Fixed Size Confidence Regions for Parameters of Stationary Processes Based on a Minimum Contrast Estimator

Takayuki Shiohama
Institute of Economic Research, Hitotsubashi University,
Naka 2-1, Kunitachi, Tokyo 186-0003, Japan

Abstract

For parameters of stationary processes with zero mean and spectral density, sequential procedures are proposed for constructing fixed size confidence ellipsoidal regions for unknown parameters using a minimum contrast estimator. The confidence ellipsoids are shown to be asymptotically consistent and the associated stopping rules are shown to be asymptotically efficient as the size of the region becomes small when the assumed parametric model is correct. Monte Carlo simulations are given to investigate the performance of our proposed sequential procedures.

1 Introduction

It is well documented in literature that sequential sampling methods provide a useful way of constructing confidence intervals or regions for parameters with a fixed size and a prescribed coverage probability. Chow and Robbins (1965) proposed a sequential sampling rule for constructing a fixed-width confidence interval for an unknown mean with a prescribed probability and developed its asymptotic theory. This sampling rule is referred to as the "Chow-Robbins procedure." For details, refer to Chapter 8 of Ghosh, Mukhopadyay and Sen (1997). Their ideas have been used to develop sequential fixed size confidence regions for parameters associated with dependent and independent observations. For independent, identically distributed observations, we refer the reader to Srivastava (1967), Khan (1969), Yu (1989), Woodroofe (1982), and Chang and Martinsek (1992).

With regard to time series, Sriram (1987) developed a point and interval estimation for the mean of a first order autoregressive (AR(1)) model. Fakhre-Zakeri and Lee (1992, 1993) later considered a sequential point and fixed-width confidence interval estimation for the mean of a scalar- or vector-valued linear process. Sequential procedures dealing with both point estimation and fixed accuracy confidence sets of unknown autoregressive coefficients have been considered by Lee (1994). Sriram (2001) proposed a stopping rule to construct a sequential fixed-size confidence ellipsoid for the parameters in threshold autoregressive (TAR) models. Shiohama and Taniguchi (2001) considered the sequential estimation problems for functional of the spectral density of a Gaussian stationary process. Recently Shiohama and Taniguchi (2004) considered the sequential point estimation problems arising in time series regression models.

In this article, we assume that the observations are stationary processes with parametric spectral density $f_\theta(\lambda)$, where θ is an unknown parameter. In order to estimate θ, we use a minimum contrast estimator, $\hat{\theta}_n^{(MCE)}$, which minimizes the criterion $D(f_\theta, \hat{f}_n) = \int_{-\pi}^{\pi} K\{f_\theta(\lambda)/\hat{f}_n(\lambda)\}d\lambda$ with respect to θ, where $\hat{f}_n(\lambda)$ is a non-parametric spectral estimator of $f_\theta(\lambda)$, and $K(\cdot)$ is an appropriate function. It was shown that under appropriate conditions, the main order term of $\sqrt{n}(\hat{\theta}_n^{(MCE)} - \theta)$ can be written as $F = \sqrt{n}\int \Psi(\lambda)\{\hat{f}_n(\lambda) - f(\lambda)\}d\lambda$, where $\Psi(\lambda)$ is an integrable function. Although the nonparametric spectral estimator deviates from $f(\lambda)$ by a probability order that is greater than $n^{-1/2}$, Taniguchi (1987) showed that the integrable functionals obey the $\sqrt{n}$-consistent asymptotics, and that $\hat{\theta}_n^{(MCE)}$ is asymptotically efficient if $f = f_\theta$. Therefore, it can be seen that the integral functional F is the key quantity. The sequential estimation problem of this integral functional has been studied by Shiohama and Taniguchi (2001). The minimum contrast estimator has the following desirable property. For various spectra $f_\theta(\lambda)$, by appropriately selecting the function $K(\cdot)$ in $D(f_\theta, \hat{f}_n)$ we can obtain the non-iterative efficient estimators of θ in explicit forms, whereas with the exception of autoregressive models, the (quasi) maximum likelihood estimations procedure requires iterative methods. For details, refer to Taniguchi (1987) and Taniguchi and Kakizawa (2000).

In Section 2, we introduce the minimum contrast estimator (MCE) and construct sequential fixed size confidence regions for θ based on it. We then state the main theorem which establishes the asymptotic consistency and efficiency of our sequential procedure. Proofs are provided in Section 3. Section 4 comprises a brief discussion on estimation with fixed proportional accuracy and estimation of a particular linear combination of the components of θ. Section 5 contains several Monte Carlo simulations that demonstrate the performances of our sequential procedure based on the MCE. In this paper, we denote the set of all integers by $\mathbb{Z}$.

2 Stopping Rule and Main Theorem

Let $\{X_t, t \geq 0\}$ be a scalar-valued linear process of the form

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}, \tag{2.1}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with $E\{\varepsilon_t\} = 0$, $E\{\varepsilon_t^2\} = \sigma^2$ and $E\{\varepsilon_t^{2p}\} < \infty$, for $p > 2$. Then the process $\{X_t; t \in \mathbb{Z}\}$ is a second-order stationary process with spectral density $f(\lambda)$. Let $\boldsymbol{F}$ be the space of spectral densities defined by

$$\begin{aligned} \boldsymbol{F} \quad = \quad & \left\{ f; f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} a_j e^{-ij\lambda} \right|^2, \quad \text{there exist } C < \infty \text{ and } \delta > 0 \text{ such that} \right. \\ & \left. \sum_{j=0}^{\infty} (1 + j^2)|a_j| \leq C, \left| \sum_{j=0}^{\infty} a_j z^j \right| \leq \delta, \quad \text{for all } |z| \leq 1 \right\}. \end{aligned}$$

Denote $I_n(\lambda)$, the periodogram constructed from a realization $\{X_1, \ldots, X_n\}$, by

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t e^{it\lambda} \right|^2.$$

We estimate $f(\lambda)$ by the weighted averages of the periodogram $I_n(\lambda)$, with a spectral window $W_n(\lambda)$ as the weight, i.e.,

$$\hat{f}_n(\lambda) = \int_{-\pi}^{\pi} W_n(\lambda - \mu) I_n(\mu) d\mu. \tag{2.2}$$

The following conditions are imposed on $W_n(\cdot)$.

(A.1) (i) $W_n(\lambda)$ can be expressed as

$$W_n(\lambda) = \frac{1}{2\pi} \sum_{l=-M}^{M} w\left(\frac{l}{M}\right) e^{-il\lambda}.$$

(ii) $w(x)$ is a continuous, even function with $w(0) = 1$, and satisfies

$$\begin{cases} |w(x)| \leq 1, \\ \int_{-\infty}^{\infty} w(x)^2 dx < \infty, \lim_{x \to 0} \frac{1 - w(x)}{|x|^2} = \kappa_2 < \infty. \end{cases}$$

(iii) $M = M(n)$ satisfies

$$n^{1/4} M + M/n^{1/2} \to 0 \quad \text{as} \quad n \to \infty.$$

Concrete examples of $W_n(\cdot)$ satisfying (A.1) can be found in Hannan (1970), Brillinger (1981), and Robinson (1983). We then define the criterion that measures the nearness of f_θ to f as

$$D(f_\theta, f) = \int_{-\pi}^{\pi} K\{f_\theta(\lambda)/f(\lambda)\} d\lambda.$$

The following are examples of $K(\cdot)$:

$$\begin{aligned}
&\text{(i)} \quad K(x) = \log x + x^{-1}, \\
&\text{(ii)} \quad K(x) = -\log x + x, \\
&\text{(iii)} \quad K(x) = (\log x)^2, \\
&\text{(iv)} \quad K(x) = x \log x - x, \\
&\text{(v)} \quad (x^\alpha - 1)^2, \quad 0 < \alpha < \infty.
\end{aligned}$$

We impose the following assumptions on $K(\cdot)$ and $f_\theta(\lambda)$.

(A.2) (i) $K(x)$ is a three times continuously differentiable function in $(0, \infty)$ and has a unique minimum at $x = 1$.

(ii) The spectral model $f_\theta(\lambda)$ is three times continuously differentiable with respect to θ, and every component of the second derivative $\partial^2 f_\theta / \partial\theta\partial\theta'$ is continuous in λ.

In order to estimate the unknown θ, since $f(\lambda)$ is unknown, we estimate $f(\lambda)$ by a nonparametric estimator (2.2) satisfying (A.1). Therefore, a semiparametric estimator $\hat{\theta}_n^{(MCE)}$ of θ is defined as

$$\hat{\theta}_n^{(MCE)} = \underset{\theta \in \Theta}{\text{argmin}} D(f_\theta, \hat{f}_n(\lambda)). \tag{2.3}$$

Suppose that Assumptions (A.1) and (A.2) hold and $f = f_\theta$, then

$$\sqrt{n}(\hat{\theta}_n^{(MCE)} - \theta) \xrightarrow[\mathscr{L}]{} N(\mathbf{0}, F(\theta)^{-1}), \tag{2.4}$$

where

$$F(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta} \log f_\theta(\lambda) \frac{\partial}{\partial\theta'} \log f_\theta(\lambda) d\lambda, \tag{2.5}$$

which is referred to as the Fisher information matrix in time series analysis; refer to Taniguchi (1987) and Taniguchi and Kakizawa (2001). If we select an appropriate $K(\cdot)$, the minimum contrast estimator (2.3) provides explicit, non-iterative, and efficient estimators for various spectral parameterizations.

Suppose that the spectral density $f_\theta(\lambda)$ is parameterized as

$$f_\theta(\lambda) = S\{A_\theta(\lambda)\}, \tag{2.6}$$

where $A_\theta(\lambda) = \sum_j \theta_j \exp(ij\lambda)$ and $S(\cdot)$ is a continuously three times differentiable bijective function. In order to obtain non-iterative estimators, the following relation should be imposed;

$$K\left[\frac{S(A_\theta(\lambda))}{f(\lambda)}\right] = c_1(\lambda)A_\theta(\lambda)^2 + c_2(\lambda)A_\theta(\lambda) + c_3(\lambda) + c_4 \log S\{A_\theta(\lambda)\}, \tag{2.7}$$

where $c_i(\lambda), i = 1, 2, 3,$ are functions that are independent of θ, and c_4 is a constant that is independent of θ and λ. If we estimate an innovation-free parameter $\theta = (\theta_1, \ldots, \theta_q)'$, then from Theorem 6 of Taniguchi (1987), we observe that the the non-iterative estimator is given by

$$\hat{\theta}_n^{(MCE)} = \hat{R}^{-1}\hat{r}, \tag{2.8}$$

where

$$\hat{R} = [\hat{R}(j-l)] = \int_{-\pi}^{\pi} G_1(\hat{f}_n(\lambda)) \cos(j-l)\lambda d\lambda, \tag{2.9}$$

and

$$\hat{r} = [\hat{r}(l)] = \int_{-\pi}^{\pi} G_2(\hat{f}_n(\lambda)) \cos l\lambda d\lambda. \tag{2.10}$$

In this case, $G_i(\cdot),\ i = 1, 2,$ satisfies a uniform Lipschitz condition (of order 1) in $[-\pi, \pi]$.

For AR models with spectral density $f_\theta(\lambda) = \sigma^2/2\pi|\sum_{j=0}^p \theta_j e^{ij\lambda}|^{-2}$, where $\theta_0 = 1$ and $\sum_{j=0}^q \theta_j z^j \neq 0$ for $|z| \leq 1$, select $K_{AR}(x) = \log x + \frac{1}{x}$; therefore, the non-iterative estimator is obtained by selecting $G_1(x) = G_2(x) = x$. For MA models with spectral density $f_\theta(\lambda) = \sigma^2/2\pi|\sum_{j=0}^p \theta_j e^{ij\lambda}|^2$, where $\theta_0 = 1$ and $\sum_{j=0}^q \theta_j z^j \neq 0$ for $|z| \leq 1$, select $K_{MA}(x) = -\log x + x$; therefore, the non-iterative estimator is obtained by selecting $G_1(x) = G_2(x) = x^{-1}$. In the case where $f_\theta(\lambda) = \sigma^2 \exp\left[\sum_{j=0}^q \theta_j \cos(j\lambda)\right]$, $\theta_0 = 1$ (refer to Bloomfield (1973)), select $K_E(x) = (\log(x))^2$; therefore, the non-iterative estimator is obtained by selecting $G_1(x) = 1/2$ and $G_2(x) = \log x$.

Based on the asymptotic normality result for $\hat{\theta}_n^{(MCE)}$ in (2.4), it follows that

$$n(\hat{\theta}_n^{(MCE)} - \theta)' F(\hat{\theta}_n^{(MCE)})(\hat{\theta}_n^{(MCE)} - \theta) \xrightarrow[\mathscr{L}]{} \chi^2(q), \quad \text{as} \quad n \to \infty, \tag{2.11}$$

where

$$F(\hat{\theta}_n^{(MCE)}) = \frac{1}{4\pi}\left[\int_{-\pi}^{\pi}\frac{\partial}{\partial\theta}\log f_\theta(\lambda)\frac{\partial}{\partial\theta'}\log f_\theta(\lambda)d\lambda\right]_{\theta=\hat{\theta}_n^{(MCE)}}. \tag{2.12}$$

For any $d > 0$, let

$$R_n = \left\{\theta \in \mathbb{R}^q : (\theta - \hat{\theta}_n^{(MCE)})'F(\hat{\theta}_n^{(MCE)})(\theta - \hat{\theta}_n^{(MCE)}) \le d^2\lambda(F(\hat{\theta}_n^{(MCE)}))\right\}, \tag{2.13}$$

where $\lambda(F(\hat{\theta}_n^{(MCE)}))$ is the smallest eigenvalue of $F(\hat{\theta}_n^{(MCE)})$. Then, R_n defines an ellipsoid with a maximum axis equal to $2d$ ($d > 0$), and it is in this sense that the size of the ellipsoid is fixed. Moreover, for any $\alpha \in (0,1)$, $n_0(d)$ is determined by

$$n_0(d) = \text{smallest integer} \ge a^2/d^2\lambda(F(\theta)), \tag{2.14}$$

where a^2 satisfies $P[\chi^2(q) \le a^2] = 1-\alpha$, and $\lambda(F(\theta))$ is the smallest eigenvalue of the covariance matrix $F(\theta)$. From (2.13), for $\theta \in \Theta$, we have

$$\lim_{d\to 0} P(\theta \in R_{n_0(d)}) = 1-\alpha. \tag{2.15}$$

This result in (2.15) shows that for a small value of d, the sample size $n_0(d)$ yields an ellipsoidal confidence region of a fixed size and a prescribed coverage probability. However, the sample size $n_0(d)$ cannot be used in practice because it depends on unknown parameters. In order to overcome this, we define a stopping rule

$$T_d = \inf\left\{n \ge m, n \ge a^2/d^2\lambda(F(\hat{\theta}_n^{(MCE)}))\right\}, \tag{2.16}$$

where m is the initial sample size. The confidence ellipsoid R_{T_d} has the length of the major axis equal to $2d$. Moreover, we have the following theorems whose proofs are provided in Section 3.

Theorem 2.1 *Suppose that Assumptions (A.1) and (A.2) hold, and $\theta \in \Theta$. Then, for the stopping rule T_d defined in (2.16), the following holds.*

$$(i) \quad T_d/n_0(d) \to 1 \quad a.s. \quad as\ d \to 0, \tag{2.17}$$

where $d_0(d)$ is as in (2.14), and

$$(ii) \quad \sqrt{T_d}(\hat{\theta}_{T_d}^{(MCE)} - \theta) \underset{\mathscr{L}}{\longrightarrow} N(\mathbf{0}, F(\theta)^{-1}) \tag{2.18}$$

$$(iii) \quad \lim_{d\to 0} P[\theta \in R_{T_d}] = 1-\alpha \ \textit{(asymptotic consistency)}. \tag{2.19}$$

Theorem 2.2 *Suppose that Assumptions (A.1) and (A.2) hold, and $\theta \in \Theta$. Then, for the stopping rule T_d and $n_0(d)$ defined in (2.16) and (2.14), respectively, the following holds.*

$$(i) \quad \{T_d/n_0(d); 0 < d < 1\} \ \text{ is uniformly integrable} \tag{2.20}$$

and

$$(ii) \quad \lim_{d\to 0} E(T_d/n_0(d)) = 1 \ \text{ (asymptotic efficiency).} \tag{2.21}$$

The third part of Theorem 2.1 states that the coverage probability of the sequential fixed size confidence ellipsoid is asymptotically, as the size of the ellipsoid approaches zero, the desired value $1-\alpha$. Theorem 2.2 asserts that this is achieved with an expected sample size that is asymptotically equivalent to the nonrandom sample size that would have been used, had $\lambda(F(\theta))$ been known.

3 Proofs

In this section, we present proofs for Theorems 2.1 and 2.2. The proofs for these Theorems are based on the following lemmas. Let $\theta = (\theta_1, \dots, \theta_q)'$ and $\hat{\theta}_n^{(MCE)} = (\hat{\theta}_{n,1}^{(MCE)}, \dots, \hat{\theta}_{n,q}^{(MCE)})'$. $||\{\cdot\}||_p$ denotes the L_p-norm, i.e., $||\{\cdot\}||_p = [E|\{\cdot\}|^p]^{1/p}$.

Lemma 3.1 *Suppose that (A.1) and (A.2) hold. If $f = f_\theta$, where θ is the innovation free parameter, then*

$$\max_{1\le i\le q} \left\| \hat{\theta}_{n,i}^{(MCE)} - \theta_i \right\|_p = O(M \cdot n^{-1/2}). \tag{3.1}$$

Proof. To prove (3.1), it suffices to show that for any constant vector $\alpha = (\alpha_1, \dots, \alpha_q)'$

$$\left\| \alpha'(\hat{\theta}_n^{(MCE)} - \theta) \right\|_p = O(M \cdot n^{-1/2}). \tag{3.2}$$

From (2.8), note that

$$\begin{aligned} \alpha'(\hat{\theta}_n^{(MCE)} - \theta) &= \alpha'(\hat{R}^{-1}\hat{r} - R^{-1}r) \\ &= \alpha'(R^{-1}(\hat{r} - r) + (\hat{R}^{-1} - R^{-1})\hat{r}) \\ &= \alpha'(R^{-1}(\hat{r} - r) + \hat{R}^{-1}(\hat{R} - R)R^{-1}\hat{r}), \end{aligned} \tag{3.3}$$

where

$$R = [R(j-l)] = \int_{-\pi}^{\pi} G_1(f_\theta(\lambda)) \cos(j-l)\lambda d\lambda \tag{3.4}$$

and

$$r = [r(l)] = \int_{-\pi}^{\pi} G_2(f_\theta(\lambda)) \cos l\lambda d\lambda. \tag{3.5}$$

Using the Minkowski inequality, we observe

$$\begin{aligned} \left\| \alpha'(\hat{\theta}_n^{(MCE)} - \theta) \right\|_p &\leq \left\| \alpha' R^{-1}(\hat{r} - r) \right\|_p + \left\| \alpha' \hat{R}^{-1}(\hat{R} - R) R^{-1}) \hat{r} \right\|_p \\ &= L_1 + L_2 \qquad \text{(say)}. \end{aligned} \tag{3.6}$$

We first evaluate L_1 in (3.6). From the Minkowski inequality,

$$\begin{aligned} \|L_1\|_p &= \left\| \sum_{i,j=1}^{q} \alpha_i R_{ij}^{-1} (\hat{r}(j) - r(j)) \right\|_p \\ &\leq \sum_{i,j=1}^{q} |\alpha_i| |R_{ij}^{-1}| \, \|\hat{r}(j) - r(j)\|_p , \end{aligned} \tag{3.7}$$

where R_{ij}^{-1} is the (i,j)th element of R^{-1}. Note that $G_2(\cdot)$ satisfies the Lipschitz condition of order 1, (2.10), and (3.5); we observe that for some constants $K_1 > 0$ and $K_2 > 0$,

$$\begin{aligned} \|\hat{r}(j) - r(j)\|_p &= \left\| \int_{-\pi}^{\pi} (G_2(\hat{f}_n(\lambda) - G_2(f_\theta(\lambda))) \cos j\lambda d\lambda \right\|_p \\ &\leq \left\| \int_{-\pi}^{\pi} K_1 |\hat{f}_n(\lambda) - f_\theta(\lambda)| \cos j\lambda d\lambda \right\|_p \\ &\leq K_1 \left\| \sup_{|\lambda| \leq \pi} |\hat{f}_n(\lambda) - f_\theta(\lambda)| \int_{-\pi}^{\pi} \cos j\lambda d\lambda \right\|_p \\ &\leq K_2 \left\| \sup_{|\lambda| \leq \pi} |\hat{f}_n(\lambda) - f_\theta(\lambda)| \right\|_p . \end{aligned} \tag{3.8}$$

From the equation (4.7) by Shiohama and Taniguchi (2004), we have

$$\left\| \sup_{|\lambda| \leq \pi} \left| \hat{f}_n(\lambda) - f_\theta(\lambda) \right| \right\|_p = (M \cdot n^{-1/2}). \tag{3.9}$$

Hence $L_1 = O(M \cdot n^{-1/2})$. L_2 can be evaluated as follows:

$$\begin{aligned} \|L_2\|_p &= \left\| \sum_{i,j,k,l=1}^{q} \alpha_i \hat{R}_{ij}^{-1} (\hat{R}(j-k) - R(j-k)) R_{kl}^{-1} \hat{r}(l) \right\|_p \\ &\leq \sum_{i,j,k,l=1}^{q} |\alpha_i| |R_{kl}^{-1}| \left\| \hat{R}_{ij}^{-1} \hat{r}(l) (\hat{R}(j-k) - R(j-k)) \right\|_p \\ &\leq \sum_{i,j,k,l=1}^{q} |\alpha_i| |R_{kl}^{-1}| \left\| \hat{R}_{ij}^{-1} \hat{r}(l) \right\|_{2p} \left\| \hat{R}(j-k) - R(j-k) \right\|_{2p} . \end{aligned} \tag{3.10}$$

As before we observe

$$\left\|\hat{R}(j-k)-R(j-k)\right\|_{2p}=O(M\cdot n^{-1/2}). \tag{3.11}$$

Since

$$\left\|\hat{R}_{ij}^{-1}\hat{r}(l)\right\|_{2p}=O(1), \tag{3.12}$$

the desired result can be obtained. □

Lemma 3.2 *Suppose that (A.1) and (A.2) hold. If $f=f_\theta$, where θ is the innovation-free parameter, then we have*

$$\max_{1\le i,j\le k}\left\|F_{ij}(\hat{\theta}_n^{(MCE)})-F_{ij}(\theta)\right\|_{p/2}=O(M\cdot n^{-1/2}), \tag{3.13}$$

where $F_{ij}(\hat{\theta}_n^{(MCE)})$ and $F_{ij}(\theta)$ are the (i,j)th elements of $F(\hat{\theta}_n^{(MCE)})$ and $F(\theta)$, respectively.

Proof. On the basis of the mean-value theorem, we have

$$F_{ij}(\hat{\theta}_n^{(MCE)})-F_{ij}(\theta)=\frac{\partial}{\partial\theta}F_{ij}(\theta^*)(\hat{\theta}_n^{(MCE)}-\theta), \tag{3.14}$$

where $\|\hat{\theta}_n^{(MCE)}-\theta^*\|\le\|\hat{\theta}_n^{(MCE)}-\theta\|$. On the basis of Theorem 3 by Taniguchi (1987), we have $\hat{\theta}_n^{(MCE)}\underset{p}{\longrightarrow}\theta$, which implies that $\theta^*\underset{p}{\longrightarrow}\theta$; hence,

$$\begin{aligned}\left\|F_{ij}(\hat{\theta}_n^{(MCE)})-F_{ij}(\theta)\right\|_{p/2} &= \left\|\frac{\partial}{\partial\theta}F_{ij}(\theta^*)(\hat{\theta}_n^{(MCE)}-\theta)\right\|_{p/2}\\ &\le \left\|\frac{\partial}{\partial\theta}F_{ij}(\theta^*)\right\|_{p}\left\|(\hat{\theta}_n^{(MCE)}-\theta)\right\|_{p}. \end{aligned} \tag{3.15}$$

From (A.2) we have $\|\partial/\partial\theta F_{ij}(\theta^*)\|_p=O(1)$. Hence, from Lemma 3.1, we obtain (3.13). □

Lemma 3.3 *Under the same assumptions as those in Lemma 3.2, we have*

$$\|\lambda(F(\hat{\theta}_n^{(MCE)}))-\lambda(F(\theta))\|_{p/2}=O(M\cdot n^{-1/2}). \tag{3.16}$$

In particular, for any $\varepsilon>0$,

$$P(|\lambda(F(\hat{\theta}_n^{(MCE)}))-\lambda(F(\theta))|>\varepsilon)=O(M^p\cdot n^{-p/2}). \tag{3.17}$$

Proof of Theorem 2.1 In order to prove (i), we observe from (3.17) and the Borel-Cantelli lemma that

$$\lambda(F(\hat{\theta}_n^{(MCE)})) \to \lambda(F(\theta)) \text{ a.s. as } n \to \infty. \tag{3.18}$$

Let $f(n) = n\lambda(F(\hat{\theta}_n^{(MCE)}))/\lambda(F(\theta))$ and $t = a^2/d^2\lambda(F(\theta)) = n_0(d) \to \infty$ as $d \to 0$. Then the conditions of Lemma 1 of Chow and Robbins (1965) are satisfied, and hence

$$\lim_{t\to\infty} T_d/t = \lim_{d\to 0} T_d/n_0(d) = 1 \text{ a.s..}$$

It is clear that (ii) implies (iii). So, only (ii) needs to be proved. From Theorem 5 of Taniguchi (1987) we have

$$\sqrt{n}\left(\hat{\theta}_n^{(MCE)} - \theta\right) = \sqrt{n}\int_{-\pi}^{\pi} \rho_{f_\theta}\{\hat{f}_n(\lambda) - f_\theta(\lambda)\}d\lambda + o_p(1), \tag{3.19}$$

where

$$\rho_{f_\theta} = \left[\int_{-\pi}^{\pi} \frac{\partial}{\partial\theta}\log f_\theta(\lambda)\frac{\partial}{\partial\theta'}\log f_\theta(\lambda)\right]^{-1}\frac{\partial}{\partial\theta} f_\theta^{-1}(\lambda). \tag{3.20}$$

This implies that the limiting distribution of $\sqrt{n}(\hat{\theta}_n^{(MCE)} - \theta)$ is described by the integral functional of the spectral density. Let $\xi_n = \sqrt{n}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_n(\lambda) - f_\theta(\lambda)\}d\lambda$. To show (ii), we need to show that the sequence $\{\xi_n, n \geq 1\}$ is uniformly continuous in probability, that is,

$$P\left\{\max_{0\leq k\leq n\delta}\|\xi_{n+k} - \xi_n\| \geq \varepsilon\right\} < \varepsilon \text{ for all } n \geq 1, \tag{3.21}$$

where $\|\cdot\|$ is the Euclidian norm. Essentially,

$$\begin{aligned}
& \|\xi_{n+k} - \xi_n\| \\
= & \left\|\sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_{n+k}(\lambda) - f_\theta(\lambda)\}d\lambda - \sqrt{n}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_n(\lambda) - f_\theta(\lambda)\}d\lambda\right\| \\
= & \left\|\sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\hat{f}_{n+k}(\lambda)d\lambda - \sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\hat{f}_n(\lambda)d\lambda + \sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\hat{f}_n(\lambda)d\lambda\right. \\
& \left. -\sqrt{n}\int_{-\pi}^{\pi}\rho_{f_\theta}\hat{f}_n(\lambda)d\lambda - (\sqrt{n+k} - \sqrt{n})\int_{-\pi}^{\pi}\rho_{f_\theta}f_\theta(\lambda)d\lambda\right\| \\
\leq & \left\|(\sqrt{n+k} - \sqrt{n})\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_n(\lambda) - f_\theta(\lambda)\}d\lambda\right\| + \left\|\sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_{n+k}(\lambda) - \hat{f}_n(\lambda)\}d\lambda\right\|
\end{aligned}$$

It can be observed that

$$\sqrt{n}\left\|\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_n(\lambda) - f_\theta(\lambda)\}d\lambda\right\| = O_p(1).$$

We also observe that

$$\begin{aligned}
&\max_{0\le k\le n\delta}\left\|\sqrt{n+k}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_{n+k}(\lambda)-\hat{f}_n(\lambda)\}d\lambda\right\| \\
= \;& O_p\left(\sqrt{n}\sqrt{1+\delta}\max_{0\le k\le n\delta}\left\|\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_{n+k}(\lambda)-f_\theta(\lambda)\}d\lambda-\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_n(\lambda)-f_\theta(\lambda)\}d\lambda\right\|\right) \\
= \;& O_p(\sqrt{\delta}), \qquad (3.22)
\end{aligned}$$

which, along with (3.22), implies (3.21). Therefore, using Anscombe's theorem we obtain

$$\sqrt{T_d}\int_{-\pi}^{\pi}\rho_{f_\theta}\{\hat{f}_{T_d}(\lambda)-f_\theta(\lambda)\}d\lambda \xrightarrow[\mathscr{L}]{} N(\mathbf{0},F(\theta)^{-1}) \text{ as } d\to 0, \qquad (3.23)$$

which implies (ii). □

On the basis of Theorem 2.1, $d^2T_d\lambda(F(\theta))/a^2 \to 1$ a.s. as $d \to 0$. Hence, to prove the asymptotic efficiency, it is sufficient to show that $\{d^2T_d : d \in (0,1)\}$ is uniformly integrable.

Proof of Theorem 2.2. Let $\delta > 0$ and $K_d = [a^2/d^{-2}\lambda^{-1}(F(\theta))d(1+\delta)]+1$. Then, for $k \ge K_d$ and some $\eta > 0$, it can be shown that

$$\begin{aligned}
P[T_d \ge k] &\le P\left[\left|\lambda^{-1}(F(\hat{\theta}_k^{(MCE)}))-\lambda^{-1}F(\theta)\right|\ge\eta\right] \\
&= O(M^p\cdot k)^{-p/2}), \qquad (3.24)
\end{aligned}$$

where (3.17) is used to obtain the last equation. This implies that $\sum_{k\ge1}P(T_d>k)<\infty$. Based on this and on the arguments by Woodroofe (1982), it follows that

$$\{d^2T_d : d\in(0,1)\} \;\; \textit{is uniformly integrable.} \qquad (3.25)$$

Hence, $\lim_{d\to0}E(T_d/n_0(d)) = 1$. □

4 Some Related Fixed Size Confidence Sets

Fixed proportional accuracy confidence ellipsoids

Suppose $\theta_i, i = 1,\dots,q$, are nonzero and at least one of the parameter values is near the origin, then a smaller confidence ellipsoid can be constructed for θ which gives us an improvement in the accuracy of the estimates of small coordinates. One approach is to construct an ellipsoidal region such that the statistical distance between $\hat{\theta}_n^{(MCE)}$ and θ is less than a certain fraction of the true value of $\theta_{(1)} = \min_{1 \le j \le q} |\theta_j|$. This yields the following ellipsoidal region.

$$\Gamma_n = \left\{ z : (z - \hat{\theta}_n^{(MCE)})' F(\hat{\theta}_n^{(MCE)})(z - \hat{\theta}_n^{(MCE)}) \le d^2 \lambda(F(\hat{\theta}_n^{(MCE)}))\hat{\theta}_{(1),n} \right\} \tag{4.1}$$

for $d > 0$, where $\hat{\theta}_{(1),n} = \min_{1 \le j \le q} |\hat{\theta}_{n,j}^{(MCE)}|$. Γ_n defines an ellipsoid having the length of the major axis equal to $2d\sqrt{\hat{\theta}_{(1),n}}$.

For any given $\alpha \in (0,1)$ and $d > 0$, it is desired to have

$$P[\theta \in \Gamma_n] \approx 1 - \alpha. \tag{4.2}$$

Since $\hat{\theta}_n^{(MCE)} \to \theta$ almost surely, $\hat{\theta}_{(1),n} \to \theta_{(1)}$ a.s. as $n \to \infty$, and therefore,

$$(\hat{\theta}_n^{(MCE)} - \theta)' F'(\hat{\theta}_n^{(MCE)})(\hat{\theta}_n^{(MCE)} - \theta)/\hat{\theta}_{(1),n} \underset{\mathscr{L}}{\longrightarrow} \chi^2(q)/\theta_{(1)}. \tag{4.3}$$

Hence, to satisfy (4.2), we define a sample size

$$t_0(d) = \text{smallest integer } \ge a^2/[d^2 \lambda(F(\theta))\theta_{(1)}], \tag{4.4}$$

where a^2 and $\lambda(F(\theta))$ are defined as in (2.14). Since both $\lambda(F(\theta))$ and $\theta_{(1)}$ are unknown, it is impossible to decide the sample size in advance. This suggests a stopping time

$$N_d = \inf\{n \ge m : n \ge a^2/[d^2 \lambda(F(\hat{\theta}_n^{(MCE)}))\hat{\theta}_{(1),n}]\}. \tag{4.5}$$

Then we obtain the following theorem.

Theorem 4.1 *Suppose that Assumptions (A.1) and (A.2) hold, and $\theta \in \Theta$. Then, for the stopping rule N_d defined in (4.5), the following holds.*

$$(i) \quad N_d/t_0(d) \to 1 \quad a.s. \quad as\ d \to 0, \tag{4.6}$$

$$(ii) \quad \lim_{d \to 0} P[\theta \in \Gamma_{N_d}] = 1 - \alpha, \tag{4.7}$$

where $t_0(d)$ is as in (4.4) and

$$(iii) \quad \{N_d/t_0(d); 0 < d < 1\} \quad \textit{is uniformly integrable}, \tag{4.8}$$

$$(iv) \quad \lim_{d \to 0} E[N_d/t_0(d)] = 1. \tag{4.9}$$

Confidence interval for a linear combination of θ

In practice, we may only be interested in a particular linear combination of the components of θ, rather than the entire vector. That is, for some $C \in \mathbb{R}^q$, $\|C\| \neq 0$, we can construct a fixed-width confidence interval for $C'\theta$. It follows from the asymptotic normality of $\hat{\theta}_n^{(MCE)}$ that, as $n \to \infty$,

$$\sqrt{n}(C'\hat{\theta}_n^{(MCE)} - C'\theta) \underset{\mathscr{L}}{\longrightarrow} N(0, C'F(\theta)^{-1}C). \tag{4.10}$$

If $F(\theta)$ were known, then for a given $d > 0$, $\alpha \in (0,1)$, and the sample size is determined by

$$h_0(d) = \text{smallest interger} \geq z_{\alpha/2}^2/[d^2C'F(\theta)^{-1}C]. \tag{4.11}$$

From (4.10), we have

$$\lim_{d \to 0} P(C'\theta \in [C'\hat{\theta}_n^{(MCE)} - d, C'\hat{\theta}_n^{(MCE)} + d]) = 1 - \alpha, \tag{4.12}$$

where $z_{\alpha^2}^2$ satisfies $\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1 - \alpha$. However, since $F(\theta)$ is unknown, the sample size $h_0(d)$ cannot be used. As observed previously, (4.11) suggests the stopping rule

$$H_d = \inf\{n \geq m : n \geq z_{\alpha/2}^2/[d^2C'F(\hat{\theta}_n^{(MCE)})^{-1}C]\}. \tag{4.13}$$

We have the following theorem.

Theorem 4.2 *Suppose that Assumptions (A.1) and (A.2) hold, and $\theta \in \Theta$. Then, for the stopping rules H_d and $h_0(d)$ defined in (4.13) and (4.11), respectively, the following hold.*

$$(i) \quad H_d/h_0(d) \to 1 \quad a.s. \quad as\ d \to 0, \tag{4.14}$$

$$(ii) \quad \lim_{d \to 0} P\left[C'\theta \in [C'\hat{\theta}_{H_d}^{(MCE)} - d, C'\hat{\theta}_{H_d}^{(MCE)} + d]\right] = 1 - \alpha. \tag{4.15}$$

Furthermore,

$$(iii) \quad \{H_d/h_0(d); 0 < d < 1\} \quad \text{is uniformly integrable}, \tag{4.16}$$

$$(iv) \quad \lim_{d \to 0} E[H_d/h_0(d)] = 1. \tag{4.17}$$

Theorems 4.1 and 4.2 can be proved using arguments similar in the proofs for Theorems 2.1 and 2.2, and are therefore omitted for brevity.

5 Simulations

In this section, we present some Monte Carlo simulations to verify that our sequential procedures are asymptotically consistent and efficient. We consider the standard Gaussian AR(1), MA(1), and AR(2) models.

1. AR(1): $X_t = \theta X_{t-1} + \varepsilon_t,\ \ \theta = 0.1, 0.2, \ldots, 0.9;$
2. MA(1): $X_t = \varepsilon_t + \theta\varepsilon_t,\ \ \theta = 0.1, 0.2, \ldots, 0.9;$
3. AR(2): $X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \varepsilon_t$, where $(\theta_1, \theta_2) = (1.1, -0.24), (0.6, -0.4), (0.2, -0.35), (0.2, 0.35), (0.05, 0.35)$, and $(1.5, -0.75)$.

The innovations $\{\varepsilon_t\}$ are standard normal distributions. For the simulation study we set $n_0(d) = 100, 200, 300, 400, 500$, and the coverage probability as 0.90. The initial sample size m is chosen as $m = 2$ for AR(1) and AR(2) models, and $m = 50$ for MA(1) models. For each choice of parameters, 1000 replications of the series were generated. The results are presented in Tables 1, 2, and 3.

As observed from Tables 1 and 2, the expected sample sizes and the coverage probabilities depend on the value of θ. For smaller values of θ, the agreement between the asymptotic theory and the simulations is better regardless of the value of d. On the other hand, the bias in T_d increases substantially with an increase in θ, while the relative discrepancy decreases with a decrease in d. For AR(1) models, the rate of convergence of the coverage probability is not strongly affected by θ, while that for MA(1) models deteriorates with an increase in θ. It is also noted that for large values of θ, the standard deviations of the expected sample size for MA(1) models are larger than those for AR(1) models.

The situation for AR(2) models appears quite similar to what has been observed for AR(1) models. Let π_1 and π_2 be the roots of the characteristic equation $1 - \theta_1 z - \theta_2 z^2 = 0$. The absolute values of π_1 and π_2 are also provided in Table 3. In case of AR(2) models with absolute values of roots close to the unit circle, the behavior of sequential procedure exhibits a poor performance.

Table 1. Simulation results of 90% confidence interval for AR(1) model.

d	$n_0(d)$	T_d ($s.d.^*$)	$T_d/n_0(d)$	$c.p.^*$.	d	$n_0(d)$	T_d ($s.d.^*$)	$T_d/n_0(d)$	$c.p.^*$
$\theta = 0.1$					$\theta = 0.2$				
0.164	100	99.69 (2.23)	0.9969	.917	0.161	100	100.47 (3.80)	1.0047	.904
0.116	200	199.97 (2.65)	0.9998	.914	0.114	200	200.53 (5.37)	1.0026	.913
0.094	300	299.97 (3.48)	0.9999	.901	0.093	300	300.08 (7.14)	1.0003	.890
0.082	400	400.03 (3.87)	1.0001	.908	0.081	400	400.71 (7.87)	1.0018	.914
0.073	500	499.80 (4.50)	0.9996	.912	0.072	500	500.51 (8.86)	1.0010	.901
$\theta = 0.3$					$\theta = 0.4$				
0.157	100	101.09 (5.66)	1.0109	.910	0.151	100	102.26 (7.54)	1.0226	.921
0.111	200	201.73 (8.01)	1.0086	.907	0.107	200	202.65 (11.50)	1.0132	.907
0.091	300	302.34 (10.19)	1.0078	.900	0.087	300	302.91 (13.94)	1.0097	.901
0.078	400	401.80 (11.96)	1.0045	.914	0.075	400	402.61 (16.44)	1.0065	.907
0.070	500	502.00 (13.82)	1.0040	.897	0.067	500	502.99 (18.44)	1.0060	.906
$\theta = 0.5$					$\theta = 0.6$				
0.142	100	104.44 (10.23)	1.0444	.886	0.132	100	107.78 (12.59)	1.0778	.874
0.100	200	203.93 (14.74)	1.0196	.909	0.093	200	207.24 (19.35)	1.0362	.896
0.082	300	306.13 (18.59)	1.0204	.895	0.076	300	308.66 (25.04)	1.0289	.886
0.071	400	406.76 (22.41)	1.0169	.909	0.068	400	407.99 (29.15)	1.0200	.883
0.064	500	506.65 (24.13)	1.0121	.908	0.059	500	508.28 (32.79)	1.0166	.898
$\theta = 0.7$					$\theta = 0.8$				
0.117	100	111.95 (16.01)	1.1195	.880	0.098	100	120.87 (19.95)	1.2087	.863
0.083	200	213.76 (24.86)	1.0688	.888	0.070	200	222.82 (31.27)	1.1141	.883
0.068	300	311.95 (32.15)	1.0398	.889	0.057	300	326.49 (42.46)	1.0883	.858
0.059	400	412.87 (38.80)	1.0322	.879	0.049	400	426.50 (46.99)	1.0662	.890
0.053	500	515.06 (41.72)	1.0301	.890	0.044	500	527.35 (53.29)	1.0547	.897
$\theta = 0.9$									
0.072	100	150.65 (24.05)	1.5065	.768					
0.051	200	257.24 (39.67)	1.2862	.827					
0.041	300	352.45 (62.97)	1.1748	.864					
0.036	400	454.84 (61.81)	1.1371	.891					
0.032	500	553.63 (71.55)	1.1073	.900					

*s.d., standard deviation; c.p., coverage probability.

Table 2. Simulation results of 90% confidence interval for MA(1) model.

d	$n_0(d)$	T_d $(s.d.^*)$	$T_d/n_0(d)$	$c.p.^*$	d	$n_0(d)$	T_d $(s.d.^*)$	$T_d/n_0(d)$	$c.p.^*$
$\theta = 0.1$					$\theta = 0.2$				
0.164	100	99.62 (2.75)	0.9962	.881	0.161	100	100.04 (4.96)	1.0004	.880
0.116	200	199.75 (2.98)	0.9988	.887	0.114	200	200.43 (5.60)	1.0022	.894
0.094	300	299.92 (3.68)	0.9997	.878	0.093	300	300.22 (7.00)	1.0007	.910
0.082	400	399.77 (4.12)	0.9994	.902	0.081	400	400.36 (8.00)	1.0009	.903
0.073	500	499.55 (4.60)	0.9991	.896	0.072	500	500.17 (9.28)	1.0003	.901
$\theta = 0.3$					$\theta = 0.4$				
0.157	100	100.89 (6.49)	1.0089	.895	0.151	100	101.56 (9.14)	1.0156	.874
0.111	200	200.98 (9.10)	1.0049	.885	0.107	200	201.96 (12.13)	1.0098	.904
0.091	300	301.08 (11.28)	1.0036	.881	0.087	300	302.31 (15.16)	1.0077	.884
0.078	400	401.07 (12.63)	1.0027	.893	0.075	400	401.66 (18.01)	1.0042	.885
0.070	500	501.00 (14.41)	1.0020	.887	0.067	500	502.38 (19.04)	1.0048	.903
$\theta = 0.5$					$\theta = 0.6$				
0.142	100	101.88 (14.14)	1.0188	.835	0.132	100	103.96 (18.26)	1.0396	.804
0.100	200	203.80 (16.51)	1.0190	.892	0.093	200	206.09 (22.72)	1.0305	.874
0.082	300	303.54 (20.25)	1.0118	.883	0.076	300	310.76 (25.03)	1.0359	.882
0.071	400	403.67 (22.64)	1.0092	.889	0.068	400	408.46 (30.05)	1.0212	.892
0.064	500	505.28 (25.19)	1.0106	.914	0.059	500	507.92 (33.09)	1.0158	.880
$\theta = 0.7$					$\theta = 0.8$				
0.117	100	106.92 (25.49)	1.0692	.728	0.098	100	111.41 (37.91)	1.1141	.530
0.083	200	211.36 (33.20)	1.0568	.801	0.070	200	217.90 (56.10)	1.0895	.590
0.068	300	313.84 (38.40)	1.0461	.859	0.057	300	325.51 (67.35)	1.0850	.650
0.059	400	419.07 (39.83)	1.0477	.852	0.049	400	432.80 (75.80)	1.0820	.674
0.053	500	518.10 (46.63)	1.0362	.883	0.044	500	537.25 (78.29)	1.0745	.750
$\theta = 0.9$									
0.072	100	133.05 (64.05)	1.3305	.529					
0.051	200	233.29 (102.83)	1.1665	.374					
0.041	300	337.05 (132.64)	1.1235	.335					
0.036	400	425.08 (158.88)	1.0627	.286					
0.032	500	538.13 (181.26)	1.0763	.299					

*s.d., standard deviation; c.p., coverage probability.

Table 3. Simulation results of 90% confidence interval for AR(2) models.

d	$n_0(d)$	$\bar{T}_d$ $(s.d.^*)$	$\bar{T}_d/n_0(d)$	$c.p.^*$
$\theta_1 = 1.1$, $\theta_2 = -0.24$, $\lvert\pi_1\rvert = 1.25$, $\lvert\pi_2\rvert = 3.33$				
0.286	100	92.45 (5.85)	0.9245	.662
0.202	200	190.15 (8.00)	0.9508	.912
0.165	300	288.42 (9.10)	0.9614	.854
0.143	400	387.02 (11.33)	0.9676	.888
0.128	500	486.87 (12.94)	0.9737	.901
$\theta = 0.6$, $\theta_2 = -0.4$, $\lvert\pi_1\rvert = \lvert\pi_2\rvert = 1.58$				
0.235	100	101.52 (7.62)	1.0152	.776
0.166	200	201.36 (11.50)	1.0068	.906
0.136	300	302.10 (14.93)	1.0070	.898
0.118	400	402.12 (17.97)	1.0053	.888
0.105	500	502.06 (20.70)	1.0041	.891
$\theta_1 = 0.2$, $\theta_2 = -0.35$, $\lvert\pi_1\rvert = \lvert\pi_2\rvert = 1.69$				
0.215	100	101.17 (8.01)	1.0117	.725
0.152	200	201.81 (12.61)	1.0091	.905
0.124	300	302.68 (15.59)	1.0089	.903
0.108	400	401.32 (18.47)	1.0033	.896
0.096	500	502.60 (19.65)	1.0052	.907
$\theta_1 = 0.2$, $\theta_2 = 0.35$, $\lvert\pi_1\rvert = 1.43$, $\lvert\pi_2\rvert = 2.00$				
0.230	100	101.68 (11.21)	1.0168	.749
0.163	200	202.61 (16.64)	1.0130	.909
0.133	300	302.25 (21.18)	1.0075	.894
0.115	400	402.05 (25.00)	1.0051	.897
0.103	500	503.00 (26.95)	1.0060	.903
$\theta_1 = 0.05$, $\theta_2 = 0.35$, $\lvert\pi_1\rvert = 1.62$, $\lvert\pi_2\rvert = 1.76$				
0.209	100	107.61 (9.06)	1.0761	.748
0.148	200	208.14 (15.17)	1.0407	.900
0.120	300	308.85 (19.53)	1.0295	.893
0.104	400	407.45 (24.70)	1.0186	.897
0.093	500	508.01 (27.89)	1.0160	.909
$\theta_1 = 1.5$, $\theta_2 = -0.75$, $\lvert\pi_1\rvert = \lvert\pi_2\rvert = 1.15$				
0.193	100	151.41 (7.62)	1.5141	.045
0.137	200	241.99 (11.49)	1.2099	.611
0.112	300	331.24 (28.77)	1.1041	.751
0.097	400	441.86 (34.84)	1.1047	.707
0.087	500	549.12 (34.86)	1.0982	.714

*s.d., standard deviation; c.p., coverage probability.

References

[1] Bloomfield, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* **60** 217-226.

[2] Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory:* Expanded ed. Holt, Rinehart and Winston, New York.

[3] Chang, Y. -C. I. and Martinsek, A. T. (1992). Fixed size confidence regions for parameters of a logistic regression models. *Ann. Statist.* **20** 1953-1969.

[4] Chow, Y. S. and Robbins, H. E. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean, *Ann. Math. Statist.* **36** 457-462.

[5] Fakhre-Zakeri I. and Lee, S. (1992). Sequential estimation of the mean of a linear process. *Seq. Anal.* **11** 181-197.

[6] Fakhre-Zakeri I. and Lee, S. (1993). Sequential estimation of the mean vector of a multivariate linear process. *J. Multivariate Anal.* **47** 196-209.

[7] Ghosh, M., Mukhopadhyay, N. and Sen, P.K. (1997). *Sequential Estimation.* Wiley, New York.

[8] Hannan, E. J. (1970). *Multiple Time Series.* Wiley, New York.

[9] Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear process, *Ann. Statist.* **10** 132-153 (Correction: (1993) **21** 1115-1117).

[10] Khan, R. A. (1969). A general method of determining fixed-width confidence intervals. *Ann. Math. Statist.* **40** 704-709.

[11] Lee, S. (1994). Sequential estimation for the parameter of a stationary autoregressive model. *Seq. Anal.* **13** 301-317.

[12] Robinson, P. M. (1983). Review of various approaches to power spectrum estimation. *Handbook of Statistics* Vol. 3, Eds.: Brillinger and Krishnaiah. Amsterdam: North-Holland, 343-368.

[13] Shiohama, T. and Taniguchi, M. (2001). Sequential estimation for a functional of the spectral density of a Gaussian stationary process. *Ann. Inst. Statist. Math.* **53** 142-158.

[14] Shiohama, T. and Taniguchi, M. (2004). Sequential estimation for time series regression models. *J. Statist. Plan. Inf.* **123** 295-312.

[15] Sriram, T. N. (1987). Sequential estimation of the mean of a first order stationary process. *Ann. Statist.* **15** 1079-1090.

[16] Sriram, T. N. (2001). Fixed size confidence regions for parameters of threshold AR(1) models. *J. Statist. Plan. Inf.* **97** 293-304.

[17] Srivastava, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *J. Roy. Statist. Soc. Ser.B* **29** 132-140.

[18] Taniguchi, M. (1987). Minimum contrast estimation for spectral density of stationary processes, *J. Roy. Statist. Soc. Ser. B* **49** 315-325.

[19] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Springer, New York.

[20] Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia.

[21] Yu, K. F. (1989). On fixed-width confidence intervals associated with maximum likelihood estimation. *Journ. Theor. Probab.* **2** 193-199.