Economic Modelling 49 (2015) 387–394

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Approximate Non-Similar critical values based tests vs Maximized Monte Carlo tests

Sivagowry Sriananthakumar *

School of Economics, Finance and Marketing, RMIT University, GPO Box 2476 Melbourne 3001, Australia

ARTICLE INFO

Article history:
Accepted 6 May 2015
Available online 17 June 2015

Keywords:
Nuisance parameters
Simulated Annealing
Approximate Non-Similar critical values
Maximized Monte Carlo
Asymptotic Maximized Monte Carlo

ABSTRACT

Testing in the presence of nuisance parameters is a problem often faced by researchers; consequently, a number of ways are suggested in the literature to manage this situation. Among these, Maximized Monte Carlo (MMC) tests or asymptotically valid MMC (AMMC) tests are becoming popular. The MMC type tests have certain advantages as well as disadvantages. This paper introduces a simple way to obtain Approximate Non-Similar (ANS) critical values using a global optimizer called Simulated Annealing (SA). All three methods are applied in the dynamic linear regression model context. As expected the AMMC approach is certainly less time consuming than the MMC approach. Overall the AMMC approach seems best in terms of power properties; however the ANS approach takes negligible time compared to its competitors. Though the ANS approach controls the sizes well it can be slightly less powerful than its competitors.

1. Introduction

In practice one often has to conduct hypothesis testing in the presence of a number of nuisance parameters, some of which may make the test non-similar in the sense that the test's size[1] varies with the values of these nuisance parameters. There is evidence that similar tests may be less powerful than non-similar tests (Andrews, 2011; Andrews et al., 2008; King and McAleer, 1987; McAleer, 1995). Also similar tests are generally asymptotic in nature (Hansen, 2003; Andrews, 2011). As economists often use sample sizes of under one hundred, it is preferable that their testing procedures have reliable finite-sample properties. A number of studies show that asymptotic tests cannot always be reliable in small samples (MacKinnon, 2009; Neto and Lima, 2010; Sriananthakumar, 2013). Consequently one needs to rely on Monte Carlo simulations or a bootstrapping method to derive critical values or p values of non-similar tests in finite samples. Bootstrap testing can work very well in some cases, but it is, in general, neither as easy nor as reliable as practitioners often appear to believe. Procedures such as the double bootstrap and fast double bootstrap may help, but this is by no means guaranteed. Also, if the rejection probabilities depend strongly on one or more nuisance parameters and those parameters were hard to estimate reliably one cannot expect a parametric bootstrap to work well (MacKinnon, 2009).

In the literature different approaches to control the sizes of the tests when testing in the presence of nuisance parameters are suggested. The classical approach to non-similar tests is to find exact non-similar critical values, for which sizes are never greater than the nominal significance level for all possible values of the nuisance parameters. Such critical values typically have to be obtained using the Monte Carlo method (Andrews et al., 2008; King and McAleer, 1987; Palomares and Roldan, 2006; Silvapulle and King, 1991). Other popular approaches include using bounds type tests and confidence intervals, as suggested in Dufour (1990), and replacing unknown nuisance parameters with consistent estimates and then relying on asymptotic theory (Moreira, 2009). However exact bounds tests can be less powerful than non-similar tests. Forchini (2005) showed that any test with size bounded from above by a known constant has potentially very low power and a large type II error.

Kiviet and Dufour (2003) and Dufour (2006) suggested a MMC approach which involves maximizing a simulated p value of a test statistic over the nuisance parameter space using SA. Dufour (2006) also suggested (and proved) AMMC tests which use a consistent set of estimators of the nuisance parameters.[2] The maximum p value of an AMMC test can be obtained by maximizing a simulated p value over a subset of nuisance parameter space (for example a confidence set for nuisance parameters) instead of an entire nuisance parameter space. Thus the AMMC approach will be less time-consuming than the MMC approach (Dufour, 2006; Phipps and Byron, 2007). The great advantage of the MMC tests is that they yield exact tests whenever the distribution of the test statistic can be simulated as a function of the nuisance parameters, also no additional assumption on distribution is needed (Dufour and Valéry, 2009).

* Tel.: +61 39925 1456; fax: +61 39925 5986.
E-mail address: Sivagowry.Sriananthakumar@rmit.edu.au.

[1] Throughout this paper the term size is used to denote probability of a type I error.

[2] This is always feasible as long as a consistent point estimate of the nuisance parameters is available.

http://dx.doi.org/10.1016/j.econmod.2015.05.006
0264-9993/

Although the MMC tests are gaining popularity (Beaulieu et al., 2013; Dufour and Tarek, 2006; Dufour and Valéry, 2009; Frederic and Olivier, 2006; Thomas et al., 2007), it is criticized for the following three reasons: (1) it can be computationally demanding (2) MMC based actual rejection frequency may be very much less than the level of the test and may, in consequence, be severely lacking power and (3) it is possible to get a much larger p value for nuisance parameter values remote from the ones that actually generated the data (MacKinnon, 2009).

This paper proposes a simple method for obtaining ANS critical values of general non-similar tests. This involves allowing SA to find values for the nuisance parameters (over the nuisance parameter space) such that the size of a non-similar test (for an arbitrary critical value) is at its maximum. Then the exact size critical value (that is, size equals to the nominal size) is calculated for those values of the nuisance parameters. As explained later, in Section 2.3, if this approach is repeated until the maximum size found is equal to the nominal level, one may determine an exact non-similar critical value. However, this iterative procedure can be extremely time-consuming and may lead to a monotonically non-increasing sequence of critical values which is strictly decreasing until the procedure converges. A more practical approach might be to stop the iterative process after one full round of the procedure and hope the exact critical value obtained at that stage is close to the exact non-similar critical value.[3] Critical values obtained this way can be regarded as ANS critical values. A number of studies use a similar approach to controlling the sizes of non-similar tests (Inder, 1986; King and McAleer, 1987; Silvapulle and King, 1991). These studies typically involve finding the maximum critical value of the test for a range of nuisance parameter values under the null hypothesis and using this critical value for further inferences. Such an approach may be suitable when few nuisance parameters are present; and may not work well when the number of nuisance parameters increases. This is what motivated the author to propose an SA based approach to obtain ANS critical values of general non-similar tests. Despite its popularity in science and engineering fields, SA is seldom used in econometrics (Dufour, 2006; Dufour and Valéry, 2009; Sriananthakumar, 2013; Sriananthakumar and King, 2006). An aim of this paper is to promote this useful algorithm in econometrics.

In this paper, the ANS approach is applied to two non-similar tests, namely, the Durbin–Watson (DW) test (Durbin and Watson, 1950, 1951) and Durbin's (1970) t test in the context of the dynamic linear regression model. The (initial) critical values for the DW and Durbin's t tests are obtained from the approximate small disturbance asymptotic (ASDA) distribution (Grant, 1987; Inder, 1986) and large-sample distribution of the test statistics, respectively (see Section 3.1 for more details). The sizes of the DW and Durbin's t tests are calculated for a variety of nuisance parameter values and design matrices, in order to check whether the SA based ANS critical values are indeed working well in terms of controlling the sizes of the tests over the nuisance parameter space. In addition, the MMC based DW test and the AMMC based DW test are also considered. Because these approaches are extremely time-consuming their application was restricted to one data set and DW test only.

Both (ANS and MMC) approaches control the sizes of the tests over the nuisance parameter space rather well. Interestingly both the methods produce similar sizes. The MMC test seems slightly more powerful than the ANS critical values based tests. In terms of computational time, there is a vast difference between the two methods. For example, MMC based tests take months to produce the required results whereas ANS based tests take hours to produce the same.[4] The AMMC approach also controls the sizes rather well except when the dynamic parameter becomes closer to 1 and σ gets bigger. Overall the AMMC approach is more powerful than its competitors. While the AMMC approach certainly takes less run time than the MMC, the ANS approach is clearly preferable in this regard.

The main contribution of this paper is in proposing a simple method to obtaining ANS critical values of general non-similar tests using SA. Unlike the MMC and AMMC approaches, this approach is less time-consuming and seems to work well in the presence of nuisance parameters. The SA based approach proposed in this paper can also be used to check which asymptotic approximation (ASDA or large sample) is best in finite samples or (for example) whether the standard normal distribution or Student's t distribution is more appropriate for the null distribution of Durbin's t test statistic in finite samples. However, this method can also be criticized (as can the MMC type approach) because it is possible to get a much larger size for nuisance parameter values distant from the ones that actually generated the data. Use of small SA parameter values may alleviate this problem and produce reasonably good sizes in a short time, as is the case in this study.

The plan of this paper is as follows. The theory, including how SA can be effectively used to obtain ANS critical values, is discussed in Section 2. This theory is applied in Section 3 to the problem of testing for autocorrelation in the dynamic linear regression model. Section 4 explains the Monte Carlo data generating process while Section 5 presents the details of the Monte Carlo experiment and its main findings. Finally, some concluding remarks are given in Section 6.

2. Theory

In explaining the theory behind this approach, let y be an observable $n \times 1$ vector which has probability density

$$f(y; \vartheta, \phi, \psi)$$

where ϑ, ϕ and ψ are $u \times 1$, $v \times 1$ and $w \times 1$ vectors of unknown parameters. Suppose we wish to test

$$H_0 : \vartheta = \vartheta_0 \text{ against } H_a : \vartheta > \vartheta_0 (\text{or } \vartheta < \vartheta_0 \text{ or } \vartheta \neq \vartheta_0),$$

where ϑ_0 is a known $u \times 1$ vector. Then ϕ and ψ are vectors of nuisance parameters. Suppose we have a test statistic $T(y)$ with a null distribution which is invariant with respect to ϕ but depends on ψ. The $T(y)$ test sizes vary with values of ψ, making the test non-similar.

$$\text{Let } s(\psi) = \Pr[T(y) > c | f(y; \vartheta_0, \phi, \psi)]. \tag{1}$$

We wish to obtain the critical value c such that

$$\underset{\psi}{Sup}\ s(\psi) = \alpha, \tag{2}$$

where α is the desired significance level of the test.

Unfortunately, typically the finite sample distribution of $T(y)$ under H_0 (in the presence of unavoidable nuisance parameters) is unknown and we have to simulate it by the Monte Carlo method. The Monte Carlo method results in a step-type function for probability given in Eq. (1) which typically has zero derivatives almost everywhere, except on isolated points where it is not differentiable. Further the supremum (2) is typically not unique (that is, several values of nuisance parameters can yield the same supremum). Therefore standard derivative based optimization routines do not work in this case. However, the required maximizations can be performed by using special optimization techniques, such as SA, which do not require differentiability.

2.1. A brief introduction to SA

An intelligent random-search technique, SA was developed by Kirkpatrick et al. in 1983 to deal with highly nonlinear problems. SA's

[3] This assumes little change in the values of the nuisance parameters that maximize size for different critical values.

[4] The ANS method takes less than 45 min for a large data set ($n = 76$) and for a small data set ($n = 20$) it takes less than 4 min.