Statistics and Probability Letters 80 (2010) 1348–1353

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

The asymptotic efficiency of improved prediction intervals

Paul Kabaila *, Khreshna Syuhada

La Trobe University, Australia
Institut Teknologi Bandung, Indonesia

ARTICLE INFO

Article history:
Received 29 October 2009
Received in revised form 20 April 2010
Accepted 26 April 2010
Available online 10 May 2010

Keywords:
Asymptotic efficiency
Estimative prediction limit
Improved prediction limit

ABSTRACT

We consider the Barndorff-Nielsen and Cox (1994, p. 319) method of modifying an estimative prediction interval to obtain an improved prediction interval with better conditional coverage properties. The parameter estimator, on which this improved interval is based, is assumed to have the same asymptotic distribution as the conditional maximum likelihood estimator. This improved interval depends strongly on the asymptotic conditional bias of this estimator, which can be very sensitive to small changes in this estimator. We show, however, that the asymptotic efficiency of this improved prediction interval does *not* depend on this bias.

1. Introduction

Suppose that $\{Y_t\}$ is a discrete-time stochastic process with probability distribution determined by the parameter vector θ, where the Y_t are continuous random variables. Also suppose that $\{Y^{(t)}\}$ is a Markov process, where $Y^{(t)} = \left(Y_{t-p+1}, \ldots, Y_t\right)$. For example, $\{Y_t\}$ may be an AR(p) process or an ARCH(p) process. The available data are $Y_1, \ldots, Y_n$. Suppose that we are concerned with k-step-ahead prediction where k is a specified positive integer. We use lower case to denote the observed value of a random vector. For example, $y^{(n)}$ denotes the observed value of the random vector $Y^{(n)}$.

Firstly, suppose that our aim is to find an upper prediction limit $z(Y_1, \ldots, Y_n)$, for Y_{n+k}, such that it has coverage probability conditional on $Y^{(n)} = y^{(n)}$ equal to $1-\alpha$ i.e. such that

$$P_\theta\left(Y_{n+k} \le z(Y_1, \ldots, Y_n) \mid Y^{(n)} = y^{(n)}\right) = 1 - \alpha$$

for all θ and $y^{(n)}$. The desirability of a prediction limit or interval having coverage probability $1-\alpha$ conditional on $Y^{(n)} = y^{(n)}$ has been noted by a number of authors. In the context of an AR(p) process, this has been noted by Phillips (1979), Stine (1987), Thombs and Schucany (1990), Kabaila (1993), McCullough (1994), He (2000), Kabaila and He (2004) and Vidoni (2004). In the context of an ARCH(p) process, this has been noted by Christoffersen (1998), Kabaila (1999), Vidoni (2004) and Kabaila and Syuhada (2008).

Define $z_\alpha(\theta, y^{(n)})$ by the requirement that $P_\theta\left(Y_{n+k} \le z_\alpha(\theta, y^{(n)}) \mid Y^{(n)} = y^{(n)}\right) = 1-\alpha$ for all θ and $y^{(n)}$. The estimative $1-\alpha$ prediction limit is defined to be $z_\alpha(\widehat{\Theta}, Y^{(n)})$, where $\widehat{\Theta}$ is an estimator of θ with the same asymptotic distribution as the conditional maximum likelihood estimator of θ. This prediction limit may not have adequate coverage probability properties, unless n is very large. In Section 2, we recap the argument (due to Cox, 1975, p. 50) showing that the coverage probability of $z_\alpha(\widehat{\Theta}, Y^{(n)})$ conditional on $Y^{(n)} = y^{(n)}$ that is $1-\alpha+O(n^{-1})$.

Methods for obtaining prediction limits with better asymptotic coverage properties, conditional on $Y^{(n)} = y^{(n)}$, than the estimative prediction limit have been described by Cox (1975), Barndorff-Nielsen and Cox (1994, p. 319), Corcuera (2001,

* Corresponding address: Department of Mathematics and Statistics, La Trobe University, Victoria 3086, Australia. Tel.: +61 3 9479 2594; fax: +61 3 9479 2466.
E-mail address: P.Kabaila@latrobe.edu.au (P. Kabaila).

0167-7152/$ – see front matter
doi:10.1016/j.spl.2010.04.016

section 4), Vidoni (2004, 2009) and Kabaila and Syuhada (2008). For convenience, we refer to Barndorff-Nielsen and Cox (1994, p. 319) as BNC94. The key advantages of the BNC94 method are its simplicity, ease of explanation and accessibility to a relatively wide audience. In Section 2, we recap the argument that shows that the BNC94-improved $1-\alpha$ prediction limit has coverage probability conditional on $Y^{(n)} = y^{(n)}$ that is $1-\alpha+O(n^{-3/2})$.

The BNC94 method is applicable not only for $\widehat{\Theta}$ the conditional maximum likelihood estimator or θ, but also for any estimator $\widehat{\Theta}$ with the same asymptotic distribution as this estimator. There are many possible choices for $\hat{\Theta}$. For example, for a stationary Gaussian AR(1) model, commonly-used estimators of the autoregressive parameter include least-squares, Yule–Walker and Burg estimators. It is natural to ask the following question. What difference does it make to the BNC94-improved prediction limit which estimator $\widehat{\Theta}$ is used? The BNC94-improved $1-\alpha$ prediction limit is obtained from the estimative $1-\alpha$ prediction limit using a correction that includes the asymptotic bias of $\widehat{\Theta}$ conditional on $Y^{(n)} = y^{(n)}$. This asymptotic conditional bias can be very sensitive to small changes in the estimator $\widehat{\Theta}$ (Kabaila and Syuhada, 2007, Section 4 and Syuhada, 2008). For example, the Yule–Walker and least-squares estimators of the autoregressive parameter of a stationary Gaussian AR(1) model have quite different asymptotic conditional biases. So, the form of the BNC94-improved $1-\alpha$ prediction limit will typically depend very strongly on the choice of $\widehat{\Theta}$. What difference does this make to the asymptotic efficiency of the BNC94-improved prediction limit? In Section 2, we show that the asymptotic efficiency of this improved prediction limit is *not* influenced by which estimator $\widehat{\Theta}$ is used.

We extend these results to prediction *intervals* as follows. In Section 3, we show that the estimative $1-\alpha$ prediction interval has coverage probability conditional on $Y^{(n)} = y^{(n)}$ is $1-\alpha+O(n^{-1})$. In this section, we also present a modification of an estimative $1-\alpha$ prediction interval, analogous to the BNC94 modification of an estimative prediction limit, to obtain an improved $1-\alpha$ prediction interval with better coverage properties. We show that this improved $1-\alpha$ prediction interval has coverage probability conditional on $Y^{(n)} = y^{(n)}$ that is $1-\alpha+O(n^{-3/2})$. This modification involves the use of a correction that includes the asymptotic bias of $\widehat{\Theta}$ conditional on $Y^{(n)} = y^{(n)}$. Does it make any difference to the asymptotic efficiency of this improved prediction interval which estimator $\widehat{\Theta}$ of θ is used? In Section 3, we show that it does *not* make any difference which estimator is used. In Section 5, we present an illustration of this result.

The fact that the asymptotic efficiencies of these prediction limits and intervals do not depend on which estimator $\widehat{\Theta}$ of θ is used, provided that it has the same asymptotic distribution as the conditional maximum likelihood estimator, has the following consequence. We can use that estimator $\widehat{\Theta}$ whose asymptotic conditional bias is easiest to find. Usually, this will be conditional maximum likelihood estimator whose asymptotic conditional bias can be found using the formula of Vidoni (2004, p. 144).

When the extensive algebraic manipulations required to find the BNC94-improved prediction limit or interval are too messy, we can use the Kabaila and Syuhada (2008) simulation-based methodology to find approximations to this improved prediction limit or interval. The implications of the asymptotic efficiency results described in Sections 2 and 3 for these approximations are described in Section 4.

2. Asymptotic efficiency result for improved prediction limits

In this section we recap the argument, due to Cox (1975), that the conditional coverage probability of the estimative $1-\alpha$ upper prediction limit is $1-\alpha+O(n^{-1})$. We then recap the argument, due to BNC94, that the conditional coverage probability of their improved $1-\alpha$ upper prediction limit is $1-\alpha+O(n^{-3/2})$. This prediction limit includes a correction term that depends strongly on the asymptotic conditional bias of $\widehat{\Theta}$. We show, however, that the asymptotic efficiency of this prediction limit (which we measure to $O(n^{-1})$) does *not* depend on this bias (which we also measure to $O(n^{-1})$).

Let $F(\,\cdot\,;\theta,y^{(n)})$ denote the cumulative distribution function of Y_{n+k}, conditional on $Y^{(n)} = y^{(n)}$. Also, let $f(\,\cdot\,;\theta,y^{(n)})$ denote the probability density function corresponding to this cumulative distribution function. Assume, as do Cox (1975), BNC94 and Vidoni (2004), that

$$E_\theta\big(\widehat{\Theta}-\theta \mid Y^{(n)} = y^{(n)}\big) = b(\theta,y^{(n)})n^{-1} + \cdots \tag{1}$$

$$E_\theta\big((\widehat{\Theta}-\theta)(\hat{\Theta}-\theta)^T \mid Y^{(n)} = y^{(n)}\big) = i^{-1}(\theta) + \cdots \tag{2}$$

where $i(\theta)$ denotes the expected information matrix. We assume that every element of $i(\theta)$ is $O(n^{-1})$. Henceforth, we use the Einstein summation notation that repeated indices are implicitly summed over.

Define $H_\alpha(\theta|y^{(n)}) = P_\theta\big(Y_{n+k} \le z_\alpha(\widehat{\Theta},y^{(n)}) \mid Y^{(n)} = y^{(n)}\big)$, which is the conditional coverage probability of the $1-\alpha$ estimative prediction limit. Using the fact that the distribution of Y_{n+k} given $(Y_1,\ldots,Y_n) = (y_1,\ldots,y_n)$ depends only on $y^{(n)}$, it may be shown that $H_\alpha(\theta|y^{(n)}) = E_\theta\big(F(z_\alpha(\widehat{\Theta},y^{(n)});\theta,y^{(n)}) \mid Y^{(n)} = y^{(n)}\big)$. Now define $G_\alpha(\widehat{\Theta};\theta|y^{(n)}) = F(z_\alpha(\widehat{\Theta},y^{(n)});\theta,y^{(n)})$. Thus $H_\alpha(\theta|y^{(n)}) = E_\theta\big(G_\alpha(\widehat{\Theta};\theta|y^{(n)}) \mid Y^{(n)} = y^{(n)}\big)$. We now use the stochastic expansion

$$G_\alpha(\widehat{\Theta};\theta|y^{(n)}) = G_\alpha(\theta;\theta|y^{(n)}) + \left.\frac{\partial G_\alpha(\hat{\theta};\theta|y^{(n)})}{\partial\hat{\theta}_i}\right|_{\hat{\theta}=\theta}(\widehat{\Theta}_i-\theta_i) + \frac{1}{2}\left.\frac{\partial^2 G_\alpha(\hat{\theta};\theta|y^{(n)})}{\partial\hat{\theta}_r\partial\hat{\theta}_s}\right|_{\hat{\theta}=\theta}(\widehat{\Theta}_r-\theta_r)(\widehat{\Theta}_s-\theta_s) + \cdots. \tag{3}$$