Statistics and Probability Letters 83 (2013) 1677–1682

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

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

Jackknife estimation with a unit root

Marcus J. Chambers [a,1], Maria Kyriacou [b,*]

[a] Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England, United Kingdom

[b] Economics Division, School of Social Sciences, University of Southampton, Highfield, SO17 1BJ, United Kingdom

ARTICLE INFO

Article history:
Received 23 August 2012
Received in revised form 1 February 2013
Accepted 20 March 2013
Available online 30 March 2013

Keywords:
Jackknife
Bias reduction
Unit root

ABSTRACT

We study jackknife estimators in a first-order autoregression with a unit root. Non-overlapping sub-sample estimators have different limit distributions, so the jackknife does not fully eliminate first-order bias. We therefore derive explicit limit distributions of the numerator and denominator to calculate the expectations that determine optimal jackknife weights. Simulations show that the resulting jackknife estimator produces substantial reductions in bias and RMSE.

1. Introduction

A longstanding bias reduction method whose properties have been less widely explored in autoregressive models is the jackknife (Quenouille, 1956; Tukey, 1958). In recent work, Chambers (2013) investigated jackknife methods in a stationary autoregression and Phillips and Yu (2005) used the jackknife to estimate the parameters of a continuous time model and associated bond option prices. In all the above contributions the properties of the jackknife as a bias reduction device are confirmed by significant bias reductions in simulation experiments. As observed by Quenouille (1956) in his original paper, autoregressive parameter estimators typically suffer from negative bias. The nature of the bias in stationary autoregression has been extensively studied and its properties are well understood. For example, early contributions to this topic can be found in Marriott and Pope (1954), Kendall (1954) and Shenton and Johnson (1965). This bias is large in the unit root case and it is, therefore, particularly interesting to ascertain the extent of bias reduction that can be achieved by jackknife methods in this setting.

The focus of this paper is jackknife estimation in unit root regression. We consider the jackknife proposed by Phillips and Yu (2005) based on non-overlapping sub-samples which was found to perform well by Chambers (2013). In the presence of a unit root, we show that the jackknife in its usual formulation fails to fully eliminate the first-order bias. The source of this failure lies in the different limit distributions of the sub-samples. These distributions motivate a set of optimal jackknife weights that ensure the first-order bias is fully removed. Simulations reveal that the 'optimal' jackknife estimator proposed here produces further bias and root mean squared error (RMSE) reductions.

The following notation is used throughout. The symbol $\overset{d}{=}$ denotes equality in distribution, $\overset{d}{\to}$ convergence in distribution, $\overset{p}{\to}$ convergence in probability, $\Rightarrow$ weak convergence of the relevant probability measures, and $W(r)$ a Wiener process on $C[0,1]$, the space of continuous real-valued functions on the unit interval. Functionals of $W(r)$, such as $\int_0^1 W(r)^2 dr$ are denoted $\int_0^1 W^2$.

* Corresponding author. Tel.: +44 23 80593645.
E-mail addresses: mchamb@essex.ac.uk (M.J. Chambers), M.Kyriacou@soton.ac.uk (M. Kyriacou).

1 Tel.: +44 1206 872756.

0167-7152/$ – see front matter
http://dx.doi.org/10.1016/j.spl.2013.03.016

2. Jackknife estimation with a unit root

Let the data be generated by the random walk

$$y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim iid\, N(0, \sigma^2),\ t = 1, \ldots, n, \tag{1}$$

where y_0 can be any observed $O_p(1)$ random variable. Ordinary least squares (OLS) regression gives

$$y_t = \hat{\beta} y_{t-1} + \hat{\epsilon}_t, \quad t = 1, \ldots, n, \tag{2}$$

where $\hat{\epsilon}_t$ is the regression residual and the coefficient satisfies

$$n(\hat{\beta} - 1) = \frac{n^{-1} \sum_{t=1}^{n} y_{t-1}\epsilon_t}{n^{-2} \sum_{t=1}^{n} y_{t-1}^2} \Rightarrow \frac{\int_0^1 WdW}{\int_0^1 W^2} \quad \text{as } n \to \infty. \tag{3}$$

The limit distribution in (3) is skewed and the estimator suffers from significant negative finite sample bias. Assuming $y_0 = 0$, Phillips (1987, Theorem 7.1) demonstrated the validity of an asymptotic expansion given by

$$n(\hat{\beta} - 1) \stackrel{d}{=} \frac{\int_0^1 WdW}{\int_0^1 W^2} - \frac{\eta}{\sqrt{2n}\int_0^1 W^2} + O_p(n^{-1}), \tag{4}$$

where $\eta \sim N(0, 1)$ and is distributed independently of W. Taking expectations in (4) and noting that the expected value of the leading term is -1.781 (see, for example, Table 7.1 of Tanaka, 1996), the bias satisfies

$$E(\hat{\beta} - 1) = -\frac{1.781}{n} + O(n^{-2}), \tag{5}$$

an expansion that motivates the use of the jackknife as a method of bias reduction.

The jackknife offers a simple method of reducing bias by eliminating the leading bias term. The jackknife estimator combines the full-sample estimator, $\hat{\beta}$, with a set of m sub-sample estimators, $\hat{\beta}_j$ ($j = 1, \ldots, m$). The weights assigned to these components depend on the type of sub-sampling employed. In stationary autoregression, Chambers (2013) compares alternative methods of sub-sampling and finds non-overlapping ones to perform best in reducing bias. The jackknife estimator is

$$\hat{\beta}_J = \kappa_m \hat{\beta} + \delta_m \frac{1}{m} \sum_{j=1}^{m} \hat{\beta}_j, \tag{6}$$

where the non-overlapping weights are given by $\kappa_m = m/(m-1)$ and $\delta_m = -1/(m-1)$ and the sub-sample length is ℓ, with $n = m \times \ell$. These weights are determined on the assumption that each sub-sample estimator also satisfies (5):

$$E(\hat{\beta}_j - 1) = -\frac{1.781}{\ell} + O(\ell^{-2}), \quad j = 1, \ldots, m. \tag{7}$$

Under such circumstances the jackknife estimator is capable of completely eliminating the $O(n^{-1})$ bias term in the estimator as compared to $\hat{\beta}$.

2.1. Sub-sample limit distribution

The sub-sample estimators, however, do not share the same limit distribution as the full-sample estimator, which means that the bias expansions for the sub-sample estimators are incorrect. The reason is that the sub-sample initial value is the accumulated sum of all previous innovations and the initialization is not eliminated in the asymptotics. Let $\tau_j = \{(j-1)\ell + 1, \ldots, j\ell\}$ denote the set of integers in sub-sample j ($j = 1, \ldots, m$). Under (1), the observations in sub-sample j satisfy

$$y_t = y_{t-1} + \epsilon_t = y_{(j-1)\ell} + \sum_{i=(j-1)\ell+1}^{t} \epsilon_i, \quad t \in \tau_j, \tag{8}$$

and so the initial value, $y_{(j-1)\ell}$, is $O_p\left(\sqrt{(j-1)\ell}\right)$ rather than $O_p(1)$ or a constant.

The sub-sample estimator can be written

$$\hat{\beta}_j - 1 = \frac{\sum_{t\in\tau_j} y_{t-1}\epsilon_t}{\sum_{t\in\tau_j} y_{t-1}^2}, \quad j = 1, \ldots, m. \tag{9}$$