normal approximation to R (and W, S as well) is of order $O_p(n^{-1/2})$, (see for example Barndorff-Nielsen (1986)) that is,

$$R - Z \sim O_p(n^{-1/2}),$$

where Z denotes a standard normal random variable. Recently, Barndorff-Nielsen (1986, 1991) proposed the modification to R by

$$R^* = R + R^{-1}\log(U/R)$$

which under H_0 is asymptotically distributed $N(0,1)$ to a high degree of approximation where U is a statistic depending on a ancillary statistic, say, A and has the same sign as R. He showed that

$$R^* - Z \sim O_p(n^{-3/2})$$

conditioning on A. Thus, a one-sided test for (3.1) based on R^* and the $N(0,1)$ null distribution is expected to perform better in terms of size in small samples than does the $N(0,1)$ approximation for R, W or S. In practice, finding the ancillary statistic A can be difficult. To handle this problem, DiCiccio and Martin (1993) proposed to replace U by a suitably chosen T to obtain

$$R^{**} = R + R^{-1}\log(T/R)$$

with

$$R^{**} - Z \sim O_p(n^{-1}).$$

The derivation of T is based on Bayesian methods. Typically, T is a known function of the derivatives of $L(\theta, \gamma)$ and a prior density $\pi(\theta, \gamma)$, where θ and γ are replaced either with unrestricted ML estimates or H_0 restricted ML estimates, and $\pi(\theta, \gamma)$ is chosen to satisfy certain conditions. While the normal approximation to R^{**} in small samples is not as good as to R^*, it is better than that to R. For further discussion on the use of higher-order asymptotics to improve accuracy of this form of one-sided LR test, see Pierce and Peters (1992).

3.3 General Tests

Instead of taking square roots of the two-sided LR, Wald and LM statistics for the case $p = 1$, another approach to testing against an inequality restricted alternative (3.1) is to follow the traditional ways of constructing one-sided LR, Wald and LM tests. This involves comparing the log-likelihood function under H_a to that under H_0 for the LR test, comparing ML estimates of θ under H_a to those under H_0 for the Wald test, and comparing scores under H_a to those under H_0 for the LM (KT) test. This in turn involves two issues: inequality (H_a) restricted ML estimation, and more importantly finding the asymptotic distributions of the test statistics under H_0. These one-sided tests can be used for $p \geq 1$. In the rest of this subsection, we review some general results.

Gouriéroux, Holly and Monfort (1980) (also see Farebrother (1988) for a summary) considered testing against (3.1). Let $\delta = (\theta', \gamma')'$ and $\delta_0 = (\theta_0', \gamma_0')'$ where γ_0 is the true unknown value, of the parameter γ. It is assumed that γ_0 is an interior point of Γ. Let λ be the score vector that corresponds to θ, scaled by dividing by sample size, that is, $\lambda = n^{-1}\partial L/\partial\theta$. Let $\tilde{\delta} = (\tilde{\theta}', \tilde{\gamma}')'$ be the solution (inequality restricted ML estimate) of

$$\begin{cases} \max n^{-1} L(\theta, \gamma) \\ \text{subject to } \theta \geq \theta_0, \ \gamma \in \Gamma, \end{cases}$$

so

$$\tilde{d} = n^{-1} \frac{\partial L}{\partial \delta}\bigg|_{\delta = \tilde{\delta}} = \begin{bmatrix} \tilde{\lambda} \\ 0 \end{bmatrix} \tag{3.6}$$

and

$$\tilde{\lambda}'(\tilde{\theta} - \theta_0) = 0, \tag{3.7}$$

where the last equality is the Kuhn-Tucker condition (see Kuhn and Tucker (1951)) which means that the components of $\tilde{\lambda}$ in (3.6) must be either zero or negative. This is because the presence of a positive component in $\tilde{\lambda}$ implies that $L(\theta, \gamma)$ has not yet been maximized. Equation (3.6) can be viewed as the definition of $\tilde{\lambda}$. Let $\hat{\delta}_0 = (\theta_0', \hat{\gamma}_0')'$ be the solution of

$$\begin{cases} \max n^{-1} L(\theta, \gamma) \\ \text{subject to } \theta = \theta_0, \gamma \in \Gamma, \end{cases}$$

so

$$\hat{d}_0 = n^{-1} \frac{\partial L}{\partial \delta}\bigg|_{\delta=\hat{\delta}_0} = \begin{bmatrix} \hat{\lambda}_0 \\ 0 \end{bmatrix}.$$

Note that $\hat{\delta}_0$ is the ML estimate of δ under the null hypothesis. Gouriéroux *et al.* noted that for (3.1), the one-sided LR, Wald and LM test statistics are, respectively,

$$s_{LR} = 2\big(L(\tilde{\delta}) - L(\hat{\delta}_0)\big), \tag{3.8}$$

$$s_W = n(\tilde{\theta} - \theta_0)'(\tilde{I}^{\theta\theta})^{-1}(\tilde{\theta} - \theta_0), \tag{3.9}$$

$$s_{KT} = n(\hat{\lambda}_0 - \tilde{\lambda})'\hat{I}_0^{\theta\theta}(\hat{\lambda}_0 - \tilde{\lambda}), \tag{3.10}$$

where $\tilde{I}^{\theta\theta}$ denotes $I^{\theta\theta}$ evaluated under the alternative hypothesis with $\tilde{\delta} = (\tilde{\theta}', \tilde{\gamma}')'$ and $\hat{I}_0^{\theta\theta}$ denotes $I^{\theta\theta}$ evaluated at the null hypothesis with $\hat{\delta}_0 = (\theta_0', \hat{\gamma}_0')'$. The critical regions are the upper tails of the null distributions of these tests. Obviously, (3.8) and (3.9) are exactly the general definitions of the LR and Wald tests, respectively, as discussed before. Note the presence of the inequality restricted score $\tilde{\lambda}$ in (3.10). This implies the computational advantage of the two-sided LM test only requiring estimates under H_0 disappears largely because $\tilde{\lambda}$ is obtained under the alternative hypothesis.

Gouriéroux *et al.* show that asymptotically $s_{LR} = s_W = s_{KT}$ under H_0 so the three tests are asymptotically equivalent under the null hypothesis. (A similar result for generalized linear models is proved in Silvapulle (1994)). This also implies the three tests are asymptotically equivalent under local alternatives. The asymptotic distribution of s_{KT} under H_0 (s_{LR} and s_W as well) is given by

$$\Pr(s_{KT} < c) = \sum_{i=0}^{p} w(p, i) \Pr(\chi_i^2 < c) \tag{3.11}$$

for $c \in \mathrm{R}$. This is a probability mixture of independent chi-squared distributions, χ_i^2, with different degrees of freedom. The weights $w(p,i), i = 0, \ldots, p$, represent the asymptotic probability of the event that under H_0 any i elements of $\tilde{\theta} - \theta_0$ are strictly positive, and the remaining $p - i$ elements are zeros (thus $\sum_{i=0}^{p} w(p,i) = 1$). By (3.6) and (3.7), this is also the event that i corresponding elements in $\tilde{\lambda}$ are zero and the remaining $p-i$ elements are strictly negative. Thus, the weights are dependent on the distribution of $\tilde{\lambda}$ which is information matrix, and therefore density function, $f(x \mid \theta, \gamma)$, specific. We note that $w(p,0)$ is the asymptotic probability that $\tilde{\theta} - \theta_0 < 0$ strictly, and χ_0^2 is the degenerate distribution with unit mass at zero. Let α be the significance level of the test. If the $w(p,i)$'s are known, the positive value c satisfying $\Pr(s_{KT} > c) = \alpha$ asymptotically under H_0 can be found by solving

$$\alpha = \sum_{i=1}^{p} w(p,i) \Pr(\chi_i^2 > c). \tag{3.12}$$

Note that $w(p,0)$ is not involved. Also note that α cannot be larger than $1 - w(p,0)$ as asymptotically $\Pr(s_{KT} > 0) = 1 - w(p,0)$ and $\Pr(s_{KT} = 0) = w(p,0)$. Because $w(p,0) \leq \frac{1}{2}$ in all cases, this causes no difficulty for the conventional choice of α. The probability (3.12) can be calculated if a subroutine to compute probabilities from chi-squared distributions is available.

3.4 Partially One-Sided Testing

A generalization of the above problem involves testing

$$H_0 : \theta_1 = \theta_{10}, \theta_2 = \theta_{20}$$

against

$$H_a : \theta_1 \geq \theta_{10}, \theta \neq (\theta'_{10}, \theta'_{20})', \tag{3.13}$$

where for $\theta = (\theta'_1, \theta'_2)', \theta_1$ is $p_1 \times 1, \theta_2$ is $p_2 \times 1, p_1 + p_2 = p$, and θ_2 is unrestricted. This is a partially one-sided testing problem. By defining $\tilde{\delta} = (\tilde{\theta}'_1, \tilde{\theta}'_2, \tilde{\gamma}')'$ as the ML estimator under the alternative hypothesis, $\hat{\delta}_0 = (\theta'_{10}, \theta'_{20}, \hat{\delta}'_0)'$ as the ML estimator under H_0, and the corresponding scaled scores as, respectively,

$$n^{-1} \frac{\partial L}{\partial \delta}\bigg|_{\delta = \tilde{\delta}} = \begin{bmatrix} \tilde{\lambda}_1 \\ 0 \\ 0 \end{bmatrix},$$

and

$$n^{-1}\frac{\partial L}{\partial \delta}\bigg|_{\delta=\hat{\delta}_0} = \begin{bmatrix} \hat{\lambda}_{10} \\ \hat{\lambda}_{20} \\ 0 \end{bmatrix},$$

then the LR, Wald and KT tests are given by rejecting H_0 for large values of

$$s_{LR} = 2(L(\tilde{\delta}) - L(\hat{\delta}_0)), \tag{3.14}$$

$$s_W = n\begin{bmatrix} \tilde{\theta}_1 - \theta_{10} \\ \tilde{\theta}_2 - \theta_{20} \end{bmatrix}' (\tilde{I}^{\theta\theta})^{-1} \begin{bmatrix} \tilde{\theta}_1 - \theta_{10} \\ \tilde{\theta}_2 - \theta_{20} \end{bmatrix}, \tag{3.15}$$

$$s_{KT} = n\begin{bmatrix} \tilde{\lambda}_1 - \lambda_{10} \\ -\tilde{\lambda}_{20} \end{bmatrix}' \hat{I}_0^{\theta\theta} \begin{bmatrix} \tilde{\lambda}_1 - \lambda_{10} \\ -\tilde{\lambda}_{20} \end{bmatrix}, \tag{3.16}$$

respectively. As in the purely one-sided case, these tests are asymptotically identical under H_0. The asymptotic distribution under H_0 is given by

$$\Pr(s_{KT} < c) = \sum_{i=0}^{p_1} w(p_1, i)\Pr(\chi^2_{p_2+i} < c). \tag{3.17}$$

The probability weights have a similar interpretation to those in (3.11).

So far we have presented the main features of the one-sided or partially one-sided LR, Wald and KT tests. The results are more general than they appear. Instead of (3.13), one may be interested in testing functions of the parameters. That is, testing

$$H_0 : h_1(\theta) = 0, h_2(\theta) = 0$$

against

$$H_a : h_1(\theta) \geq 0, \quad \text{excluding } H_0, \tag{3.18}$$

for some functions $h_1(\cdot)$ and $h_2(\cdot)$ where $h_1(\cdot)$ is $m_1 \times 1$ and $h_2(\cdot)$ is $m_2 \times 1$. By regarding $h_i(\theta)$ as a parameter transformation ($\phi_i = h_i(\theta)$) and calculating the information matrix and scores accordingly, the results (3.14) to (3.17) remain the same with respect to $\phi = (\phi_1', \phi_2')'$ with p_i in (3.17) replaced by m_i, $i = 1, 2$. Of course, additional regularity conditions on the differentiability of $h_i(\cdot)$ are assumed.

Kodde and Palm (1986) constructed a test for (3.18) from another angle. They considered comparing the minimum distance, D_0, from $n^{1/2}\hat{\phi}$ to the space of H_0 to the distance,

D_1, from $n^{1/2}\hat{\phi}$ to the space of $H_0 \cup H_a$, where $\hat{\phi}$ is any unrestricted consistent estimator of ϕ. For example, $\hat{\phi}$ can be the ML estimator $\hat{\phi} = h(\hat{\theta})$. Suppose Σ is the asymptotic covariance matrix of $n^{1/2}\hat{\phi}$ and the distances D_0 and D_1 are in the metric of Σ^{-1}, see below. Then the test statistic is $D = D_0 - D_1$. The test has the asymptotic distribution implied by (3.17) (with $p_1 = m_1$, $p_2 = m_2$).

3.5 Probability Weights

The essential step in the construction of one-sided LR and Wald tests, and of the KT test is to find the probability weights $w(p, i)$ which determine the mixture of chi-squared distributions for the asymptotic null distribution. Let $I_j^i, j = 1, 2, \ldots, 2^p$ denote 2^p subsets of $\{1, 2, \ldots, p\}$ where i indicates the number of elements in I_j^i, let M_i be the collection of $\binom{p}{i}$ I_j^i's with a particular i value. Associated with M_i, let A_i be the event that any i elements of $\tilde{\theta} - \theta_0$ are strictly larger than zero, and the remaining $p - i$ elements are zero. Thus, $I_j^i \in M_i$ determines which i elements in $\tilde{\theta} - \theta_0$ are strictly positive. Note A_i can happen in $\binom{p}{i}$ mutually exclusive ways according to $I_j^i \in M_i$. We use $\sum_{M_i}$ to denote summation over these $\binom{p}{i}$ exclusive subevents. Recall that $w(p, i) = P(A_i)$. We first discuss how to determine the weights by considering the limiting case as $n \to \infty$.

The weights depend on the asymptotics that lead to the null distribution (3.11). An essential assumption is that the unrestricted ML estimator $n^{1/2}(\hat{\theta} - \theta_0)$ is asymptotically normal with $I^{\theta\theta}$ as its covariance matrix. Since $I^{\theta\theta}$ can be consistently estimated, it can be treated as if it is known. Heuristically, by regarding x as $n^{1/2}(\hat{\theta} - \theta_0)$ and Σ as $I^{\theta\theta}$, the weights for the problem (3.1) are equal to the weights in the specific problem of testing

$$H_0 : \theta = 0 \text{ against } H_a : \theta > 0 \text{ when } x \sim N(\theta, \Sigma) \tag{3.19}$$

where Σ is $p \times p$ and known. Kudô (1963) proposed the one-sided LR test for this case. In particular, because x is normal and Σ is known, the null distribution (3.11) was proved to hold exactly in finite samples for the LR test. The same exact null distribution also holds for testing linear regression coefficients when the error vector is normal with known covariance matrix; see subsection 3.7.

Because a multi-variate normal distribution is characterized by its covariance matrix Σ, the weights are functions of Σ and that part of the parameter space under test, denoted as C. Therefore we may denote the weights as $w(p,i) = w(p,i \mid \Sigma, C)$, where i is the degrees of freedom of the chi-squared distribution to which the weight $w(p,i \mid \Sigma, C)$ is attached. Typically when C is implied by $H_0 \cup H_a$ from (3.19); i.e., $C = \{\theta; \theta \geq 0\}, w(p,i \mid \Sigma, C)$ is simply written as $w(p,i \mid \Sigma)$.

Within the framework of (3.19), the determination of $P(A_i) = w(p,i \mid \Sigma)$ becomes relatively straightforward. Given the observation $x, \tilde{\theta}$ is the solution of

$$D_1 = \min_{\theta \geq 0} (x-\theta)' \Sigma^{-1} (x - \theta);$$

i.e., $\tilde{\theta}$ is the projection of x onto the cone $C = \{\theta; \theta \geq 0\}$ in the metric of Σ^{-1}. (D_1 is called the distance from x to C in the metric of Σ^{-1}.) $P(A_i)$ is the probability of the set of those x in the sample space whose projections on $C = \{\theta; \theta \geq 0\}$ contain exactly i elements that are larger than zero. From this, it is obvious that if $\Sigma = I$, due to symmetry,

$$w(p,i \mid I) = \binom{p}{i} 2^{-p}, \qquad i = 0, 1, \ldots, p.$$

Kudô (1963) demonstrates that for a general Σ,

$$w(p,i \mid \Sigma) = \sum_{M_i} P((\Sigma_{M_i'})^{-1}) P(\Sigma_{M_i;M_i'}) \tag{3.20}$$

where M_i' is the complement of $M_i, \Sigma_{M_i'}$ is the covariance matrix of the marginal distribution of those elements of x whose coordinates belong to M_i', $\Sigma_{M_i;M_i'}$ is the covariance matrix of the conditional distribution of elements of x whose coordinates belong to M_i given the remaining elements are zero, $P(A)$ is the probability of $y \geq 0$ for $y \sim N(0, A), P(\Sigma_{M_0;M_0'}) = 1$ and $P((\Sigma_{M_p'})^{-1}) = 1$. As an example, for $p = 2$, it is easily seen that $w(2,1 \mid \Sigma) = 0.5$, $w(2,0 \mid \Sigma) = \cos^{-1} \rho / 2\pi$ and $w(2,2 \mid \Sigma) = 0.5 - w(2,0 \mid \Sigma)$ where ρ is the correlation coefficient between x_1 and x_2 for $x = (x_1, x_2)' \sim N(0, \Sigma)$.

Now consider testing $H_0 : R\theta = 0$ against $H_a : R\theta > 0$ given $x \sim N(\theta, \Sigma)$ where R is a known $k \times p$ matrix with $k \leq p$ and of rank k. This is the case where $C = \{\theta; R\theta \geq 0\}$. When $k = p$, by regarding $R\theta$ as a new parameter vector, it is easy to see that $w(p, i \mid$

$\Sigma, C) = w(p, i \mid R\Sigma R')$, $i = 0, 1, \ldots, p$, and therefore Kudô's formulae can be used to find $w(p, i \mid R\Sigma R')$. When $k < p$, $w(p, p-k+i \mid \Sigma, C) = w(k, i \mid R\Sigma R')$, $i = 0, 1, \ldots, k$, and $w(p, j \mid \Sigma, C) = 0$, $j = 0, 1, \ldots, p-k-1$. As we see in this example, when the parameter space under the alternative hypothesis is specified as a cone such as C, the basic idea is to transform the parameters to get a testing problem of the form of (3.19), so the weight formulae can be applied (see Shapiro (1988), pp.54-55, for more details). This idea can also be found in Wolak (1988). Note however that in practice, when p (the number of restrictions) is greater than 4, calculation of the weights becomes very complicated. Computer programs for calculating the weights based on (3.20) have been written by Bohrer and Chow (1978).

3.6 Tests That Avoid Weight Calculations

To avoid the difficulty of calculating the weights, Kodde and Palm (1986) provided lower and upper bounds of the critical values for the general case of (3.18). These can be used for cases such as (3.1) and (3.13) and for special cases such as (3.22) in the next subsection. This method is quite easy to apply in practice. However the inconclusive region becomes wider as p_2 increases.

Rogers (1986) proposed a modified KT test for (3.1) with $\theta_0 = 0$. This test avoids the calculation of ML estimates under the alternative hypothesis. Specifically, this involves constructing 2^p cones C_{0j}^i that partition R^p according to the 2^p sets $I_j^i, j = 1, 2, \ldots, 2^p$. Let W_0 be I_0^{-1} where I_0 is the asymptotic information matrix I evaluated at δ_0. The structure of C_{0j}^i is determined by elements of W_0. Also define $d_0 = n^{-1}\partial L/\partial\delta \mid_{\delta=\delta_0}$. Let

$$V_0 = \left[[I_p \vdots 0]W_0[I_p \vdots 0]'\right]^{-1}.$$

Construct the $p \times (p+q)$ matrix

$$Z_0 = \left[I_p \vdots V_0[I_p \vdots 0]W_0[0 \vdots I_q]\right]$$

and let B_{0j} be the $(p+q) \times (p+q)$ symmetric matrix of zeros whose top $p \times p$ block is

$$V_0 - \Big(([I_p - \mathrm{diag}(e_j)] \vdots 0)V_0([I_p - \mathrm{diag}(e_j)] \vdots 0)'\Big)^+$$

in which $(\cdot)^+$ denotes the Moore-Penrose inverse and e_j is the $p \times 1$ vector whose k^{th} element is one if $k \in I_j^i$ and zero otherwise, where I_j^i is associated with the cone C_{0j}^i. The construction of these cones and matrices is motivated by the fact that asymptotically $n^{-1/2} Z_0 d_0 \in C_{0j}^i$ if and only if the i elements (whose coordinates are elements of I_j^i on which C_{0j}^i is based) of $n^{\frac{1}{2}}\tilde{\theta}$ are strictly positive and the remaining $n-i$ elements are zero. Rogers shows that under H_0,

$$s_{KT} - n d_0' W_0 B_{0j} W_0 d_0 \xrightarrow{P} 0 \qquad \text{if } n^{1/2} Z_0 d_0 \in C_{0j}^i, \tag{3.21}$$

and the unconditional limiting distribution of $n d_0' W_0 B_{0j} W_0 d_0$ is χ_i^2. In practice, $d_0' W_0 B_{0j} W_0 d_0$ is not observable, and it is estimated by $\hat{d}_0' W_0 \hat{B}_{0j} \hat{W}_0 \hat{d}_0$. Note that unlike the notation in Rogers' paper, we have used $\hat{}_0$ to denote estimation under H_0 and have omitted the subscript n. Define

$$Q_n = n d_0' W_0 B_{0j} W_0 d_0 \qquad \text{if } n^{1/2} Z_0 d_0 \in C_{0j}^i,$$

and

$$S_n = \hat{d}_0' \hat{W}_0 \hat{B}_{0j} \hat{W}_0 \hat{d}_0 \qquad \text{if } n^{1/2} \hat{Z}_0 \hat{d}_0 \in \hat{C}_{0j}^i,$$

where $\hat{C}_{0j}^i$ is C_{0j}^i when δ_0 is replaced by $\hat{\delta}_0$. Rogers shows that under H_0

$$S_n - Q_n \xrightarrow{P} 0.$$

This, combined with (3.21), implies $S_n - s_{KT} \xrightarrow{P} 0$. Thus, the test based on S_n and the KT test are asymptotically equivalent under the null hypothesis of $\theta = 0$.

The calculation of S_n involves calculating $n^{1/2} \hat{Z}_0 \hat{d}_0$ first as Z_0 is not I_{0j}^i specific, identifying the set $\hat{C}_{0j}^i$ (therefore the set I_{0j}^i) such that $n^{1/2} \hat{Z}_0 \hat{d}_0 \in \hat{C}_{0j}^i$, and consequently finding $\hat{B}_{0j}$ which is I_{0j}^i specific. The calculation does not require inequality restricted ML estimation. The asymptotic distribution of S_n under H_0 is still (3.11) as $S_n - s_{KT} \xrightarrow{P} 0$ implies S_n and s_{KT} have the same asymptotic distribution. The weights are

$$w(p,i) = P(Z_0 x \in C_{\cdot i}) = P(A_i)$$

where $Z_0 x \sim N(0, Z_0 I_0 Z_0')$, $C_{\cdot i}$ is the union of the $p!/(i!(p-i)!)$ cones C_{0j}^i for fixed i. The weights $w(p,i)$ can be found from (3.20). They may also be estimated by integrating the $N(0, \hat{Z}_0 \hat{I}_0 \hat{Z}_0')$ density over the sets $\hat{C}_{\cdot i}$.

To avoid calculating all the weights, Rogers suggested that the test be conducted based on critical values obtained separately from component individual χ^2 distributions. Specifically, let c_{i,α^*} denote the α^*-level critical values of the χ_i^2 distribution for $i = 1, \ldots, p$. H_0 is accepted if $n^{1/2}\hat{Z}_0\hat{d}_0 \in \hat{C}\cdot_i$ for one of $i = 1, \ldots, p$ and $S_n \leq c_{i,\alpha^*}$ for the corresponding i, is rejected if $S_n > c_{i,\alpha^*}$ for the same i, and is accepted whenever $n^{1/2}\hat{Z}_0\hat{d}_0 \in \hat{C}\cdot_0$. The overall asymptotic size of this strategy is

$$\sum_{i=1}^{p} \alpha^* w(p,i) = \alpha^*(1 - w(p,0)).$$

Thus if a test of asymptotic size α is desired, one should take $\alpha^* = \alpha/(1 - w(p,0))$. Using this testing strategy, the burden of calculating weights reduces to calculating $w(p,0)$ only, which is very simple from (3.20) once I is estimated. Rogers (p.350) pointed out possibilities of further simplification.

The modified KT test may be used when the null hypothesis is on the boundary of the admissible parameter space. The validity of this is guaranteed if, among other regularity conditions, the normalized score vector under the null is asymptotically normal with mean zero.

3.7 The Linear Regression Model

More specific results may be obtained for the linear regression model (2.5) and (2.6). To avoid possible confusion in notation, we replace $\Omega(\theta)$ in (2.6) by $\Omega(\eta)$.

The problem of interest is one of testing

$$H_0 : R\beta = r$$

against

$$H_a : R\beta > r, \tag{3.22}$$

where R is a known $p \times k$ matrix with rank $p < k$, and r is $p \times 1$ and known. This is in the form of (3.18) with θ replaced by β. Gouriéroux, Holly and Monfort (1982) show that, if $\sigma^2\Omega(\eta)$ is known,

$$\begin{aligned}
s_{LR} &= 2(L(\tilde{\beta},\sigma^2,\eta) - L(\hat{\beta}_0,\sigma^2,\eta)) \\
&= -(y - X\tilde{\beta})'(\sigma^2\Omega(\eta))^{-1}(y - X\tilde{\beta}) + (y - X\hat{\beta}_0)'(\sigma^2\Omega(\eta))^{-1}(y - X\hat{\beta}_0), \\
s_W &= (R\tilde{\beta} - r)'[R(X'(\sigma^2\Omega(\eta))^{-1}X)^{-1}R']^{-1}(R\tilde{\beta} - r), \\
s_{KT} &= (\hat{\lambda}_0 - \tilde{\lambda})'R(X'(\sigma^2\Omega(\eta))^{-1}X)R'(\hat{\lambda}_0 - \tilde{\lambda})/4,
\end{aligned}$$

and $s_{LR} = s_W = s_{KT}$, where $\hat{\beta}_0$ is the equality constrained (under H_0) estimator of β. These tests have exact null distributions as probability mixtures of chi-squared distributions and the degenerate distribution at zero, as in (3.11). If $\sigma^2\Omega(\eta)$ is unknown, σ^2 and η in s_{LR}, s_W and s_{KT} above are replaced in the conventional way by the ML estimators $\hat{\sigma}_0^2$ and $\hat{\eta}_0$ under H_0 and $\tilde{\sigma}^2$ and $\tilde{\eta}$ under H_a, respectively. The familiar finite sample inequality $s_W \geq s_{LR} \geq s_{KT}$ still holds, and the above null distribution holds asymptotically.

Farebrother (1986) considered the case where $\Omega(\eta) = I_n$ so the error covariance matrix is unknown up to the scalar parameter σ^2. He found

$$s_{LR} = n\log\left(1 + \frac{s_W}{n}\right),$$

$$s_{LR} = -n\log\left(1 - \frac{s_{KT}}{n}\right),$$

so $s_W \geq s_{LR} \geq s_{KT}$. In addition, the exact null distribution of s_{KT} is a probability mixture of (central) *Beta* distributions so that

$$\Pr(s_{KT}/n < c) = \sum_{i=0}^{p} w(p,i)\Pr(B(i/2,(n-k+p-i)/2) < c),$$

where $B(\cdot,\cdot)$ denotes the *Beta* random variable. Thus, all three tests have the same critical regions in the sample space. For the same problem, Hillier (1986) used similarity arguments to obtain the one-sided LR test, whose critical region is characterized by the tails of $F(i, n-k+p-i)$ distributions. The equivalence of this critical region to that by Farebrother may be checked by using the fact that $B(i/2, j/2) = iF(i,j)/(iF(i,j)+j)$. Note that for the special case of $p = 1$, s_W (say) does not reduce to the familiar one-sided t-test or its asymptotically identical variant W given by (3.3). For $\Omega(\eta) = I_n$, Silvapulle (1992a) developed tests of one-sided hypotheses which are robust against non-normal error distributions. These tests are based on M-estimators and bounded influence estimators.

For the linear regression model (2.5), Wolak (1987, 1989) considered the problem of testing $H_0 : R\beta > r$ against $H_a : R\beta \neq r$. This problem is of interest when one cannot be sure about ruling out the possibility of $R\beta \geq r$ not being true. For an interesting example, see Silvapulle (1992b). Wolak (1989) was able to find the null distributions of the KT (Wald and LR) statistics for respectively testing $R\beta = r$ against $R\beta > r$ and $R\beta > r$ against $R\beta \neq r$ as probability mixtures of products of two independent χ^2 distributions when $\sigma^2\Omega(\eta)$ is known. If $\sigma^2\Omega(\eta)$ is unknown up to finite parameters, these null distributions hold asymptotically. Also $s_{LR} = s_W = s_{KT}$ if $\sigma^2\Omega(\eta)$ is known. If $\sigma^2\Omega(\eta)$ is unknown, $s_{LR} \leq s_W \leq s_{KT}$ in finite samples and the three tests are equivalent asymptotically. By utilizing primal and dual relationships in quadratic programming, he also established the equivalence of the KT (Wald and LR) test for $R\beta > r$ against $R\beta \neq r$ to the KT (Wald and LR) test of $H_0 : \lambda = 0$ against $H_a : \lambda > 0$ where λ is the vector of Lagrange multipliers implied in the inequality restricted ML estimation of β. Thus power studies can be concentrated on cases which test a point null against one-sided alternatives.

In subsection 3.5, we discussed ways of calculating the probability weights in the asymptotic null distributions of the test statistics. Another way to deal with these weights is to use Monte Carlo methodology. Under this principle, in addition to simulating the weights, one can also consider certain variants of the one-sided LR, Wald or KT tests. For example, Farebrother (1990) considered using

$$\max_i b_i(R\hat{\beta} - r)/Var(b_i(R\hat{\beta} - r))$$

as a test statistic for testing (3.22), where b_i is the ith row of $(R(X'X)^{-1}R')^{-1}$. Though critical values of this test may also be obtained by numerical integration, Farebrother pointed out that simulation techniques can be employed for the same kind of statistics used when R has row rank deficiency.

3.8 Asymptotic LMMP Tests

In subsection 2.2, King and Wu's (1993) LMMP test was introduced as a multi-parameter generalization of the LB test of a single parameter. The test is based on the sum of the scores of the parameters under test evaluated at the null hypothesis with critical regions given by (2.10). If there are difficulties in finding exact critical values or if nuisance parameters are present, an asymptotic version of the test for the general problem outlined in subsection 3.1 can be constructed as follows. H_0 is rejected for large values of

$$\ell' \frac{\partial \log f(x \mid \theta, \hat{\gamma}_0)}{\partial \theta}\bigg|_{\theta=\hat{\theta}_0} (n\ell' \hat{I}_0^{\theta\theta} \ell)^{-1/2} \tag{3.23}$$

which under H_0 is asymptotically distributed $N(0,1)$, where ℓ is the $p \times 1$ vector of ones.

Since King and Wu first presented their test procedure (King and Wu (1989)), a range of applications of both the exact test (2.10) and the asymptotic test (3.23) have been suggested and/or investigated by various authors. Leybourne and McCabe (1992) and McCabe and Leybourne (1994) proposed their use in testing for multiple random-walk coefficients in nonlinear and linear regression models, respectively. King and Shively (1993) suggested the use of the exact LMMP invariant (LMMPI) test for two random coefficient testing problems which they had reparameterized to overcome difficulties caused by nuisance parameters that are present only under the alternative hypothesis. Their Monte Carlo results suggest that at least for testing for the presence of a single Rosenberg (1973) AR(1) regression coefficient, the test has good small-sample power properties. Shively (1993) also recommended the use of the exact LMMPI test in the case of testing for AR(p) disturbances with missing observations in the linear regression model. His empirical power investigation suggests the test has good power properties against AR(p) disturbances. Wu and Bhatti (1989) investigated the power of the exact LMMPI test for three-stage block effects in linear regression disturbances.

Lee and King (1993) proposed the use of the asymptotic LMMP test for the presence of autoregressive conditional heteroscedastic (ARCH) and generalized ARCH (GARCH) regression disturbances. Their Monte Carlo results show that the test has better power properties and possibly more accurate critical values than the standard two-sided LM test for second-order ARCH and GARCH disturbances. They also found the asymptotic LMMP test to be reasonably robust under nonnormality.

Wu (1991) and King and Wu (1989, 1993) give general formulae for exact LMMPI tests of the error covariance matrix in the linear regression model (2.5). They investigated the power of exact LMMPI tests for various testing problems and found that for certain problems and regression matrices, the test can perform relatively poorly especially on the boundary of the parameter space. Baltagi, Chang and Li (1992) report similar findings in the case of testing for two-way error components in (2.5) as do Ara and King (1993) and Rahman and King (1993) (also see Rahman and King (1994)) for the asymptotic LMMP test for multiple Hildreth-Houck (1968) random regression coefficients in (2.5). These problems with power are caused by situations in which there is a degree of cancellation occurring in the sum of scores. The test appears to work well when scores are positively correlated and break down, particularly near boundaries, when scores are negatively correlated.

Among other things, Ara and King (1993) and Rahman and King (1993) investigate the application of asymptotic LMMP tests to marginal likelihoods or equivalently the likelihood of a maximal invariant in the case of testing linear regression disturbances. Ara and King (1993) assume all nuisance parameters can be eliminated by invariance while Rahman and King (1993) allow for nuisance parameters which cannot be eliminated. In the latter case in particular, this approach appears to result in a slightly more accurate test than (3.23) applied to the full likelihood function.

4. Concluding Remarks

After surveying the literature on one-sided testing, it is not difficult to conclude that testing one parameter is much easier than testing multiple parameters. Consequently the literature on one-sided testing of one parameter is much better developed than that on one-sided multi-parameter testing. Locally best tests and point optimal tests typically work well in the one parameter case while if there are nuisance parameters that cannot be eliminated by invariance, similarity or marginal likelihood arguments, true one-sided LR, Wald or LM tests can be applied. When it comes to one-sided multi-parameter testing, it is not clear we have all the answers. A ULB test is unlikely to exist and there is evidence that LMMP tests do not necessarily work well in all situations. We suspect the same is true of point optimal tests. There are the general asymptotic one-sided tests based on inequality-restricted ML estimates that were discussed in subsection 3.3. A question mark hangs over these because in the one-sided case there are examples where these tests differ from known uniformly most powerful tests for large significance levels. In addition, when the number of parameters under test is large, the calculation of the weights for the asymptotic distribution under the null hypothesis is not easy. Because they are problem specific, they typically need to be estimated using simulation methods.

One option in the multi-parameter case that has not been discussed above is multiple single-parameter testing. This involves a sequence of separate one-sided tests of single parameters, a good example being separate t-tests of different coefficients in the linear regression model. The main difficulty with these procedures is in controlling the overall probability of a Type I error. They are known as induced tests and almost always involve the use of probability inequalities to approximately control overall size. An excellent survey of this literature is given by Savin (1984). With the exception of Savin's (1980, 1984) work, there are very few comparisons of the small-sample power properties of induced tests and multi-parameter tests. In general, induced tests often have a relative power advantage along or near axes when the null hypothesis involves zero parameter values.

In conclusion, the incorporation of inequality information into statistical procedures can help improve the quality or accuracy of statistical inference. How this is best done in the case of testing multiple parameters still appears to be an open question which requires further research.

Acknowledgement

This research was supported in part by an ARC grant. We are grateful to Alan Morgan for research assistance and to Ishaq Bhatti, Brendan McCabe and Mervyn Silvapulle for constructive comments on earlier drafts.

References

1. Andrews, D.W.K., Lee, I. and Ploberger, W. (1994). Optimal changepoint tests for linear regression. *J. Econometrics*, forthcoming.

2. Andrews, D.W.K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, forthcoming.

3. Ansley, C.F., Kohn, R. and Shively, T.S. (1992). Computing p-values for the generalized Durbin-Watson and other invariant test statistics. *J. Econometrics* 54, 277-300.

4. Ara, I. and King, M.L. (1993). Marginal likelihoood based tests of regression disturbances. Paper presented at the 1993 Australasian Meeting of the Econometric Society, Sydney.

5. Baltagi, B.H., Chang, Y.J., and Li, Q. (1992). Monte Carlo results on several new and existing tests for the error component model. *J. Econometrics* 54, 95-120.

6. Barlow, R.E., Bartholomew, D.J., Bremner, J.N., and Brunk, H.D. (1972). *Statistical Inference under Order Restrictions.* Wiley, New York.

7. Barndorff-Nielsen, O.E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika* 73, 307-322.

8. Barndorff-Nielsen, O.E. (1991). Modified signed log likelihood ratio. *Biometrika* 78, 557-563.

9. Berenblut, I.I. and Webb, G.I. (1973). A new test for autocorrelated errors in the linear regression model. *J. Roy. Statist. Soc.* B 35, 33-50.

10. Bhatti, M.I. (1992). Testing regression models based on sample survey data. Unpublished Ph.D. thesis, Monash University.

11. Bhatti, M.I. and King, M.L. (1990). A beta-optimal test of the equicorrelation coefficient. *Austral. J. Statist.* 32, 87-97.

12. Bhatti, M.I. and King, M.L. (1993). Testing for subblock effects in multi-stage linear regression models. Mimeo, Monash University.

13. Bohrer, R. and Chow, W. (1978). Weights for one-sided multivariate inference. *Appl. Statist.* 27, 100-104.

14. Brooks, R.D. and King, M.L. (1994). Testing Hildreth-Houck against return to normalcy random regression coefficients. *J. Quant. Econ.*, forthcoming.

15. Cox, D.R. and Hinkley, D.V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

16. Davies, R.B. (1969). Beta-optimal tests and an application to the summary evaluation of experiments, *J. Roy. Statist. Soc.* B 31, 524-538.

17. Davies, R.B. (1980). Algorithm AS155: The distribution of a linear combination of χ^2 random variables. *Appl. Statist.* 29, 323-333.

18. DiCiccio, T.J. and Martin, M.A. (1993). Simple modifications for signed roots of likelihood ratio statistics. *J. Roy. Statist. Soc.* B 55, 305-316.

19. Dufour, J-M. and King, M.L. (1991). Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors. *J. Econometrics* 47, 115-143.

20. Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression I. *Biometrika* 37, 409-428.

21. Durbin, J. and Watson, G.S. (1951). Testing for serial correlation in least squares regression II. *Biometrika* 38, 159-178.

22. Durbin, J. and Watson, G.S. (1971). Testing for serial correlation in least squares regression III. *Biometrika* 58, 1-19.

23. Evans, M.A. and King, M.L. (1985a). Critical value approximations for tests of linear regression disturbances. *Austral. J. Statist.* 27, 68-83.

24. Evans, M.A. and King, M.L. (1985b). A point optimal test for heteroscedastic disturbances. *J. Econometrics* 27, 163-178.

25. Evans, M.A. and King, M.L. (1988). A further class of tests for heteroscedasticity. *J. Econometrics* 37, 265-276.

26. Farebrother, R.W. (1986). Testing linear inequality constraints in the standard linear model. *Comm. Statist. A* 15, 7-31.

27. Farebrother, R.W. (1988). Some recent econometric work on tests for inequality constraints. *J. Quant. Econ.* 4, 227-238.

28. Farebrother, R.W. (1990). Simpler tests of linear inequality constraints in the standard linear model. Paper presented at the Sixth World Congress of the Econometric Society, Barcelona.

29. Franzini, L. and Harvey, A.C. (1983). Testing for deterministic trend and seasonal components in time series models. *Biometrika* 70, 673-682.

30. Fraser, D.A.S., Guttman, I. and Styan, G.P.H. (1976). Serial correlation and distributions on the sphere. *Comm. Statist.* A5, 97-118.

31. Gourieroux, C., Holly, A. and Monfort, A. (1980). Kuhn-Tucker, likelihood ratio and Wald tests for nonlinear models with inequality constraints on the parameters. Discussion Paper No.770, Harvard Institute of Economic Research, Harvard University.

32. Gourieroux, C., Holly, A. and Monfort, A. (1982). Likelihood ratio test, Wald test and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. *Econometrica* 50, 63-80.

33. Hildreth, C. and Houck, J.P. (1968). Some estimators for a linear model with random coefficients. *J. Amer. Statist. Assoc.* 63, 584-595.

34. Hillier, G.H. (1986). Joint tests for zero restrictions on non-negative regression coefficients. *Biometrika* 73, 657-669.

35. Hillier, G.H. (1987). Classes of similar regions and their power properties for some econometric testing problems. *Econometric Theory* 3, 1-44.

36. Imhof, P.J. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* 48, 419-426.

37. Isaacson, S.L. (1951). On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters, *Ann. Math. Statist.* 22, 217-234.

38. Judge, G.G. and Yancey, T.A. (1986). *Improved Methods of Inference in Econometrics.* North-Holland, Amsterdam.

39. King, M.L. (1981a). The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic. *J. Econometrics* 17, 51-66.

40. King, M.L. (1981b). A small sample property of the Cliff-Ord test for spatial correlation. *J. Roy. Statist. Soc.* B 43, 263-264.

41. King, M.L. (1982). Testing for a serially correlated component in regression disturbances. *Inter. Econ. Rev.* 23, 577-582.

42. King, M.L. (1983a). Testing for autoregressive against moving average errors in the linear regression model. *J. Econometrics* 21, 35-51.

43. King, M.L. (1983b). Testing for moving average regression disturbances. *Austral. J. Statist.* 25, 23-34.

44. King, M.L. (1984). A new test for fourth-order autoregressive disturbances. *J. Econometrics* 24, 269-277.

45. King, M.L. (1985a). A point optimal test for autoregressive disturbances. *J. Econometrics* 27, 21-37.

46. King, M.L. (1985b). A point optimal test for moving average regression disturbances. *Econometric Theory* 1, 211-222.

47. King, M.L. (1987a). Towards a theory of point optimal testing. *Econometric Rev.* 6, 169-218.

48. King, M.L. (1987b). Testing for autocorrelation in linear regression models: A survey. In M.L. King and D.E.A. Giles (eds.), *Specification Analysis in the Linear Model.* Routledge and Kegan Paul, London. 19-73.

49. King, M.L. (1987c). An alternative test for regression coefficient stability. *Rev. Econ. Statist.* 69, 379-381.

50. King, M.L. (1989). Testing for fourth-order autocorrelation in regression disturbances when first-order autocorrelation is present. *J. Econometrics* 41, 285-301.

51. King, M.L. and Evans, M.A. (1984). A joint test for serial correlation and heteroscedasticity. *Econ. Letters* 16, 297-302.

52. King, M.L. and Evans, M.A. (1985). The Durbin-Watson test and cross-sectional data. *Econ. Letters* 18, 31-34.

53. King, M.L. and Evans, M.A. (1988). Locally optimal properties of the Durbin-Watson test. *Econometric Theory* 4, 509-516.

54. King, M.L. and Hillier, G.H. (1985). Locally best invariant tests of the error covariance matrix of the linear regression model. *J. Roy. Statist. Soc.* B 47, 98-102.

55. King, M.L. and Shively, T.S. (1993). Locally optimal testing when a nuisance parameter is present only under the alternative. *Rev. Econ. Statist.* 75, 1-7.

56. King, M.L. and Smith, M.D. (1986). Joint one-sided tests of linear regression coefficients. *J. Econometrics* 32, 367-383.

57. King, M.L. and Wu, P.X. (1989). Locally optimal one-sided tests for multiparameter hypotheses. Paper presented at the Australasian Meeting of the Econometric Society, Armidale.

58. King, M.L. and Wu, P.X. (1993). Locally optimal one-sided tests for multiparameter hypotheses. Mimeo, Monash University.

59. Kodde, D.A. and Palm, F.C. (1986). Wald criteria for jointly testing equality and inequality restrictions. *Econometrica* 54, 1243-1248.

60. Koerts, J. and Abrahamse, A.P.J. (1969). *On the Theory and Application of the General Linear Model.* Rotterdam University Press, Rotterdam.

61. Kudô, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* 50, 403-418.

62. Kuhn, H. and Tucker, A. (1951). Non-linear programming. In J. Neyman (ed.), *Proceedings of the Second Berkeley Symposium Proceedings.* University of California Press. 481-492.

63. Lee, J.H.H. and King, M.L. (1993). A locally most mean powerful based score test for ARCH and GARCH regression disturbances. *J. Bus. Econ. Statist.* 11, 17-27.

64. Lehmann, E.L. (1986). *Testing Statistical Hypotheses.* 2nd ed. John Wiley, New York.

65. Leybourne, S.J. and McCabe, B.P.M. (1989). On the distribution of some test statistics for coefficient constancy. *Biometrika* 76, 169-177.

66. Leybourne, S.J. and McCabe, B.P.M. (1992). A simple test for parameter constancy in a nonlinear time series regression models. *Econ. Letters* 38, 157-162.

67. McCabe, B.P.M. and Leybourne, S.J. (1994). An optimal test for coefficient variation in the linear regression model. *Pak. J. Statist.* this issue.

68. Nabeya, S. and Tanaka, K. (1988). Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Ann. Statist.* 16, 218-235.

69. Neyman, J. and Pearson, E.S. (1938). Contributions to the theory of testing statistical hypotheses. II. *Statist. Research Memoirs* 2, 25-57.

70. Neyman, J. and Scott, E.L. (1967). On the use of $C(\alpha)$. optimal tests of composite hypotheses. *Bull. Inter. Statist. Instit.* 41, 477-497.

71. Nyblom, J. (1986). Testing for deterministic linear trend in time series. *J. Amer. Statist. Assoc.* 81, 545-549.

72. Nyblom, J. (1989). Testing for the constancy of parameters over time. *J. Amer. Statist. Assoc.* 84, 223-230.

73. Pierce, D.A. and Peters, D. (1992). Practical use of higher order asymptotics for multiparameter exponential families (with discussion). *J. Roy. Statist. Soc.* B 54, 701-737.

74. Rahman, S. and King, M.L. (1993). Marginal likelihood score-based tests of regression disturbances in the presence of nuisance parameters. Submitted to *J. Econometrics.*

75. Rahman, S. and King, M.L. (1994). A comparison of marginal likelihood based and approximate point optimal tests for random regression coefficients in the presence of autocorrelation. *Pak. J. Statist.* this issue.

76. Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Influence.* John Wiley, New York.

77. Rogers, A.J. (1986). Modified Lagrange multiplier tests for problems with one-sided alternatives. *J. Econometrics* 31, 341-361.

78. Rosenberg, B. (1973). The analysis of a cross-section of time series by stochastically convergent parameter regression. *Ann. Econ. Social Meas.* 2, 399-428.

79. Saikkonen, P. and Luukkonen, R. (1993). Point optimal tests for testing the order of differencing in ARIMA models. *Econometric Theory* 9, 343-362.

80. Sargan, J.D. and Bhargava, A. (1983). Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica* 51, 153-174.

81. Savin, N.E. (1980). The Bonferroni and the Scheffé multiple comparison procedures. *Rev. Econ. Stud.* 47, 255-273.

82. Savin, N.E. (1984). Multiple hypothesis testing. In Z. Griliches and M.D. Intrilligator (eds.), *Handbook of Econometrics* 2. Elsevier Science, New York.

83. Self, S.G. and Liang, K.Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *J. Amer. Statist. Assoc.* 82, 605-610.

84. Sengupta, A. (1987). On tests for equicorrelation coefficient of a standard symmetric multivariate normal distribution. *Austral. J. Statist.* 29, 49-59.

85. Sengupta, A. and Vermeire, L. (1986). Locally optimal tests for multiparameter hypotheses. *J. Amer. Statist. Assoc.* 81, 819-825.

86. Shapiro, A. (1988). Towards a unified theory of inequality constrained testing in multivariate analysis. *Inter. Statist. Rev.* 56, 49-62.

87. Shively, T.S. (1986). Issues in the analysis of stochastic coefficient regression models. Unpublished Ph.D. thesis, University of Chicago.

88. Shively, T.S. (1988a). An analysis of tests for regression coefficient stability. *J. Econometrics* 39, 367-386.

89. Shively, T.S. (1988b). An exact test for a stochastic coefficient in a time series regression model. *J. Time Series Analysis* 9, 81-88.

90. Shively, T.S. (1993). Testing for autoregressive disturbances in a time series regression with missing observations. *J. Econometrics* 57, 233-255.

91. Shively, T.S., Ansley, C.F. and Kohn, R. (1990). Fast evaluation of the distribution of the Durbin-Watson and other invariant test statistics in regression. *J. Amer. Statist. Assoc.* 85, 676-685.

92. Silvapulle, M.J. (1992a). Robust Wald-type tests of one-sided hypotheses in the linear model. *J. Amer. Statist. Assoc.* 87, 156-161.

93. Silvapulle, M.J. (1992b). Robust tests of inequality constraints and one-sided hypotheses in the linear model. *Biometrika* 79, 621-630.

94. Silvapulle, M.J. (1994). On tests against one-sided hypotheses in the generalized linear models. *Biometrics* forthcoming.

95. Silvapulle, P. (1991). Point optimal tests for non-nested time series models and illustrations. Unpublished Ph.D thesis, Monash University.

96. Silvapulle, P. and King, M.L. (1991). Testing moving average against autoregressive disturbances in the linear regression model. *J. Bus. Econ. Statist.* 9, 329-335.

97. Silvapulle, P. and King, M.L. (1993). Non-nested testing for autocorrelation in the linear regression model. *J. Econometrics* 58, 295-314.

98. Tanaka, K. (1990). The Fredholm approach to asymptotic inference on nonstationary and noninvertible time series models. *Econometric Theory* 6, 411-432.

99. Tanaka, K. and Satchell, S.E. (1989). Asymptotic properties of the maximum likelihood and nonlinear least-squares estimators for noninvertible moving average models. *Econometric Theory* 5, 333-353.

100. Thompson, M. (1982). Some results on the statistical properties of an inequality constrained least squares estimator in a linear model with two regressors. *J. Econometrics* 19, 215-232.

101. Vinod, H.D. (1973). Generalization of the Durbin-Watson statistic for higher order autoregressive processes. *Comm. Statist.* 2, 115-144.

102. Wallis, K.F. (1972). Testing for fourth order autocorrelation in quarterly regression equations. *Econometrica* 40, 617-636.

103. Wolak, F.A. (1987). An exact test for multiple inequality constraints in the linear regression model. *J. Amer. Statist. Assoc.* 82, 782-793.

104. Wolak, F.A. (1988). Duality in testing multivariate hypotheses. *Biometrika* 75, 611-615.

105. Wolak, F.A. (1989). Testing inequality constraints in linear econometric models. *J. Econometrics* 41, 205-235.

106. Wu, P.X. (1991). One-sided and partially one-sided multiparameter hypothesis testing in econometrics. Unpublished Ph.D thesis, Monash University.

107. Wu, P.X. and Bhatti, M.I. (1989). Testing for block effects and misspecification in regression models based on survey data. *Proceedings of The Third Economics Postgraduate Research Conference*, University of Western Australia, Perth.