Journal of Statistical Planning and Inference 97 (2001) 293–304

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Fixed size confidence regions for parameters of threshold AR(1) models

T.N. Sriram*, [1]

Department of Statistics, The University of Georgia, 204 Statistics Building, Athens, GA 30602-1952, USA

Received 3 May 1999; received in revised form 29 March 2000; accepted 12 October 2000

Abstract

For parameters of single and multiple threshold autoregressive models of order one, sequential procedures are proposed for constructing fixed size confidence ellipsoids. Sequential procedures are also proposed for constructing fixed proportional accuracy confidence ellipsoids and fixed width confidence intervals for linear combination of parameters. The confidence ellipsoids and intervals are shown to be asymptotically consistent and the associated stopping rules are shown to be asymptotically efficient as the size/width of the region becomes small. © 2001 Elsevier Science B.V. All rights reserved.

MSC: primary 62L12; 60M10; secondary 62F25; 60G10

Keywords: TAR models; Ergodicity; Fixed size confidence ellipsoids; Stopping rule; Asymptotic consistency; Asymptotic efficiency; Uniform integrability

1. Introduction

It is well documented in the literature that sequential sampling methods provide a useful way of constructing confidence intervals/regions (for parameters) with fixed size and prescribed coverage probability. In a seminal paper, Chow and Robbins (1965) proposed a recipe for constructing a sequential fixed width confidence interval for an unknown mean with prescribed probability. Their ideas have been used to develop sequential fixed size confidence regions for higher dimensional cases and regression models; see, for example, Gleser (1965), Albert (1966), Srivastava (1967, 1971) and Finster (1985). More recently, sequential confidence regions based on maximum likelihood estimators have been constructed by other authors, for example, Grambsch (1983, 1989), Yu (1989) and Chang and Martinsek (1992). Grambsch (1989) and Chang and

* Fax: +706-542-3391.

E-mail address: tn@stat.uga.edu (T.N. Sriram).

[1] Supported by a grant from NSA #MDA904-00-1-0025.

PII: S0378-3758(00)00246-9

Martinsek (1992) construct sequential fixed size confidence regions for parameters of a logistic regression model using different stopping rules.

For over a decade or so, there has been a steadily growing interest in using sequential methods to estimate parameters in linear time-series models. See for example, Lai and Siegmund (1983), Sriram (1987, 1988), Greenwood and Shiryaev (1992), Fakhre-Zakeri and Lee (1992, 1993) and Lee (1994). There are many situations, however, where one would *not* expect linear time-series models to be the best class of models to fit a real data set, although one may tacitly assume that the linear time-series model under consideration provides a close approximation to physical reality.

One class of non-linear time-series models which is generally agreed to be useful is the class of threshold autoregressive (TAR) models introduced by Tong (1978) and discussed extensively in Tong and Lim (1980). Recently, Lee and Sriram (1999) considered the problem of sequential point estimation of the parameters in a TAR(1) model where they studied the first-order properties of the risk of sequential procedures involved.

In this paper, we, once again, consider a TAR(1) model defined by

$$X_i = \theta_1 X_{i-1}^{+} + \theta_2 X_{i-1}^{-} + \varepsilon_i, \quad i = 1, 2, \ldots, \tag{1.1}$$

where the real parameters θ_1 and θ_2 are not necessarily equal, $\{\varepsilon_i\}$ is a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s), and $x^+ = \max(x, 0)$ and $x^- = \min(x, 0)$ for a real number x. The distribution of error ε_1 is unspecified but it is assumed throughout that $E\varepsilon_1 = 0 < E\varepsilon_1^2 = \sigma^2 < \infty$, where σ^2 is an unknown constant.

Our aim here, however, is to construct a sufficiently precise confidence ellipsoid for $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ in the two-dimensional Euclidean space. That is, we wish to construct an ellipsoidal region $\boldsymbol{R}_n$ such that the length of the major axis is equal to $2d(d > 0)$, and such that the coverage probability, $P(\boldsymbol{\theta} \in \boldsymbol{R}_n)$, is approximately equal to $1 - \alpha$ $(0 < \alpha < 1)$ for sufficiently small values of d.

It has been shown by Petruccelli and Woolford (1984) that the process $\{X_i; i \geqslant 0\}$ defined in (1.1) is ergodic if and only if

$$\boldsymbol{\theta} \in \Theta = \{(\theta_1, \theta_2)' \colon \theta_1 < 1, \theta_2 < 1 \text{ and } \theta_1\theta_2 < 1\}. \tag{1.2}$$

This implies the existence of an invariant probability distribution for $\{X_i\}$. Chan et al. (1985) have extended the above-mentioned result of Petruccelli and Woolford to a multiple-threshold AR(1) model. Furthermore, Chan et al. (1985) have shown that $E|\varepsilon_1|^k < \infty$ for some integer $k \geqslant 1$ implies that the invariant probability distribution for the chain $\{X_i\}$ has finite kth moment for each $\boldsymbol{\theta} \in \Theta$; see Chan et al. (1985), Theorem 2.3 and the remark following it. In what follows we shall assume that X_0 has as its distribution $\pi(\cdot)$, the invariant probability distribution for $\{X_i\}$, so that the process is strictly stationary. Also, we will denote $(X_0^{\pm})^k$ and $(X_i^{\pm})^k$ by $X_0^{\pm k}$ and $X_i^{\pm k}$, respectively, for $k \geqslant 1$.

Suppose we estimate the parameters θ_1 and θ_2 in (1.1) by their least-squares estimators

$$\hat{\theta}_{1,n} = \sum_{i=1}^{n} X_i X_{i-1}^{+} \bigg/ \sum_{i=1}^{n} X_{i-1}^{+2} \tag{1.3}$$

and

$$\hat{\theta}_{2,n} = \sum_{i=1}^{n} X_i X_{i-1}^{-} \bigg/ \sum_{i=1}^{n} X_{i-1}^{-2}. \tag{1.4}$$

Then the corresponding estimator of σ^2 is $\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^{n} (X_i - \hat{\theta}_{1,n} X_{i-1}^{+} - \hat{\theta}_{2,n} X_{i-1}^{-})^2$. Note that the estimators defined above are also the maximum likelihood estimators for θ_1, θ_2 and σ^2, respectively, under the assumption of normal error distribution.

It is shown in Petruccelli and Woolford (1984) that if $\boldsymbol{\theta} \in \Theta$ defined in (1.2), then the estimators $\hat{\theta}_{1,n}$, $\hat{\theta}_{2,n}$ and $\hat{\sigma}_n^2$ are strongly consistent for θ_1, θ_2 and σ^2, respectively. Furthermore, if $\boldsymbol{\theta} \in \Theta$, then it can be concluded using the result $n^{-1} \sum_{i=1}^{n} X_{i-1}^{\pm 2} \to EX_0^{\pm 2}$ almost surely (a.s.) as $n \to \infty$ (follows from ergodicity) and Theorem 3.2 of Petruccelli and Woolford (1984) that for $\Gamma_n = \mathrm{diag}(\sum_{i=1}^{n} X_{i-1}^{+2}, \sum_{i=1}^{n} X_{i-1}^{-2})$, a diagonal matrix, and $\hat{\boldsymbol{\theta}}_n = (\hat{\theta}_{1,n}, \hat{\theta}_{2,n})'$ we have

$$\sigma^{-2} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})' \Gamma_n (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{D} \chi_2^2 \quad \text{as } n \to \infty, \tag{1.5}$$

where χ_2^2 is a χ^2 random variable with two degrees of freedom. Now, for any $d > 0$, let

$$\boldsymbol{R}_n = \{\boldsymbol{z}\colon (\boldsymbol{z} - \hat{\boldsymbol{\theta}}_n)' \Gamma_n (\boldsymbol{z} - \hat{\boldsymbol{\theta}}_n) \leqslant d^2 \lambda_{\min}(\Gamma_n)\}, \tag{1.6}$$

where $\lambda_{\min}(\Gamma_n) = \min(\sum_{i=1}^{n} X_{i-1}^{+2}, \sum_{i=1}^{n} X_{i-1}^{-2})$ is the smallest eigenvalue of Γ_n defined above. Then, $\boldsymbol{R}_n$ defines an ellipsoid with length of the major axis equal to $2d$ and it is in this sense that the size of the ellipsoid is fixed. Moreover, for any $\alpha \in (0,1)$ and $n_0(d)$ determined by

$$n_0(d) = \text{smallest integer} \geqslant \sigma^2 a^2 / [d^2 \min(EX_0^{+2}, EX_0^{-2})], \tag{1.7}$$

where a^2 satisfies $P[\chi_2^2 \leqslant a^2] = 1 - \alpha$, we have from (1.5) that for $\boldsymbol{\theta} \in \Theta$

$$\lim_{d \to 0} P(\boldsymbol{\theta} \in \boldsymbol{R}_{n_0(d)}) = 1 - \alpha. \tag{1.8}$$

The result in (1.8) shows that, for small values of d, the sample size $n_0(d)$ yields an ellipsoidal confidence region of fixed size and prespecified coverage probability. However, the sample size $n_0(d)$ cannot be used in practice because it depends on the unknown parameters. To over come this, we define a stopping rule

$$T_d = \inf\{n \geqslant m\colon \lambda_{\min}(\Gamma_n) \geqslant \hat{\sigma}_n^2 a^2 / d^2\}, \tag{1.9}$$

where $\lambda_{\min}(\Gamma_n)$ is as defined in (1.6), $\hat{\sigma}_n^2$ is defined as above and $m(\geqslant 2)$ is the initial sample size. The stopping rule in (1.9) is somewhat similar to the one defined in Chang and Martinsek (1992). The confidence ellipsoid $\boldsymbol{R}_{T_d}$ has length of the major axis equal to $2d$. Moreover, we have the following theorems.

Theorem 1.1. *Suppose $\boldsymbol{\theta} \in \Theta$ defined in* (1.2). *Then, for the stopping rule T_d defined in* (1.9) *the following hold*:

$$\text{(i) for each } d > 0,\ T_d < \infty \text{ a.s. and } T_d \to \infty \text{ a.s. as } d \to 0, \tag{1.10}$$

$$\text{(ii) } T_d/n_0(d) \to 1 \text{ a.s. as } d \to 0, \tag{1.11}$$

where $n_0(d)$ is as in (1.7), *and*

$$\text{(iii) } \lim_{d \to 0} P[\boldsymbol{\theta} \in \boldsymbol{R}_{T_d}] = 1 - \alpha \text{ (asymptotic consistency).} \tag{1.12}$$

Theorem 1.2 (Asymptotic efficiency). *Assume that $\boldsymbol{\theta} \in \Theta$ and $E|\varepsilon_1|^{2p+\gamma} < \infty$ for $p > 2$ and some $\gamma > 0$. Then, for T_d and $n_0(d)$ defined in* (1.9) *and* (1.7), *respectively, the following hold*:

$$\text{(i) } \{T_d/n_0(d); 0 < d < 1\} \text{ uniformly integrable} \tag{1.13}$$

and

$$\text{(ii) } \lim_{d \to 0} E(T_d/n_0(d)) = 1. \tag{1.14}$$

The stopping rule defined in (1.9) is very different from the one defined in Lee and Sriram (1999) for the sequential point estimation of $\boldsymbol{\theta}$. The difference is essentially due to the presence of r.v. $\lambda_{\min}(\Gamma_n)$ in (1.9) in place of n that appears in Lee and Sriram's rule. This necessitates obtaining rate of convergence of $n^{-1}\lambda_{\min}(\Gamma_n)$ to its limit in probability. This rate of convergence result is of independent interest and is given in a Lemma in Section 2. This result is crucial to prove (1.13) and (1.14).

The theorems stated above are proved in Section 2. In Section 3, we construct fixed proportional accuracy confidence ellipsoids for $\boldsymbol{\theta}$ and fixed width confidence intervals for a linear combination of $\boldsymbol{\theta}$. In Section 4, the results stated above are extended to a multiple-threshold AR(1) model. We end this section with some notations which will be used throughout the rest of the paper. Let

$$\hat{\lambda}_n = \lambda_{\min}(\Gamma_n) = \min\left(\sum_{i=1}^{n} X_{i-1}^{+2}, \sum_{i=1}^{n} X_{i-1}^{-2}\right) \quad \text{and} \quad \lambda = \min(EX_0^{+2}, EX_0^{-2}). \tag{1.15}$$

2. Proofs

Proof of Theorem 1.1. Since $n^{-1}\sum_{i=1}^{n} X_{i-1}^{\pm 2} \to EX_0^{\pm 2}$ a.s. as $n \to \infty$ it follows from (1.15) that

$$n^{-1}\hat{\lambda}_n \to \lambda \text{ a.s. as } n \to \infty. \tag{2.1}$$

This, the result $\hat{\sigma}_n^2 \to \sigma^2$ a.s. and routine arguments yield (1.10) and (1.11).

Assertion (1.12) follows from (1.5), (1.11) and the Anscombe's theorem (see, for instance, Woodroofe, 1982, Theorem 1.4) because the sequence

$$\{\sigma^{-2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})'\Gamma_n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}); n \geqslant 1\} \text{ is } \textit{uniformly continuous in probability (u.c.i.p.)}. \tag{2.2}$$

The result in (2.2) follows from Lemma 3.1, display (2.2) of Lee and Sriram (1999). □

Lemma. *Assume that* $\boldsymbol{\theta} \in \Theta$ *defined in* (1.2). *If* $E|\varepsilon_1|^{2p+\gamma} < \infty$ *for* $p > 2$ *and some* $\gamma > 0$, *then for any* $\delta > 0$

$$\text{(i) } P\left\{\left|n^{-1}\sum_{i=1}^{n} X_{i-1}^{\pm 2} - EX_0^{\pm 2}\right| > \delta\right\} = \mathrm{O}(n^{-p/2}) \tag{2.3}$$

and hence

$$\text{(ii) } P\{|n^{-1}\hat{\lambda}_n - \lambda| > \delta\} = \mathrm{O}(n^{-p/2}) \tag{2.4}$$

as $n \to \infty$, *where* $\hat{\lambda}_n$ *and* λ *are as defined in* (1.15).

Proof. It is shown below that the result in (2.3) follows from a result on moment bounds for stationary, strong mixing sequences which is due to Yokoyama (1980). To this end, first observe that if $\boldsymbol{\theta} \in \Theta$, then $\{X_i; i \geqslant 0\}$ is geometrically ergodic (see Chan et al., 1985, Theorem 2.3). From this, Theorem 2.1 and Remark 2.2 of Nummelin and Tuominen (1982), and results in Doukhan (1994, p. 88, display (1′)) it follows that $\{X_i; i \geqslant 0\}$ is (geometrically) β-mixing with mixing coefficient $\beta_n = \mathrm{O}(\rho^n)$ for some $0 < \rho < 1$. Since β-mixing implies strong mixing (or, equivalently, α-mixing, see Doukhan, 1994, pp. 3, 4 and 20) we have that the α-mixing coefficient $\alpha_n = \mathrm{O}(\rho^n)$. This implies that condition (3.1) of Theorem 1 in Yokoyama (1980) is satisfied. Now, let $S_n = \sum_{i=1}^{n}(X_{i-1}^{+2} - EX^{+2})$. By the moment assumption (also see (1.2)) and an application of Theorem 1 of Yokoyama (1980) we have that there exists a constant $M > 0$ such that

$$E|S_n|^p \leqslant Mn^{p/2}, \quad n \geqslant 1. \tag{2.5}$$

For the sequence $\{X_i^{+2}\}$, the result in (2.3) now follows from the Markov inequality and (2.5). Similarly, the result in (2.3) holds for the sequence $\{X_i^{-2}\}$.

As for (2.4), the algebraic identity $\min(a,b) = \{a + b - |a - b|\}/2$ and the inequality $||x| - |y|| \leqslant |x - y|$ imply that

$$|n^{-1}\hat{\lambda}_n - \lambda| \leqslant \left|n^{-1}\sum_{i=1}^{n} X_{i-1}^{+2} - EX^{+2}\right| + \left|n^{-1}\sum_{i=1}^{n} X_{i-1}^{-2} - EX^{-2}\right|. \tag{2.6}$$

The result in (2.4) now follows from (2.6) and (2.3). Hence the lemma. □

Proof of Theorem 1.2. Define another stopping rule $\tilde{T}_d$ by

$$\tilde{T}_d = \inf\left\{n \geqslant m\colon \hat{\lambda}_n \geqslant d^{-2}a^2\left(n^{-1}\sum_{i=1}^{n}\varepsilon_i^2\right)\right\}, \tag{2.7}$$

where $\hat{\lambda}_n$ is as in (1.15). Since $\hat{\theta}_{1,n}$ and $\hat{\theta}_{2,n}$ in (1.3) and (1.4) are least-squares estimators of θ_1 and θ_2, respectively, we have that $\hat{\sigma}_n^2 \leqslant n^{-1}\sum_{i=1}^n \varepsilon_i^2$, where $\hat{\sigma}_n^2$ is as defined in Section 1. From this and (1.9) it follows that

$$T_d \leqslant \tilde{T}_d. \tag{2.8}$$

Now, let $K_d = [d^{-2}a^2(\sigma^2/\lambda)(1+\delta)]+1$ for some $\delta > 0$ and λ defined in (1.15). Then, for $k \geqslant K_d$ and some $\eta > 0$ it can be shown that

$$\begin{aligned} P[\tilde{T}_d > k] &\leqslant P\left\{\left|\left[\left(k^{-1}\sum_{i=1}^k \varepsilon_i^2\right)\Big/(k^{-1}\hat{\lambda}_k)\right] - (\sigma^2/\lambda)\right| \geqslant \eta\right\} \\ &= \mathrm{O}(k^{-p/2}), \end{aligned} \tag{2.9}$$

where the last step follows from (2.4), the result $P\{|k^{-1}\sum_{i=1}^k \varepsilon_i^2 - \sigma^2| \geqslant \delta\} = \mathrm{O}(k^{-p/2})$ which follows from Corollary 10.3.2 of Chow and Teicher (1978), and an application of Lemma 1 of Sriram (1987). Now, (2.9) implies that $\sum_{k\geqslant 1} P(\tilde{T}_d > k) < \infty$. From this and arguments as of Woodroofe (1982, display (4.9), p. 47) it follows that

$$\{d^2\tilde{T}_d; 0 < d < 1\} \text{ is uniformly integrable.} \tag{2.10}$$

Hence, the assertion in (1.13) follows from (2.8),(2.10) and the definition of $n_0(d)$ in (1.7). The assertion in (1.14) follows from (1.11) and (1.13). Hence the theorem. □

3. Related fixed size confidence regions

3.1. Fixed proportional accuracy confidence ellipsoids

Suppose θ_1 and θ_2 in (1.1) are nonzero but at least one of the parameter values is near the origin. Then, one may wish to take this into account and construct a smaller confidence ellipsoid for $\boldsymbol{\theta}$ which gives us an improvement in accuracy of estimates of small coordinates. One approach (see Chang and Martinsek, 1992; Martinsek, 1995, for instance) is to construct an ellipsoidal region such that the statistical distance between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}$ is less than a certain fraction of the true value of $\theta_{(1)} = \min(|\theta_1|, |\theta_2|)$. This yields the following ellipsoidal region:

$$\boldsymbol{E}_n = \{\boldsymbol{z}\colon (\boldsymbol{z}-\hat{\boldsymbol{\theta}}_n)'\Gamma_n(\boldsymbol{z}-\hat{\boldsymbol{\theta}}_n) \leqslant d^2\lambda_{\min}(\Gamma_n)\hat{\theta}_{(1),n}\} \tag{3.1}$$

for $d > 0$, where $\hat{\theta}_{(1),n} = \min(|\hat{\theta}_{1,n}|, |\hat{\theta}_{2,n}|)$ with $\hat{\theta}_{1,n}$ and $\hat{\theta}_{2,n}$ as defined in (1.3) and (1.4), respectively, and $\lambda_{\min}(\Gamma_n)$ is as in (1.6). $\boldsymbol{E}_n$ defines an ellipsoid with length of the major axis equal to $2d\sqrt{\hat{\theta}_{(1),n}}$.

Now, for λ defined in (1.15), $\theta_{(1)} = \min(|\theta_1|, |\theta_2|)$ and a^2 as in (1.7), define an (unknown) sample size

$$t_0(d) = \text{smallest integer} \geqslant \sigma^2 a^2/[d^2\theta_{(1)}\lambda]. \tag{3.2}$$

Once again, as in (1.8), for any $\alpha \in (0,1)$ and the sample size determined by (3.2), if $\boldsymbol{\theta} \in \Theta$ defined in (1.2) then we have from (1.5) that

$$\lim_{d\to 0} P(\boldsymbol{\theta} \in \boldsymbol{E}_{t_0(d)}) = 1-\alpha. \tag{3.3}$$

However, the sample size $t_0(d)$ cannot be used in practice since it depends on the unknown parameters. Therefore, as in (1.9), define a stopping rule

$$N_d = \inf\{n \geqslant m:\ \hat{\lambda}_n \hat{\theta}_{(1),n} \geqslant \hat{\sigma}_n^2 a^2/d^2\}, \tag{3.4}$$

where $\hat{\lambda}_n$ is as in (1.15), $\hat{\theta}_{(1),n}$ is as defined in (3.1), $\hat{\sigma}_n^2$ is as defined in Section 1, and m is an initial sample size. We then have the following theorem.

Theorem 3.1. *Suppose $\boldsymbol{\theta} \in \Theta$ defined in* (1.2). *Then, for the stopping rule N_d defined in* (3.4) *the following hold*:

$$\text{(i) for each } d > 0, \quad N_d < \infty \text{ a.s.} \quad \text{and} \quad N_d \to \infty \text{ a.s. as } d \to 0, \tag{3.5}$$

$$\text{(ii) } N_d/t_0(d) \to 1 \text{ a.s.} \quad \text{as } d \to 0, \tag{3.6}$$

where $t_0(d)$ is as in (3.2) *and*

$$\text{(iii)} \lim_{d \to 0} P[\boldsymbol{\theta} \in \boldsymbol{E}_{N_d}] = 1 - \alpha. \tag{3.7}$$

Furthermore, under the conditions of Theorem 1.2

$$\text{(iv) } \{N_d/t_0(d); 0 < d < 1\} \text{ is uniformly integrable,} \tag{3.8}$$

and

$$\text{(v) } \lim_{d \to 0} E[N_d/t_0(d)] = 1. \tag{3.9}$$

Proof of Theorem 3.1. The assertions in (3.5)–(3.7) can be proved using arguments similar to those in the Proof of Theorem 1.1. As for (3.8), define another stopping rule $\tilde{N}_d$ by

$$\tilde{N}_d = \inf\left\{n \geqslant m : \hat{\lambda}_n \hat{\theta}_{(1),n} \geqslant d^{-2} a^2 \left(n^{-1} \sum_{i=1}^{n} \varepsilon_i^2\right)\right\}. \tag{3.10}$$

Once again, as in (2.8) and (2.10), $N_d \leqslant \tilde{N}_d$ and it suffices to establish the result in (3.8) for $\tilde{N}_d$ in place of N_d.

To this end, first we obtain the rate of convergence (in probability) of $\hat{\theta}_{1,n}$ and $\hat{\theta}_{2,n}$ (see (1.3) and (1.4)) to their limits θ_1 and θ_2, respectively. Write

$$\hat{\theta}_{1,n} - \theta_1 = n^{-1} \sum_{i=1}^{n} X_{i-1}^{+} X_i \Big/ \left(n^{-1} \sum_{i=1}^{n} X_{i-1}^{+2}\right) - (\theta_1 E X_0^{+2} / E X_0^{+2}). \tag{3.11}$$

Now,

$$n^{-1} \sum_{i=1}^{n} X_{i-1}^{+} X_i - \theta_1 E X_0^{+2} = n^{-1} \sum_{i=1}^{n} X_{i-1}^{+} \varepsilon_i + \theta_1 \left[n^{-1} \sum_{i=1}^{n} (X_{i-1}^{+2} - E X_0^{+2})\right]. \tag{3.12}$$

Since $E|\varepsilon_1|^{2p} < \infty$ for $p > 2$, by a result of Sriram (1988, Lemma 1, p. 58) and the Markov inequality we have that

$$P\left\{\left|n^{-1} \sum_{i=1}^{n} X_{i-1}^{+} \varepsilon_i\right| > \delta/2\right\} = \mathrm{O}(n^{-p/2}) \quad \text{as } n \to \infty \tag{3.13}$$

for some $\delta > 0$. Therefore, from (3.12) and (2.3) we have that

$$P\left\{\left|n^{-1}\sum_{i=1}^{n} X_{i-1}^{+}X_i - \theta_1 EX^{+2}\right| > \delta\right\} = \mathrm{O}(n^{-p/2}) \quad \text{as } n \to \infty. \tag{3.14}$$

Hence, the result

$$P\{|\hat{\theta}_{1,n} - \theta_1| > \delta\} = \mathrm{O}(n^{-p/2}) \quad \text{as } n \to \infty \tag{3.15}$$

follows from (3.11), (3.14), (2.6) and Lemma 1 of Sriram (1987). Similarly, it can be shown that

$$P\{|\hat{\theta}_{2,n} - \theta_2| > \delta\} = \mathrm{O}(n^{-p/2}) \quad \text{as } n \to \infty. \tag{3.16}$$

Now, in order to establish the uniform integrability of $\{d^2\tilde{N}_d; 0 < d < 1\}$, let $\tilde{K}_d = [d^{-2}a^2\{\sigma^2/(\lambda\theta_{(1)})\}(1+\delta)] + 1$ for some $\delta > 0$, use (3.15), (3.16) and (2.4), and argue as in (2.9) to (2.10). The assertion in (3.8) now follows. The assertion in (3.9) follows from (3.6) and (3.8). Hence the theorem. □

3.2. *Confidence interval for a linear combination of* $\boldsymbol{\theta}$

In addition to constructing a fixed size confidence region for the vector $\boldsymbol{\theta}$, often it is also of interest to construct a fixed width confidence interval for a particular linear combination $\boldsymbol{c}'\boldsymbol{\theta}$ for some known $\boldsymbol{c} = (c_1, c_2)' \neq 0$. In fact, in this context, it would be of special interest to construct a fixed width confidence interval for $\theta_1 - \theta_2$ since it would help differentiate between an AR(1) model (the case $\theta_1 = \theta_2$) and the threshold model in (1.1).

It follows from Theorem 3.2 of Petruccelli and Woolford (1984) that if $\boldsymbol{\theta} \in \Theta$ (see (1.2)) then for $\hat{\boldsymbol{\theta}}_n$ defined in (1.5)

$$\sqrt{n}(\boldsymbol{c}'\hat{\boldsymbol{\theta}}_n - \boldsymbol{c}'\boldsymbol{\theta}) \xrightarrow{D} N(0, \phi^2) \quad \text{as } n \to \infty, \tag{3.17}$$

where $\phi^2 = \sigma^2[(c_1^2/EX^{+2}) + (c_2^2/EX^{-2})]$. If ϕ^2 were known, then for $\alpha \in (0,1)$ and the sample size determined by

$$k_0(d) = \text{smallest integer} \geqslant z_{\alpha/2}^2\phi^2/d^2, \tag{3.18}$$

we have from (3.17) that

$$\lim_{d\to 0} P(\boldsymbol{c}'\boldsymbol{\theta} \in [\boldsymbol{c}'\hat{\boldsymbol{\theta}}_{k_0(d)} - d, \boldsymbol{c}'\hat{\boldsymbol{\theta}}_{k_0(d)} + d]) = 1 - \alpha,$$

where $z_{\alpha/2}$ satisfies $\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1 - \alpha$. However, since ϕ^2 is unknown, the sample size $k_0(d)$ cannot be used. As before, Eq. (3.18) suggests the stopping rule

$$\tau_d = \inf\{n \geqslant m: n \geqslant d^{-2}z_{\alpha/2}^2\hat{\phi}_n^2\}, \tag{3.19}$$

where m is an initial sample size and

$$\hat{\phi}_n^2 = \hat{\sigma}_n^2\left[c_1^2\left(n\Big/\sum_{i=1}^{n} X_{i-1}^{+2}\right) + c_2^2\left(n\Big/\sum_{i=1}^{n} X_{i-1}^{-2}\right)\right]. \tag{3.20}$$

We then have the following theorem.

Theorem 3.2. *Suppose* $\boldsymbol{\theta} \in \Theta$ *defined in* (1.2). *Then the following hold for the stopping rule* τ_d *and* $k_0(d)$ *defined in* (3.19) *and* (3.18), *respectively*:

$$\text{(i) for each } d > 0, \quad \tau_d < \infty \text{ a.s.} \quad \text{and} \quad \tau_d \to \infty \text{ a.s. as } d \to 0, \tag{3.21}$$

$$\text{(ii) } \tau_d / k_0(d) \to 1 \text{ a.s.} \quad \text{as } d \to 0 \tag{3.22}$$

and

$$\text{(iii) } \lim_{d \to 0} P\{\boldsymbol{c}'\boldsymbol{\theta} \in [\boldsymbol{c}'\hat{\boldsymbol{\theta}}_{\tau_d} - d, \boldsymbol{c}'\hat{\boldsymbol{\theta}}_{\tau_d} + d]\} = 1 - \alpha. \tag{3.23}$$

Furthermore, under the conditions of Theorem 1.2

$$\text{(iv) } \{\tau_d / k_0(d); 0 < d < 1\} \text{ is uniformly integrable} \tag{3.24}$$

and

$$\text{(v) } \lim_{d \to 0} E\tau_d / k_0(d) = 1. \tag{3.25}$$

Proof of Theorem 3.2. The assertions in (3.21)–(3.23) can be proved using exactly the same arguments as in the Proof of Theorem 1.1. Assertion (3.24) can be proved using similar arguments to those in (2.7)–(2.10) and the lemma in Section 2. Assertion (3.25) also follows similarly. Hence the theorem. □

4. Extension to a multiple-threshold AR(1) model

The purpose of this section is to construct fixed size confidence regions for parameters in a multiple-threshold AR(1) model. Consider a more general TAR model defined in Tong and Lim (1980); also see Chan et al. (1985). More specifically, for any integer l, let $-\infty = r_0 < r_1 < \cdots < r_l = \infty$ and define

$$X_i = \theta_k X_{i-1} + \varepsilon_i(k) \quad \text{if } X_{i-1} \in (r_{k-1}, r_k] \tag{4.1}$$

for $1 \leqslant k \leqslant l$. Equivalently, (4.1) may be written as

$$X_i = \sum_{k=1}^{l} [\theta_k X_{i-1} + \varepsilon_i(k)] I(X_{i-1} \in (r_{k-1}, r_k]), \tag{4.2}$$

where $I(A)$ is the indicator function of the set A. In (4.1) and (4.2), the thresholds are assumed to be known, $\{\theta_k, 1 \leqslant k \leqslant l\}$ are unknown real parameters which are not necessarily equal and we assume that for each k, $1 \leqslant k \leqslant l$, $\{\varepsilon_i(k)\}$ is a sequence of i.i.d. random variables with $E\varepsilon_1(k) = 0 < E\varepsilon_1^2(k) = \sigma_k^2 < \infty$. In addition, assume that $\{\varepsilon_i(k)\}$ and $\{\varepsilon_i(j)\}$ are independent sequences for $j \neq k$ and σ_k^2, $1 \leqslant k \leqslant l$, are unknown parameters which are not necessarily equal. Incidentally, our notations in (4.1) are different from those in Chan et al. (1985) and, further, we assume that the intercept parameter in (4.1) is zero while Chan et al. (1985) assume it to be nonzero.

As mentioned in the Introduction, Chan et al. (1985) have shown that the process $\{X_i; i \geqslant 0\}$ defined in (4.1) is ergodic if and only if one of five conditions on

$\{\theta_k, 1 \leqslant k \leqslant l\}$ holds; see Theorem 2.1 of Chan et al. (1985) for details. One of the five sufficient conditions for ergodicity of $\{X_i; i \geqslant 0\}$ is

$$\boldsymbol{\theta} \in \Theta = \{(\theta_1, \ldots, \theta_l)' \colon \theta_1 < 1, \theta_l < 1 \text{ and } \theta_1\theta_l < 1\}. \tag{4.3}$$

As in the Introduction, let X_0 have as its distribution $\pi(\cdot)$, the invariant probability distribution of $\{X_i\}$ defined in (4.1). If $\boldsymbol{\theta} \in \Theta$ defined in (4.3) then by Theorem 2.3 of Chan et al. (1985), the invariant probability distribution of $\{X_i\}$ has finite pth moment provided $E|\varepsilon_1(k)|^p < \infty$, $1 \leqslant k \leqslant l$, for some $p \geqslant 1$. In what follows, let $X_i(k) = X_i I(X_i \in (r_{k-1}, r_k])$ for $1 \leqslant k \leqslant l$ and $i \geqslant 0$, and $X(k) = XI(X \in (r_{k-1}, r_k])$ for X defined above.

Suppose we estimate the parameters $\theta_1, \ldots, \theta_l$ in (4.1) by their least-squares estimators

$$\hat{\theta}_{k,n} = \sum_{i=1}^{n} X_i X_{i-1}(k) \Big/ \sum_{i=1}^{n} X_{i-1}^2(k) \tag{4.4}$$

for $1 \leqslant k \leqslant l$. Then, for each $1 \leqslant k \leqslant l$, the corresponding estimator of σ_k^2 is $\hat{\sigma}_{k,n}^2 = n^{-1} \sum_{i=1}^{n} (X_i - \hat{\theta}_{k,n} X_{i-1}(k))^2$. By Theorem 3.1 of Chan et al. (1985), if $\boldsymbol{\theta} \in \Theta$ defined in (4.3), then, for each $1 \leqslant k \leqslant l$, $\hat{\theta}_{k,n}$ and $\hat{\sigma}_{k,n}^2$ are strongly consistent for θ_k and σ_k^2, respectively. Furthermore, by arguments similar to those of Theorem 3.2 of Chan et al. (1985), we have for $\hat{\boldsymbol{\theta}}_n = (\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{l,n})'$ that $D_n^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{D} N_l(\mathbf{0}, I_l)$ as $n \to \infty$, where $D_n = \operatorname{diag}(\sum_{i=1}^{n} X_{i-1}^2(1)/\hat{\sigma}_{1,n}^2, \ldots, \sum_{i=1}^{2} X_{i-1}^2(l)/\hat{\sigma}_{l,n}^2)$ and I_l is the $l \times l$ identity matrix. This in turn implies that if $\boldsymbol{\theta} \in \Theta$, then

$$(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})' D_n (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{D} \chi_l^2 \quad \text{as } n \to \infty, \tag{4.5}$$

where χ_l^2 is a χ^2 r.v. with l degrees of freedom. Now, for any $d > 0$, let

$$\tilde{\boldsymbol{R}}_n = \{\boldsymbol{z} \colon (\boldsymbol{z} - \hat{\boldsymbol{\theta}}_n)' D_n (\boldsymbol{z} - \hat{\boldsymbol{\theta}}_n) \leqslant d^2 \lambda_{\min}(D_n)\}, \tag{4.6}$$

where $\lambda_{\min}(D_n) = \min(\sum_{i=1}^{n} X_{i-1}^2(1)/\hat{\sigma}_{1,n}^2, \ldots, \sum_{i=1}^{n} X_{i-1}^2(l)/\hat{\sigma}_{l,n}^2)$ is the smallest eigenvalue of D_n defined above. Then, $\tilde{\boldsymbol{R}}_n$ defines an ellipsoid with length of the major axis equal to $2d$. Moreover, for any $\alpha \in (0,1)$ and sample size determined by

$$m_0(d) = \text{ smallest integer} \geqslant a^2/[d^2 \lambda^*] \tag{4.7}$$

with $\lambda^* = \min(EX^2(1)/\sigma_1^2, \ldots, EX^2(l)/\sigma_l^2)$, we have from (4.5) that if $\boldsymbol{\theta} \in \Theta$ defined in (4.3), then

$$\lim_{d \to 0} P(\boldsymbol{\theta} \in \tilde{\boldsymbol{R}}_{m_0(d)}) = 1 - \alpha, \tag{4.8}$$

where a^2 in (4.7) satisfies $P[\chi_l^2 \leqslant a^2] = 1 - \alpha$. As before, $m_0(d)$ cannot be used in practice. Therefore, we define the following stopping rule:

$$S_d = \inf\{n \geqslant m \colon \lambda_{\min}(D_n) \geqslant a^2/d^2\}, \tag{4.9}$$

where $\lambda_{\min}(D_n)$ is as defined in (4.6). Then the confidence ellipsoid $\tilde{\boldsymbol{R}}_{S_d}$ has length of the major axis equal to $2d$ and we have the following theorem.

Theorem 4.1. *Suppose* $\boldsymbol{\theta} \in \Theta$ *defined in* (4.3). *Then, for the stopping rule* S_d *defined in* (4.9) *the following hold*:

$$\text{(i) for each } d > 0, \quad S_d < \infty \text{ a.s. and } \quad S_d \to \infty \text{ a.s. as } d \to 0, \tag{4.10}$$

$$\text{(ii) } S_d/m_0(d) \to 1 \text{ a.s. as } d \to 0, \tag{4.11}$$

where $m_0(d)$ *is as defined in* (4.7), *and*

$$\text{(iii) } \lim_{d \to 0} P[\boldsymbol{\theta} \in \tilde{\boldsymbol{R}}_{S_d}] = 1 - \alpha. \tag{4.12}$$

Furthermore, if for each $1 \leqslant k \leqslant l$, $E|\varepsilon_1(k)|^{2p+\gamma} < \infty$ *for* $p > 2$ *and some* $\gamma > 0$, *then*

$$\text{(iv) } \{S_d/m_0(d); 0 < d < 1\} \text{ is uniformly integrable} \tag{4.13}$$

and

$$\text{(v) } \lim_{d \to 0} E[S_d/m_0(d)] = 1. \tag{4.14}$$

Proof of Theorem 4.1. The assertions in (4.10)–(4.12) can be proved using arguments similar to those in the proof of Theorem 1.1. As for (4.13), define another stopping rule $\tilde{S}_d = \inf\{n \geqslant m: \tilde{\lambda}_n \geqslant d^{-2}a^2(\sum_{k=1}^{l} n^{-1} \sum_{i=1}^{n} \varepsilon_i^2(k))\}$, where $\tilde{\lambda}_n = \min(\sum_{i=1}^{n} X_{i-1}^2(1), \ldots, \sum_{i=1}^{n} X_{i-1}^2(l))$. Since $\hat{\theta}_{k,n}$ is the least-squares estimator of θ_k, $1 \leqslant k \leqslant l$, we have that $\hat{\sigma}_{k,n}^2 \leqslant n^{-1} \sum_{i=1}^{n} \varepsilon_i^2(k)$. From this and the fact that $\hat{\sigma}_{k,n}^2 \leqslant \sum_{k=1}^{l} \hat{\sigma}_{k,n}^2$ we have that $S_d \leqslant \tilde{S}_d$. Now, proceed as in the proof of Theorem 1.2 using a result analogous to (2.3) for $n^{-1} \sum_{i=1}^{n} X_{i-1}^2(k)$, $1 \leqslant k \leqslant l$, and (2.4) to prove (4.13). The assertion in (4.14) follows from (4.11) and (4.13). □

Remark. As in Section 3, for the multiple-TAR(1) model in (4.1), it is also possible to construct fixed proportional accuracy confidence ellipsoids and fixed width confidence intervals for linear combinations of $\boldsymbol{\theta}$ and establish their asymptotic properties. Such constructions, however, are very similar to those in Section 3. Hence, we do not explicitly state the associated results.

Acknowledgements

The author would like to thank the first referee for pointing out an error in the proof of lemma in an earlier version. The author would also like to thank the second referee for the suggestions which led to improvement in the presentation of the contents.

References

Albert, A., 1966. Fixed size confidence ellipsoids for linear regression parameters. Ann. Math. Statist. 37, 1602–1630.

Chan, K.S., Petrucelli, J.D., Tong, H., Woolford, S.W., 1985. A multiple-threshold AR(1) model. J. Appl. Probab. 22, 267–279.

Chang, Y.-C.I., Martinsek, A.T., 1992. Fixed size confidence regions for parameters of a logistic regression model. Ann. Statist. 20, 1953–1969.

Chow, Y.S., Robbins, H., 1965. On the asymptotic theory of fixed-width confidence intervals for the mean. Ann. Math. Statist. 36, 457–462.

Chow, Y.S., Teicher, H., 1978. Probability Theory: Independence, Interchangeability, Martingales. Springer, New York.

Doukhan, P., 1994. Mixing Properties and Examples. Lecture notes in Statistics. Springer, New York.

Fakhre-Zakeri, I., Lee, S., 1992. Sequential estimation of the mean of a linear process. Sequential Anal. 11, 181–197.

Fakhre-Zakeri, I., Lee, S., 1993. Sequential estimation of the mean vector of a multivariate linear process. J. Multivariate Anal. 47, 196–209.

Finster, M., 1985. Estimation in the general linear model when the accuracy is specified before data collection. Ann. Statist. 13, 663–675.

Gleser, L.J., 1965. On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. Ann. Math. Statist. 36, 463–467.

Grambsch, P., 1983. Sequential sampling based on the observed Fisher information to guarantee the accuracy of the maximum likelihood estimator. Ann. Statist. 11, 68–77.

Grambsch, P., 1989. Sequential maximum likelihood estimation with applications to logistic regression in case-control studies. J. Statist. Plann. Inference 22, 355–369.

Greenwood, P.E., Shiryaev, A.N., 1992. Asymptotic minimaxity of a sequential estimator for a first order autoregressive model. Stochastics Stochastic Rep. 38, 49–65.

Lai, T.L., Siegmund, D., 1983. Fixed accuracy estimation of an autoregressive parameter. Ann. Statist. 11, 478–485.

Lee, S., 1994. Sequential estimation for the parameters of a stationary autoregressive model. Sequential Anal. 13, 301–317.

Lee, S., Sriram, T.N., 1999. Sequential point estimation of parameters in a threshold AR(1) model. Stochastic Process. Appl. 84, 343–355.

Martinsek, A.T., 1995. Estimating a slope parameter in regression with prescribed proportional accuracy. Statist. Decisions 13, 363–377.

Nummelin, E., Tuominen, P., 1982. Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. Stochastic Process. Appl. 12, 187–202.

Petruccelli, J.D., Woolford, S.W., 1984. A threshold AR(1) model. J. Appl. Probab. 21, 270–286.

Sriram, T.N., 1987. Sequential estimation of the mean of a first order stationary autoregressive process. Ann. Statist. 15, 1079–1090.

Sriram, T.N., 1988. Sequential estimation of the autoregressive parameter in a first order autoregressive process. Sequential Anal. 7, 53–74.

Srivastava, M.S., 1967. On fixed-width confidence bounds for regression parameters and mean vector. J. Roy. Statist. Soc. Ser. B 29, 132–140.

Srivastava, M.S., 1971. On fixed-width confidence bounds for regression parameters. Ann. Math. Statist. 42, 1403–1411.

Tong, H., 1978. On a threshold model. In: Chen, C.H. (Ed.), Pattern Recognition and Signal Processing. Sijthoof & Noordhoff, Alphen aan den Rijn, The Netherlands.

Tong, H., Lim, K.S., 1980. Threshold autoregression, limit cycles and cyclical data. J. Roy. Statist. Soc. Ser. B 42, 245–292.

Woodroofe, M., 1982. Nonlinear Renewal Theory in Sequential Analysis. SIAM, Philadelphia, PA.

Yokoyama, R., 1980. Moment bounds for stationary mixing sequences. Z. Wahrsch. Verw. Gebiete. 52, 45–57.

Yu, K.F., 1989. On fixed-width confidence intervals associated with maximum likelihood estimation. J. Theoret. Probab. 2, 193–199.