Journal of Statistical Planning and Inference 140 (2010) 297 – 309

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Corrected confidence intervals for parameters in adaptive linear models

Shen-Chien Chen[a], Ruby C. Weng[b,*], Tzeeming Huang[b]

[a] Chungyu Institute of Technology, Keelung 20103, Taiwan
[b] National Chengchi University, Taipei 11623, Taiwan

ARTICLE INFO

Article history:
Received 6 December 2008
Received in revised form
10 July 2009
Accepted 10 July 2009
Available online 18 July 2009

Keywords:
Confidence sets
Threshold autoregressive model
Very weak expansions

ABSTRACT

Consider an adaptive linear model $y_t = x_t'\theta + \sigma e_t$, where $x_t = (x_{t1}, \ldots, x_{tp})'$ may depend on previous responses. Woodroofe and Coad [1999. Corrected confidence sets for sequentially designed experiments: examples. In: Ghosh, S. (Ed.), Multivariate Analysis, Design of Experiments, and Survey Sampling. Marcel Dekker, Inc., New York, pp. 135–161] derived very weak asymptotic expansions for the distributions of an appropriate pivotal quantity and constructed corrected confidence sets for θ, where the correction terms involve the limit of $\sum_{t=1}^{n} x_t x_t'/n$ (as n approaches infinity) and its derivatives with respect to θ. However, the analytic form of this limit and its derivatives may not be tractable in some models. This paper proposes a numerical method to approximate the correction terms. For the resulting approximate pivot, we show that under mild conditions the error induced by numerical approximation is $o_p(1/n)$. Then, we assess the accuracy of the proposed method by an autoregressive model and a threshold autoregressive model.

1. Introduction

Consider an adaptive linear model of the form

$$y_t = x_t'\theta + \sigma e_t, \quad t = 1, 2, \ldots, \tag{1}$$

where $e_1, e_2, \ldots$ are i.i.d. standard normal random variables and $\theta = (\theta_1, \ldots, \theta_p)'$ and $\sigma > 0$ are unknown parameters. Here "adaptive" means that $x_t = (x_{t1}, \ldots, x_{tp})'$ may depend on previous responses; that is $x_t = x_t(y_1, \ldots, y_{t-1})$. The above model is quite general and includes time series and control problems as in Lai and Wei (1982), adaptive biased coin designs as in Eisele (1994), among others.

An adaptive normal linear model with known variance was studied by Woodroofe (1989) and Woodroofe and Coad (1997). They used a Bayesian approach and Stein's (1981) identity to obtain asymptotic expansions for sampling distributions and construct corrected confidence sets for θ correct to order $o(1/n)$. The case of unknown σ was considered by Coad and Woodroofe (1998) and Woodroofe and Coad (1999), and approximations for confidence sets were evaluated for autoregressive (AR) processes, Ford–Silvey model, and certain clinical trial examples. It was shown that the maximum likelihood estimators may be severely biased in these models. This Bayesian approach starts with an approximate pivot, and employs Stein's identity to derive asymptotic expansions for the mean and variance corrections of the pivot. Then, it proceeds in the usual way to obtain the renormalized pivot, which is used to form corrected confidence sets. The correction terms have simple expressions, which involve the analytic

* Corresponding author.
E-mail address: chweng@nccu.edu.tw (R.C. Weng).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2009.07.014

forms of the limit of $\sum_{t=1}^{n} x_t x_t'/n$ (as n approaches infinity) and its derivatives with respect to the parameter θ. For example, in AR(2) models

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \sigma e_t, \quad t = 0, 1, \ldots, \tag{2}$$

the correction terms depend on the first two moments of the process and the derivatives of these moments, which are tractable. However, in some models the correction terms may not be available. Consider the following two-regime threshold autoregressive (TAR) model

$$y_t = \theta_1 y_{t-1}^{+} + \theta_2 y_{t-1}^{-} + \sigma e_t, \quad t = 0, 1, \ldots, \tag{3}$$

where

$$y_{t-1}^{+} = \begin{cases} y_{t-1} & \text{if } y_{t-1} > 0, \\ 0 & \text{if } y_{t-1} \le 0 \end{cases} \quad \text{and} \quad y_{t-1}^{-} = \begin{cases} 0 & \text{if } y_{t-1} > 0, \\ y_{t-1} & \text{if } y_{t-1} \le 0. \end{cases}$$

The TAR model was proposed by Tong (1983, 1990) to characterize certain nonlinear features of a process, and it became quite popular in the non-linear time series literature. In this model, we have

$$\frac{1}{n}\sum_{t=1}^{n} x_t x_t' = \begin{pmatrix} \frac{1}{n}\sum_{t=1}^{n}(y_{t-1}^{+})^2 & 0 \\ 0 & \frac{1}{n}\sum_{t=1}^{n}(y_{t-1}^{-})^2 \end{pmatrix} \to \begin{pmatrix} E_\theta(S^{+})^2 & 0 \\ 0 & E_\theta(S^{-})^2 \end{pmatrix}, \tag{4}$$

where S follows the stationary distribution of the process. However, it is known that the explicit analytic forms for the stationary distribution and moments of a simple TAR model are difficult to derive and are known only in certain cases where the autoregression function has special structures and the error term follows some specific distributions. For example, Anděl et al. (1984) studied model (3) with $\theta_1 = -\theta_2$, $\theta_1 \in (0, 1)$, and e_t follows N(0, 1), and Anděl and Bartoň (1986) and Loges (2004) considered the same model with Cauchy and Laplace distributions respectively. Consequently, for TAR models the corrections suggested by very weak type approximations cannot be obtained in the usual way. To address this problem, we propose to approximate the limit of $\sum_{t=1}^{n} x_t x_t'/n$ and its derivatives by combining the difference quotient method and Monte Carlo simulations. Then, we show that under mild conditions the obtained confidence intervals are accurate to order $o_p(1/n)$, where o_p is in the sense of (31).

We organize the remainder of this paper as follows. The next section gives brief review of very weak approximation for adaptive linear models. In Section 3 we describe our method and conduct error analysis. In Section 4 we assess the accuracy of the proposed method by simulation studies. Section 5 concludes the paper.

2. Review

It is known that the likelihood function is not affected by the adaptive nature of the design, so the maximum likelihood estimator of θ has the form

$$\hat{\theta}_n = \left(\sum_{t=1}^{n} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{n} x_t y_t\right),$$

provided that $\sum_{t=1}^{n} x_t x_t'$ is positive definite. The usual estimator of σ^2 is

$$\hat{\sigma}_n^2 = \frac{\sum_{t=1}^{n}(y_t - x_t'\hat{\theta}_n)^2}{n-p}.$$

Let B_n be a $p \times p$ matrix for which

$$B_n B_n' = \sum_{t=1}^{n} x_t x_t', \tag{5}$$

and define

$$Z_n = \frac{1}{\sigma} B_n'(\theta - \hat{\theta}_n), \quad T_n = \frac{1}{\hat{\sigma}_n} B_n'(\theta - \hat{\theta}_n). \tag{6}$$

These are served as the first approximate pivots for known and unknown σ respectively. The variables Z_n and T_n have exactly p-variate standard normal and t distributions respectively, in the absence of an adaptive design. For the case of an adaptive design, the bias correction is needed. To describe the correction term, we first introduce some notations. Let $P_{\theta,\sigma}$ denote the probability distribution of the model, and $E_{\theta,\sigma}$ the expectation with respect to $P_{\theta,\sigma}$. Since $\hat{\theta}_n$ is invariant and B_n is equivalent, the distributions