J. Austral. Math. Soc. (Series A) 48 (1990), 299–319

FUNCTIONAL LEAST SQUARES ESTIMATORS IN AN ADDITIVE EFFECTS OUTLIERS MODEL

SUNIL K. DHAR

(Received 21 September 1988)

Communicated by T. C. Brown

Abstract

Consider the additive effects outliers (A.O.) model where one observes $Y_{j,n} = X_j + v_{j,n}$, $0 \leq j \leq n$, with

$$X_j = \rho X_{j-1} + \varepsilon_j, \qquad j = 0, \pm 1, \pm 2, \dots, |\rho| < 1.$$

The sequence of r.v.s $\{X_j, j \leq n\}$ is independent of $\{v_{j,n}, 0 \leq j \leq n\}$ and $v_{j,n}$, $0 \leq j \leq n$, are i.i.d. with d.f. $(1 - \gamma_n)I[x \geq 0] + \gamma_n L_n(x)$, $x \in \mathbf{R}$, $0 \leq \gamma_n \leq 1$, where the d.f.s L_n, $n \geq 0$, are not necessarily known and ε_j's are i.i.d.. This paper discusses the asymptotic behavior of functional least squares estimators under the above model. Uniform consistency and uniform strong consistency of these estimators are proven. The weak convergence of these estimators to a Gaussian process and their asymptotic biases are also discussed under the above A.O. model.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*) (1985 *Revision*): 62 G 05, 62 M 10.

0. Introduction

Let F and L_n, $n \geq 0$, be distribution functions (d.f.s) on the real line $\mathbf{R}$. Throughout this paper F is assumed to have a density $f \geq 0$. Let $\{\gamma_n, n \geq 0\}$ be a sequence of numbers in $[0, 1]$ converging to 0 as $n \to \infty$. Define

$$\beta_n(x) := (1 - \gamma_n)I[x \geq 0] + \gamma_n L_n(x), \qquad x \in \mathbf{R}, \tag{0.1}$$

where $I[A]$ denotes the indicator function of the event A. Let ε_j, $j = 0, \pm 1, \pm 2, \dots$, be independent and identically distributed (i.i.d.) F random variables (r.v.s), such that the first moment of ε_0 exists and $E\varepsilon_0 = 0$. Let $v_{j,n}$, $0 \leq j \leq n$, be i.i.d. β_n r.v.s.

We consider the model in which one observes, at stage n, r.v.s $Y_{j,n}$, $0 \leq j \leq n$, satisfying

$$(0.2) \qquad Y_{j,n} = X_j + v_{j,n}, \qquad j = 0, 1, \ldots, n,$$

with $\{X_j\}$ obeying the autoregressive model of order one ($AR(1)$), viz.

$$(0.3) \qquad X_j = \rho X_{j-1} + \varepsilon_j, \qquad |\rho| < 1, \; j = 0, \pm 1, \pm 2, \ldots,$$

where $\{X_j\}$ is stationary. Moreover, $\{X_j, j \leq n\}$ is assumed to be independent of $\{v_{j,n}, 0 \leq j \leq n\}$, $n \geq 0$. This paper studies the problem of estimating ρ.

Denby and Martin [5] called the model in (0.2) and (0.3) the additive effects outliers (A.O.) model. All the above assumptions on $\{Y_j, 0 \leq j \leq n\}$, $\{X_j\}$, $\{v_{j,n}, 0 \leq j \leq n\}$ and $\{\varepsilon_j\}$ will be referred to as the model assumptions. The assumptions on $\{v_{j,n}, 0 \leq j \leq n\}$ reflect the situation in which the outliers are isolated in nature. Isolated outliers are defined by Martin and Yohai [9] as the outliers any pair of which are separated in time by a nonoutlier. In [9, Theorem 5.2 and Comment 5.1] they also made the assumption of independence of the process $\{X_j, j \leq n\}$ and $\{v_{j,n}, j = 0, 1, 2, \ldots, n\}$, $n \geq 0$.

Denby and Martin [5] studied the least squares estimator, M-estimators and a class of generalized M-estimators (GM-estimators) of ρ under the above models; they took F and L_n to be $\mathcal{N}(0, \sigma_\varepsilon^2)$ and $\mathcal{N}(0, \sigma^2)$, respectively. Under their A.O. model all of these estimators have non-vanishing asymptotic biases with a possible reduction in biases for GM-estimators.

Heathcote and Welsh [7] proposed a class of minimum distance estimators $\hat{\boldsymbol{\rho}}_n(s)$ of the vector $\boldsymbol{\rho}$ in an autoregressive model of order k, defined so as to minimize

$$M_n(\mathbf{t}, s) = -s^{-2} \log \left| (n-k)^{-1} \sum_{j=k+1}^{n} \exp\{is(X_j - \mathbf{X}'_{j-1}\mathbf{t})\} \right|^2$$

as a function of $\mathbf{t}$, where $\mathbf{X}'_{j-1} = (X_{j-1}, X_{j-2}, \ldots, X_{j-k})$.

We shall study the behavior of this class of estimators of $\rho \in (-1, 1)$, when $k = 1$, under the A.O. model (0.2)–(0.3).

Before proceeding further, observe that the process $\{X_j\}$ is stationary ergodic and X_{j-1} is independent of ε_j, $j \geq 1$. From the assumptions on $v_{j,n}$'s and X_j's it can be seen for each n, that the process $\{(X_j, v_{j,n}), 0 \leq j \leq n\}$ is stationary ergodic and hence so is $\{Y_{j,n}, 0 \leq j \leq n\}$. These observations will be used in the sequel repeatedly.

NOTATION. Throughout this paper, by $o_p(1)$ ($O_p(1)$) is meant a sequence of r.v.s that converges to zero in probability (is tight or bounded in probability); the asymptotic bias of $\hat{\rho}(s)$ is defined as the mean of the asymptotic distribution of $\sqrt{n}[\hat{\rho}(s) - \rho]$. Also, let Z_n be a r.v. with d.f. L_n, $n \geq 0$.

1. Definition of a class of estimators

Define $\mathscr{S} := [-b, -a] \cup [a, b], 0 < a < b$,

(1.1)
$$M_n(t,s) = \begin{cases} -s^{-2} \log \left| n^{-1} \sum_{j=1}^n \exp\{is[Y_{j,n} - Y_{j-1,n}t]\} \right|^2 & \text{if } s \in \mathscr{S}, \\ \frac{1}{n} \sum_{j=1}^n (Y_{j,n} - Y_{j-1,n}t)^2 & \text{if } s = 0. \end{cases}$$

Let K be a compact set containing the true parameter ρ in its nonempty interior. Then $\hat{\rho}_n(s)$ for each $s \in \mathscr{S} \cup \{0\}$ denotes a measurable minimizer of $M_n(\cdot, s)$ when restricted to K. For M_n as in (1.1), $\hat{\rho}_n$ can be uniquely defined to be sample continuous on $\mathscr{S}$. Further $\hat{\rho}_n(s)$ satisfies

(1.2)
$$\inf_{t \in K} M_n(t,s) = M_n(\hat{\rho}_n(s), s).$$

Note that $\hat{\rho}_n(0)$ is the least squares estimator which has been studied in detail by Denby and Martin [5] under the A.O. model; hence we shall not allow s to be zero.

2. Uniform (strong) consistency of $\hat{\rho}_n(s), s \in \mathscr{S}$

In this section we prove uniform (strong) consistency of the estimators $\hat{\rho}_n(s)$, $s \in \mathscr{S}$. The idea of the proof for this result is taken from [4] and [7].

LEMMA 2.1. *For each n, let $Y_{j,n} = X_{j,n} + v_{j,n}$, $j = 0, 1, \ldots, n$, be r.v.s. Let $\kappa_n = n^{-1} \sum_{j=1}^n |v_{j,n}| \wedge 1$. Let $K_n \subset \mathbb{R}^2$ be such that*

$$c_{1,n} = \sup_{(t,s) \in K_n} \{|s|\}, \quad c_{2,n} = \sup_{(t,s) \in K_n} \{|st|\} \quad \text{and} \quad c_{1,n} + c_{2,n} \leq c < \infty.$$

Then

(2.1)
$$\sup_{(t,s) \in K_n} \left| n^{-1} \sum_{j=1}^n [\exp\{is[Y_{j,n} - tY_{j-1,n}]\} - \exp\{is[X_{j,n} - tX_{j-1,n}]\}] \right| \leq [(c+2) \vee 4]\kappa_n.$$

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

PROOF. The triangle inequality and $|e^{it} - e^{is}| \le |t - s| \wedge 2$, $t, s \in \mathbf{R}$, show that the l.h.s. of (2.1) can be bounded by

$$n^{-1} \sum_{j=1}^{n} [\{c_{1,n}|v_{j,n}| + c_{2,n}|v_{j-1,n}|\} \wedge 2]. \tag{2.2}$$

The expression inside the sum in (2.2) is bounded by $[c_{1,n}|v_{j,n}| \wedge 2] + [c_{2,n}|v_{j-1,n}| \wedge 2]$. The r.h.s. of (2.2) is now dominated by $[(c_{1,n} \vee 2) + (c_{2,n} \vee 2)]\kappa_n$. From which the lemma follows.

REMARK. Under condition

$$\text{(a)} \quad \kappa_n \to 0 \text{ in probability} \quad \text{or} \quad \text{(b)} \quad \sum_{n=0}^{\infty} E\kappa_n < \infty, \tag{2.3}$$

the l.h.s. of (2.1) converges to zero in probability or a.s., the latter follows from the Markov inequality and the Borel-Cantelli Lemma applied to κ_n. In the case when $v_{j,n}$ are β_n distributed for any d.f. L_n as in (0.1), $E\kappa_n = n^{-1}(n+1)\gamma_n E\{|Z_n| \wedge 1\}$, and thus $\gamma_n = o(1)$ or $\sum \gamma_n < \infty$ become sufficient conditions for (a) or (b). Further note from here on we shall supress the n in the random variables (r.v.s) $Y_{j,n}$ and $v_{j,n}$.

LEMMA 2.2. *If $\phi_n(\mathbf{t})$ are characteristic functions on $\mathbf{R}^k$ such that $\phi_n(\mathbf{t}) \to \phi(\mathbf{t})$ for each $\mathbf{t} \in \mathbf{R}^k$, then for any compact subset K of $\mathbf{R}^k$*

$$\sup_{\mathbf{t} \in K} |\phi_n(\mathbf{t}) - \phi(\mathbf{t})| \to 0.$$

PROOF. The proof follows from Ash [1, Theorem 3.2.9].

LEMMA 2.3. *Let $X_1, X_2, \ldots$ be a sequence of strictly stationary and ergodic random vectors taking values in $\mathbf{R}^k$. Then*

$$P\left(\overline{\lim_{n}} \sup_{-\infty < x_1, x_2, \ldots, x_k < \infty} |F_n(x_1, \ldots, x_k) - F(x_1, \ldots, x_k)| = 0\right) = 1, \tag{2.4}$$

where $F_n(x_1, x_2, \ldots, x_k)$ is the joint empirical distribution function based on $X_1, \ldots, X_n$ and $F(x_1, \ldots, x_k)$ is the joint distribution of the random vector X_1.

PROOF. See [11] and [6].

To state the next lemma, let

$$D_n(t, s) = \left| \frac{1}{n} \sum_{j=1}^{n} \exp\{is(Y_j - tY_{j-1})\} - \phi_{\varepsilon_1}(s)\phi_{X_0}(s[\rho - t]) \right|, \tag{2.5}$$

$(t, s) \in K \times \mathscr{S}$, where ϕ_X denotes the characteristic function of a r.v. X.

LEMMA 2.4. *Let* $\{X_j\}$ *be as in* (0.3) *with* ε_j*'s i.i.d.. Let* $v_{j,n}$*'s be as in Lemma* 2.1. *Then*

(a) *under* (a) *of* (2.3)

$$\sup_{(t,s)\in K\times\mathscr{S}} D_n(t,s)\to 0 \quad \text{in probability} \tag{2.6}$$

and

(b) *under* (b) *of* (2.3), (2.6) *holds with convergence in probability replaced by a.s. convergence.*

PROOF. We shall give the proof of (b). The proof of (a) follows similarly. From the triangle inequality we get

(2.7)

$$D_n(t,s)\le \left|n^{-1}\sum_{j=1}^{n}\exp\{is(Y_j-Y_{j-1}t)\}-n^{-1}\sum_{j=1}^{n}\exp\{is(X_j-X_{j-1}t)\}\right| + \left|n^{-1}\sum_{j=1}^{n}\exp\{is(X_j-X_{j-1}t)\}-\phi_{\varepsilon_1}(s)\phi_{X_0}(s[\rho-t])\right|.$$

That the first term on the r.h.s. of (2.7) goes to zero a.s. follows from condition (b) of (2.3). Since $\{(X_{j-1},\varepsilon_j)\}$ is a stationary ergodic sequence, Lemma 2.3 yields

(2.8)

$$\begin{aligned}1&=P\left(\sup_{x_1,x_2\in\mathbf{R}}\left|n^{-1}\sum_{j=1}^{n}I[\varepsilon_j\le x_1, X_{j-1}\le x_2]-F_{\varepsilon_1}(x_1)F_{X_0}(x_2)\right|\to 0\right)\\ &\le P\left(\sup_{(s_1,s_2)\in K_1\times K_2}\left|\sum_{j=1}^{n}\exp\{is_1\varepsilon_j+is_2X_{j-1}\}-\phi_{\varepsilon_1}(s_1)\phi_{X_0}(s_2)\right|\to 0\right),\end{aligned}$$

where K_1 and K_2 are any compact subsets of $\mathbf{R}$. The inequality (2.8) follows from the Continuity Theorem and Lemma 2.2. In particular, taking $K_1=\mathscr{S}$ and $K_2=\{s(\rho-t): s\in\mathscr{S}, t\in K\}$ in (2.8), we get

(2.9)

$$P\left(\sum_{(t,s)\in K\times\mathscr{S}}\left|n^{-1}\sum_{j=1}^{n}\exp\{is(X_j-X_{j-1}t)\}-\phi_{\varepsilon_1}(s)\phi_{X_0}(s[\rho-t])\right|\to 0\right)=1.$$

Now the lemma follows from (2.7) and (2.9).

Next, set

$$M(t,s)=-s^{-2}\log|\phi_{\varepsilon_1}(s)\phi_{X_0}(s[\rho-t])|^2, \qquad (t,s)\in K\times\mathscr{S}. \tag{2.10}$$

Note that
(2.11)
$$M(t,s) \geq M(\rho,s) \text{ for all } s \text{ and hence } \inf_{t\in K} M(t,s) = M(\rho,s) \text{ for all } s.$$

THEOREM 2.5. *Let X_j's and $v_{j,n}$'s be as in Lemma 2.4 with*

$$(2.12)\quad |\phi_{\varepsilon_1}(s)| > 0 \text{ and } 0 < |\phi_{X_0}(s[\rho - t])| < 1, \qquad s \in \mathscr{S} \text{ and } t \in K - \{\rho\}.$$

Then the following statements hold.

(a) *Under* (a) *of* (2.3),

$$(2.13)\qquad \sup_{s\in\mathscr{S}} |\hat{\rho}_n(s) - \rho| \to 0 \quad \textit{in probability}.$$

(b) *Under* (b) *of* (2.3), (2.13) *holds with probability convergence replaced by almost sure convergence.*

PROOF. The proof of (b) is as in [4, Theorem 4.1]. We shall give the proof of (a). From (1.1) and (2.10),

$$|M_n(t,s) - M(t,s)| \leq C_0 \log[\{D_n(t,s)/|\phi_{\varepsilon_1}(s)|^2 |\phi_{X_0}(s[\rho - t])|^2\} + 1],$$

where $C_0 = \sup\{s^{-2} : s \in \mathscr{S}\}$. Since

$$\inf_{(s,t)\in\mathscr{S}\times K} |\phi_{\varepsilon_1}(s)|^2 |\phi_{X_0}(s[\rho - t])|^2 > 0,$$

Lemma 2.4 and the above inequality imply that

$$(2.14)\qquad \sup_{(t,s)\in K\times\mathscr{S}} |M_n(t,s) - M(t,s)| = o_p(1).$$

From (1.2), (2.11) and (2.14)
(2.15)
$$\sup_{s\in\mathscr{S}} |M_n(\hat{\rho}_n(s),s) - M(\rho,s)| = \sup_{s\in\mathscr{S}} \left| \inf_{t\in K} M_n(t,s) - \inf_{t\in K} M(t,s) \right| = o_p(1).$$

For $\delta > 0$, let $K(\delta) = K - \{t : |t - \rho| < \delta\}$. Then, (2.14) also implies

$$(2.16)\qquad \sup_{s\in\mathscr{S}} \left| \inf_{t\in K(\delta)} M_n(t,s) - \inf_{t\in K(\delta)} M(t,s) \right| = o_p(1).$$

Suppose that

$$(2.17)\qquad \sup_{s\in\mathscr{S}} |\hat{\rho}_n(s) - \rho| \text{ does not converge to zero in probability.}$$

Then there exists $\eta_0, \eta_1 > 0$ and a sequence of integers $n_k \uparrow \infty$ such that

$$(2.18)\qquad P\left(\sup_{s\in\mathscr{S}} |\hat{\rho}_{n_k}(s) - \rho| > \eta_1 \right) > \eta_0.$$

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

Let

$$\eta = 3^{-1} \inf_{s\in\mathcal{S}} \left| M(\rho, s) - \inf_{t\in K(\eta_1)} M(t,s) \right|.$$

From assumption (2.12), $\eta > 0$. From (2.15), (2.16) and (2.18), there exists $k_0(\eta, \eta_0)$ such that for all $k > k_0(\eta, \eta_0)$

(2.19)

$$\begin{aligned}
\eta_0 &< P\left[\sup_{s\in\mathcal{S}} |M_{n_k}(\hat\rho_{n_k}(s), s) - M(\rho, s)| < \eta, \sup_{s\in\mathcal{S}} |\hat\rho_{n_k}(s) - \rho| > \eta_1\right] + \eta_0/2 \\
&< P\left[\sup_{s\in\mathcal{S}} |M_{n_k}(\hat\rho_{n_k}(s), s) - M(\rho, s)| < \eta, \sup_{s\in\mathcal{S}} |\hat\rho_{n_k}(s) - \rho| > \eta_1 ,\right. \\
&\qquad\qquad \left. \sup_{s\in\mathcal{S}} \left| \inf_{t\in K(\eta_1)} M_{n_k}(t,s) - \inf_{t\in K(\eta_1)} M(t,s) \right| < \eta \right] \\
&\quad + P\left[\sup_{s\in\mathcal{S}} |M_{n_k}(\hat\rho_{n_k}(s), s) - M(\rho, s)| < \eta, \sup_{s\in\mathcal{S}} |\hat\rho_{n_k}(s) - \rho| > \eta_1 ,\right. \\
&\qquad\qquad \left. \sup_{s\in\mathcal{S}} \left| \inf_{t\in K(\eta_1)} M_{n_k}(t,s) - \inf_{t\in K(\eta_1)} M(t,s) \right| \ge \eta \right] + \eta_0/2 \\
&< P\left[\sup_{s\in\mathcal{S}} |M_{n_k}(\hat\rho_{n_k}(s), s) - M(\rho, s)| < \eta, \sup_{s\in\mathcal{S}} |\hat\rho_{n_k}(s) - \rho| > \eta_1 ,\right. \\
&\qquad\qquad \left. \sup_{s\in\mathcal{S}} \left| \inf_{t\in K(\eta_1)} M_{n_k}(t,s) - \inf_{t\in K(\eta_1)} M(t,s) \right| < \eta \right] + 3\eta_0/4.
\end{aligned}$$

From the definition of η, the first term on the r.h.s. of (2.19) is zero, which leads to a contradiction. Therefore (2.17) must be false and hence the result.

REMARK. If the distribution of ε_1 is infinitely divisible then $|\phi_{\varepsilon_1}| > 0$; hence so is $|\phi_{X_1}| > 0$. Also, if the distribution of ε_1 is lattice type then $|\phi_{X_1}(s)| < 1$, for all $s \in \mathbf{R} - \{0\}$, follows from $\phi_{X_1}(s) = \phi_{X_0}(\rho s)\phi_{\varepsilon_1}(s)$ and [3, Theorem 6.4.7]. Thus from the above conditions on ε_1, condition (2.12) is satisfied.

The proof of Theorem 2.5 does not use the existence of f nor does it use any of the moments of ε_0 or Z_n. Note that under the assumptions ε_j's i.i.d. and $|\rho| < 1$, $\{X_j\}$ of (0.3) is invertible and strictly stationary ergodic.

3. Weak convergence of the process $\sqrt{n}(\hat\rho_n(s) - \rho), s \in \mathcal{S}$

In this section we prove the weak convergence of $\sqrt{n}[\hat\rho_n(\cdot) - \rho]$ as a $C(\mathcal{S})$-valued random element. The idea of the proof for this result is taken from [7]. The C.L.T. given by [12] and [13] has been used to prove its finite dimensional distribution convergence. We also discuss the behavior of its asymptotic bias.

Recall from (1.1) that to minimize $M_n(t,s)$ w.r.t t is equivalent to maximizing $U_n^2(t,s) + V_n^2(t,s)$, where for $(t,s) \in K \times \mathcal{S}$,

$$U_n(t,s) = n^{-1} \sum_{j=1}^{n} \cos(s[Y_j - tY_{j-1}]), \tag{3.1}$$

$$V_n(t,s) = n^{-1} \sum_{j=1}^{n} \sin(s[Y_j - tY_{j-1}]) \quad \text{and}$$

$$U_n \equiv V_n \equiv 0, \quad \text{otherwise.}$$

Let

$$m_n \equiv 2^{-1} s^{-2} (\partial/\partial t)(U_n^2 + V_n^2).$$

Then

(3.2)

$$m_n(t,s) = \begin{cases} -\dfrac{1}{ns} \displaystyle\sum_{j=1}^{n} Y_{j-1}\{U_n \sin(s[Y_j - tY_{j-1}]) - V_n \cos(s[Y_j - tY_{j-1}])\}, \\ \hfill (t,s) \in K \times \mathcal{S}, \\ 0, \quad \text{otherwise.} \end{cases}$$

By the Taylor series expansion, we obtain

$$m_n(\hat{\rho}_n(s), s) = m_n(\rho, s) + \frac{\partial}{\partial t} m_n(t,s)|_{t=\overline{\rho}(s)}[\hat{\rho}_n(s) - \rho], \tag{3.3}$$

where

$$|\overline{\rho}_n(s) - \rho| \le |\hat{\rho}_n(s) - \rho|, \qquad s \in \mathcal{S}. \tag{3.4}$$

Theorem 3.1 below shows that $(\partial/\partial t) m_n(t,s)|_{t=\overline{\rho}(s)}$, uniformly in s, converges in probability to a negative number. Hence

$$\sup_{s \in \mathcal{S}} |m_n(\hat{\rho}_n(s), s)| = o_p(1). \tag{3.5}$$

Thus from (3.3) and (3.5) we see that in order to prove the weak convergence of $\sqrt{n}(\hat{\rho}_n(s) - \rho)$, it suffices to study the weak convergence of the process

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

$m_n(\rho, s)$, $s \in \mathscr{S}$. Before stating the next theorem, note that

$$(3.6)\qquad \frac{\partial}{\partial t} m_n(t,s) = s^{-2}\left\{U_n(t,s)\frac{\partial^2}{\partial t^2}U_n(t,s) + V_n(t,s)\frac{\partial^2}{\partial t^2}V_n(t,s) + \left[\frac{\partial}{\partial t}V_n(t,s)\right]^2 + \left[\frac{\partial}{\partial t}U_n(t,s)\right]^2\right\},$$
$$t \in K, s \in \mathscr{S},$$

$$\frac{\partial}{\partial t}U_n(t,s) = -sn^{-1}\sum_{j=1}^{n} Y_{j-1}\sin(s[Y_j - tY_{j-1}]),$$

$$\frac{\partial}{\partial t}V_n(t,s) = sn^{-1}\sum_{j=1}^{n} Y_{j-1}\cos(s[Y_j - tY_{j-1}]),$$

$$(3.7)\qquad \frac{\partial^2}{\partial t^2}U_n(t,s) = -s^2n^{-1}\sum_{j=1}^{n} Y_{j-1}^2\cos(s[Y_j - tY_{j-1}]),$$

$$\frac{\partial^2}{\partial t^2}V_n(t,s) = -s^2n^{-1}\sum_{j=1}^{n} Y_{j-1}^2\sin(s[Y_j - tY_{j-1}]),$$

for all $(t,s) \in K \times \mathscr{S}$. From here on it will be understood that sup is taken over all $s \in \mathscr{S}$, unless specified otherwise.

Theorem 3.1. *In addition to the assumptions of Theorem* 2.5(a) *and all the model assumptions* (0.1)–(0.3), *assume* $\overline{\lim}_n EZ_n^2 < \infty$. *Then*

$$(3.8)\qquad \sup_{s\in\mathscr{S}}\left|\frac{\partial}{\partial t}m_n(t,s)|_{t=r(s)} + |\phi_{\varepsilon_1}(s)|^2 EX_0^2\right| = o_p(1),$$

with $r = \hat{\rho}_n$ *or* $\overline{\rho}_n$.

Proof. Throughout this proof we shall need sup of each of the following random functions

$$(3.9)\qquad |U_n|, |V_n|, \left|\frac{\partial}{\partial t}U_n\right|, \left|\frac{\partial}{\partial t}V_n\right|, \left|\frac{\partial^2}{\partial t^2}U_n\right|, \left|\frac{\partial^2}{\partial t^2}V_n\right|$$

taken over all $(t,s) \in K \times \mathscr{S}$, to be bounded in probability, which is evident from (3.1), (3.7), $E\varepsilon_0^2 < \infty$, the Stationary Ergodic Theorem and

$$(3.10)\qquad En^{-1}\sum_{j=1}^{n} v_{j-1}^2 = \gamma_n EZ_n^2 \to 0,$$

which in turn follows from $\overline{\lim}_n EZ_n^2 < \infty$.

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

We shall now prove (3.8) with $r = \overline{\rho}_n$; the proof for $r = \hat{\rho}_n$ is exactly the same. From the triangle inequality the expression inside the sup on the l.h.s. of (3.8) can be bounded by

$$(3.11)\qquad \left|\frac{\partial}{\partial t} m_n(t,s)|_{t=\overline{\rho}(s)} - \frac{\partial}{\partial t} m_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right| + \left|\frac{\partial}{\partial t} m_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}} - \frac{\partial}{\partial t} m_n(t,s)|_{\substack{t=\rho\\ Y=X}}\right| + \left|\frac{\partial}{\partial t} m_n(t,s)|_{\substack{t=\rho\\ Y=X}} + |\phi_{\varepsilon_1}(s)|^2 E X_0^2\right|$$

Here, and in what follows, $|_{\substack{t=\rho\\ Y=X}}$ means that in the given expression replace Y_j by X_j and t by ρ, etc. The first term in (3.11) can be dominated by

$$(3.12)\qquad C_0\Bigg[\left|\frac{\partial^2}{\partial t^2} U_n(t,s)|_{t=\overline{\rho}(s)} - \frac{\partial^2}{\partial t^2} U_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right| + |U_n(\overline{\rho}(s),s) - U_n(\overline{\rho}(s),s)|_{Y=X}| \left|\frac{\partial^2}{\partial t^2} U_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right| + \left|\frac{\partial^2}{\partial t^2} V_n(t,s)|_{t=\overline{\rho}(s)} - \frac{\partial^2}{\partial t^2} V_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right| + |V_n(\overline{\rho}(s),s) - V_n(\overline{\rho}(s),s)|_{Y=X}| \left|\frac{\partial^2}{\partial t^2} V_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right| + \left|\left[\frac{\partial}{\partial t} U_n(t,s)|_{t=\overline{\rho}(s)}\right]^2 - \left[\frac{\partial}{\partial t} U_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right]^2\right| + \left|\left[\frac{\partial}{\partial t} V_n(t,s)|_{t=\overline{\rho}(s)}\right]^2 - \left[\frac{\partial}{\partial t} V_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}}\right]^2\right|\Bigg].$$

That the sup norms of the second, fourth, fifth and sixth terms in (3.12) go to zero in probability follows from (3.1), (3.7), (3.9), (3.10), the Lipschitz property of the sine and cosine functions, the Stationary Ergodic Theorem and

$$(3.13)\qquad \sup_{S\in\mathscr{S}} |\overline{\rho}(s)| = O_p(1),$$

which in turn follows from Theorems 2.5(a) and (3.4). As the sine and cosine share similar properties, it now remains only to show that the sup norm of the third term in (3.12) goes to zero in probability in order to prove that the

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

sup norm of (3.12) goes to zero in probability. Accordingly

(3.14)

$$\left|\frac{\partial^2}{\partial t^2}V_n(t,s)|_{t=\overline{p}(s)} - \frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\overline{p}(s)\\ Y=X}}\right|$$
$$\leq s^2 n^{-1}\left|\sum_{j=1}^{n}\sin(s[Y_j-\overline{p}(s)Y_{j-1}])[Y_{j-1}^2-X_{j-1}^2]\right|$$
$$+ s^2 n^{-1}\left|\sum_{j=1}^{n}X_{j-1}^2\{\sin(s[Y_j-\overline{p}(s)Y_{j-1}])-\sin(s[X_j-\overline{p}(s)X_{j-1}])\}\right|.$$

Let $C_1 = \sup\{s^2\}$ and $C_2 = \sup\{|s|^3\}$; then the r.h.s. of (3.14) can be dominated by

$$(3.15)\qquad C_1\left[2n^{-1}\sum_{j=1}^{n}|X_{j-1}v_{j-1}| + n^{-1}\sum_{j=1}^{n}v_{j-1}^2\right]$$
$$+ C_2\left[n^{-1}\sum_{j=1}^{n}X_{j-1}^2|v_j| + |\overline{p}(s)|n^{-1}\sum_{j=1}^{n}X_{j-1}^2|v_{j-1}|\right].$$

This follows from (0.2), the fact that the sine function is bounded by 1 and Lipschitz of order 1, with constant 1. That the sup norm of the expression (3.15) goes to zero in probability follows from (0.1), (3.10), (3.13), $E\varepsilon_0^2 < \infty$, $\overline{\lim}_n EZ_n^2 < \infty$, the Markov inequality applied to each of the averages in (3.15) and the independence of $\{X_j, j \leq n\}$ and $\{v_j, 0 \leq j \leq n\}$. This completes the proof that the sup norm of the first term of (3.11) is $o_p(1)$. We shall now show that the sup norm of the second term in (3.11) is $o_p(1)$. It can be dominated by

(3.16)

$$C_0\left|\frac{\partial^2}{\partial t^2}U(t,s)|_{\substack{t=\overline{p}(s)\\ Y=X}}U_n(\overline{p}(s),s)|_{Y=X} - \frac{\partial^2}{\partial t^2}U_n(t,s)|_{\substack{t=\rho\\ Y=X}}U_n(\rho,s)|_{Y=X}\right|$$
$$+ C_0\left|\frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\overline{p}(s)\\ Y=X}}V_n(\overline{p}(s),s)|_{Y=X} - \frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\rho\\ Y=X}}V_n(\rho,s)|_{Y=X}\right|$$
$$+ C_0\left|\left[\frac{\partial}{\partial t}V_n(t,s)\right]^2\Big|_{\substack{t=\overline{p}(s)\\ Y=X}} - \left[\frac{\partial}{\partial t}V_n(t,s)\right]^2\Big|_{\substack{t=\rho\\ Y=X}}\right|$$
$$+ C_0\left|\left[\frac{\partial}{\partial t}U_n(t,s)\right]^2\Big|_{\substack{t=\overline{p}(s)\\ Y=X}} - \left[\frac{\partial}{\partial t}U_n(t,s)\right]^2\Big|_{\substack{t=\rho\\ Y=X}}\right|.$$

From (3.4), (3.7), (3.9), the Lipschitz property of the sine and cosine functions, Theorem 2.5(a) and the Stationary Ergodic Theorem, the sup norm of the third and fourth terms in (3.16) goes to 0 in probability. Since the sine and cosine functions satisfy similar properties, to prove that the sup norm of the expression in (3.16) converges to zero in probability we only need now prove that the sup norm of the second term in (3.16) converges to zero in probability. It can be dominated by

$$(3.17)\quad C_0\left|\left[\frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\overline{\rho}(s)\\ Y=X}} - \frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\rho\\ Y=X}}\right] V_n(\overline{\rho}(s),s)|_{Y=X}\right| + C_0\left|[V_n(\overline{\rho}(s),s)|_{Y=X} - V_n(\rho,s)|_{Y=X}]\frac{\partial^2}{\partial t^2}V_n(t,s)|_{\substack{t=\rho\\ Y=X}}\right|.$$

From (3.1), (3.4), (3.9), the Lipschitz property of the sine function and Theorem 2.5 (a), we get that the sup norm of the second term in (3.17) goes to zero in probability. Let $s^* = \sup\{|s|\}$. The sup norm of the first term in (3.17) can be dominated by

$$(3.18)\qquad C_0 n^{-1}\sum_{j=1}^{n} X_{j-1}^2\left[\left\{s^*|X_{j-1}\sup_{s\in\mathcal{S}}|\hat{\rho}_n(s)-\rho|\right\}\wedge 2\right];$$

this follows from (3.1), (3.4) and the sine function being bounded by 1 and the inequality $|\sin(s) - \sin(t)| \le |s-t| \wedge 2$, $t, s \in \mathbf{R}$. It remains to prove that (3.18) goes to zero in probability. Let $\varepsilon > 0$ be arbitrary. Then for all $n > n_0(m)$,

$$(3.19)\quad P\left(n^{-1}\sum_{j=1}^{n} X_{j-1}^2\left[\left\{s^*|X_{j-1}|\sup_{s\in\mathcal{S}}|\hat{\rho}_n(s)-\rho|\right\}\wedge 2\right] > \varepsilon\right)$$
$$\le P\left(n^{-1}\sum_{j=1}^{n} X_{j-1}^2\left[\left\{s^*|X_{j-1}|\sup_{s\in\mathcal{S}}|\hat{\rho}_n(s)-\rho|\right\}\wedge 2\right] > \varepsilon, \sup_{s\in\mathcal{S}}|\hat{\rho}_n(s)-\rho| \le \frac{1}{m}\right) + \frac{1}{m}$$
$$\le \varepsilon^{-1}EX_0^2\left[\left\{s^*|X_0|\frac{1}{m}\right\}\wedge 2\right] + \frac{1}{m}.$$

This follows from (2.16), the Markov inequality and the stationarity of $\{X_j\}$. In (3.19), taking limit as $n \to \infty$ and then as $m \to \infty$, we see that $E\varepsilon_0^2 < \infty$ and the D.C.T. give that (3.18) converges to zero in probability.

Heathcote and Welsh [7, Theorem 2] proved that the sup norm of the third term in (3.11) goes to zero a.s. This completes the proof of (3.8) as well.

In order to prove the finite dimensional distribution convergence of the process $\sqrt{n}[\hat{\rho}(s) - \rho]$, $s \in \mathcal{S}$, we shall use [12, Theorem 2.1] and [13] to prove the needed C.L.T. To achieve this consider the following lemmas. For the definition of α_ξ-mixing set process see [12].

LEMMA 3.2. *Let θ_n, ω be Borel measurable functions from $\mathbf{R}^2$ to $\mathbf{R}$ and $\xi_{j,n} = \theta_n(Y_{j-1,n}, Y_{j,n})$, $1 \le j \le n$, $n \ge 1$. Let $Y_{j,n} = \omega(X_j, v_{j,n})$, where $\{X_j, j = 0, \pm1, \pm2, \dots\}$ is a stationary process that is strongly α_X-mixing* [8, *Definition* 17.2.1] *and the sequence of independent r.v.s $\{v_{j,n}, 0 \le j \le n\}$ is independent of $\{X_j, 0 \le j \le n\}$, $n \ge 0$. Then $\xi = \{\xi_n, n \ge 1\}$, where $\xi_n = \{\xi_{j,n}, 1 \le j \le n\}$, is strongly α_ξ-mixing with*

$$\alpha_\xi \le \alpha_X. \tag{3.20}$$

PROOF. Let $\theta = (\xi_{1,n}, \dots, \xi_{j,n})$, $\phi = (\xi_{j+k,n}, \dots, \xi_{n,n})$, $I = [(\xi_{1,n}, \cdots, \xi_{j,n}) \in B_1]$ and $J = [(\xi_{j+k,n}, \dots, \xi_{n,n}) \in B_2]$, where $B_1 \in \mathscr{B}(\mathbf{R}^j)$, the Borel σ-field, and $B_2 \in \mathscr{B}(\mathbf{R}^{n-j-k+1})$, $1 \le j \le n$, $2 + j \le j + k \le n$. Let us suppress the n in $v_{j,n}$ and define

$$\phi_{1,j}\colon \mathbf{R}^{j+1} \to \mathbf{R}^j, \qquad [x_0, x_1, \dots, x_j] \to [\theta_n(x_0, x_1), \dots, \theta_n(x_{j-1}, x_j)],$$

$$T_{\mathbf{v},j+1}\colon \mathbf{R}^{j+1} \to \mathbf{R}^{j+1}, \qquad [x_0, x_1, \dots, x_j] \to [\omega(x_0, v_0), \dots, \omega(x_j, v_j)],$$

and $T_{\mathbf{v}^*,n-j-k+2}$ by replacing, in $T_{\mathbf{v},j+1}$, $\mathbf{v}$ by $\mathbf{v}^*$ and j by $n - j - k + 1$, where $\mathbf{v} = (v_0, v_1, \dots, v_j)$ and $\mathbf{v}^* = (v_{j+k-1}, \dots, v_n)$. Then,

$$\begin{aligned}
&|P(I \cap J) - P(I)P(J)| \\
&\quad = |P[\{\omega(X_0, v_0), \dots, \omega(X_j, v_j)\} \in \phi_{1,j}^{-1}(B_1), \\
&\qquad\qquad \{\omega(X_{j+k-1}, v_{j+k-1}), \dots, \omega(X_n, v_n)\} \in \phi_{1,n-j-k+1}^{-1}(B_2)] \\
&\qquad - P[\{\omega(X_0, v_0), \dots, \omega(X_j, v_j)\} \in \phi_{1,j}^{-1}(B_1)] \\
&\qquad\qquad \cdot P[\{\omega(X_{j+k-1}, v_{j+k-1}), \dots, \omega(X_n, v_n)\} \in \phi_{1,n-j-k+1}^{-1}(B_2)]| \\
&\quad \le E[|P[(X_0, \dots, X_j) \in T_{\mathbf{v},j+1}^{-1}\phi_{1,j}^{-1}(B_1), (X_{j+k-1}, \dots, X_n) \\
&\qquad\qquad \in T_{\mathbf{v},n-j-k+2}^{*-1}\phi_{1,n-j-k+1}^{-1}(B_2)] \\
&\qquad - P[(X_0, \dots, X_j) \in T_{\mathbf{v},j+1}^{-1}\phi_{1,j}^{-1}(B_1)]P[(X_{j+k-1}, \dots, X_n) \\
&\qquad\qquad \in T_{\mathbf{v},n-j-k+2}^{*-1}\phi_{1,n-j-k+1}^{-1}(B_2)]| \,|(v_0, \dots, v_j, v_{j+k-1}, \dots, v_n)].
\end{aligned} \tag{3.21}$$

Inequality (3.21) holds because for all $k \ge 2$, $(v_0, \dots, v_j, v_{j+k-1}, \dots, v_n)$ is a sequence of independent r.v.s and independent of $(X_0, \dots, X_n)$. From the definition of strongly α-mixing sequence of stationary r.v.s, and the stationarity of $\{X_j\}$, we see that the r.h.s. of (3.21) can be bounded by α_X. Taking sup over all B_1 and B_2 in $\mathscr{B}(\mathbf{R}^j)$ and $\mathscr{B}(\mathbf{R}^{n-j-k+1})$ and then taking max twice first over all j's in the set $\{1 \le j \le n - k\}$ and next over all k's in the set $\{k\colon k \le n - 1\}$, we get that (3.20) holds.

REMARK. The proof of the Lemma 3.2 goes through even when ω and θ_n are replaced by ω_n and $\theta_{i,n}$ for each j and n.

LEMMA 3.3. *Define* $\xi_{j,n}$ *as in Lemma 3.2 and satisfying all the conditions there.*

(3.22) *In addition, let* $\{v_{j,n}, 0 \le j \le n\}$ *be identically distributed* β_n, *with* $\gamma_n \in [0,1]$, $\gamma_n = o(1)$ *and* α_X *satisfying, for any* $\eta > 0$, $\sum_{j=1}^{\infty} \alpha_X(j)^\eta < \infty$.

(3.23) *Further let* θ *and* h, *be real valued Borel measurable functions with* θ *defined on* $\mathbf{R}^2$ *and* h *on* $\mathbf{R}$, *such that* $\theta_n(x,y) \le Ch(x)$ *and* $\theta_n(x,y) \to \theta(x,y)$ *for each* $x, y \in \mathbf{R}$, $0 < C \in \mathbf{R}$.

Let for some $\delta > 0$,

$$E|h(\omega(X_0))|^{2+\delta} < \infty \quad \text{and} \quad \sup_n E \int |h(\omega(X_0, z))|^{2+\delta} dL_n(z) < \infty, \tag{3.24}$$

where $\omega(\cdot) = \omega(\cdot, 0)$; *then*

$$n^{-1}\sigma_n^2 = n^{-1} \operatorname{Var} \sum_{j=1}^{n} \xi_{j,n} \to \tau^2, \tag{3.25}$$

where

$$\begin{aligned}\tau^2 = {} & \operatorname{Var}[\theta(\omega(X_0), \omega(X_1)] \\ & + 2\sum_{j=1}^{\infty} \operatorname{Cov}[\theta\{\omega(X_0), \omega(X_1)\}, \theta\{\omega(X_j), \omega(X_{j+1})\}].\end{aligned} \tag{3.26}$$

PROOF. From the definition of $Y_{j,n}$'s and the conditions satisfied by X_j's and $v_{j,n}$'s, $\{Y_{j,n}, 0 \le j \le n\}$ is stationary, and hence we can write

$$n^{-1}\sigma_n^2 = \operatorname{Var}(\xi_{1,n}) + \frac{2}{n}\sum_{j=1}^{n-1}(n-j)\operatorname{Cov}[\theta_n(Y_{0,n}, Y_{1,n}), \theta_n(Y_{j,n}, Y_{j+1,n})]. \tag{3.27}$$

From (0.1), the definition of $Y_{j,n}$'s and the conditions satisfied by X_j's and $v_{j,n}$'s,

$$\begin{aligned}\operatorname{Var}(\xi_{1,n}) = {} & (1-\gamma_n)^2 E\theta_n^2\{\omega(X_0), \omega(X_1)\} \\ & + \gamma_n(1-\gamma_n)E\int \theta_n^2\{\omega(X_0), \omega(X_1, z)\}\, dL_n(z) \\ & + \gamma_n E\int \theta_n^2\{\omega(X_0, z), \omega(X_1, v_{1,n})\}\, dL_n(z) - (E\xi_{1,n})^2,\end{aligned} \tag{3.28}$$

which in turn converges to the first term on the r.h.s. of (3.26). The above convergence follows from $\gamma_n = o(1)$, (3.23), (3.24) and the D.C.T. From (0.1), the definition of $Y_{j,n}$'s and the conditions satisfied by X_j's and $v_{j,n}$'s for $j \ge 2$, we get

(3.29) $\operatorname{Cov}[\theta_n(Y_{0,n}, Y_{1,n}),\ (Y_{j,n}, Y_{j+1,n})]$ = Sum of the terms of the type $\gamma_n^k(1-\gamma_n)^{4-k}\operatorname{Cov}[\phi_{n2}(X_0, X_1), \phi_{n3}(X_j, X_{j+1})]$, where $\phi_{n2}(X_0, X_1)$ can be one of the following,

$$\theta_n\{\omega(X_0), \omega(X_1)\},$$
$$\int \theta_n\{\omega(X_0), \omega(X_1, z)\}\, dL_n(z),$$
$$\int \theta_n\{\omega(X_0, z), \omega(X_1)\}\, dL_n(z) \quad \text{or}$$
$$\iint \theta_n\{\omega(X_0, z), \omega(X_1, u)\}\, dL_n(z) dL_n(u),$$

$\phi_{n3}(X_j, X_{j+1})$ is similarly defined and $0 \le k \le 4$.

From computations similar to those in (3.29), $\gamma_n = o(1)$, (3.23), (3.24) and applying the D.C.T., one can show that for $j = 1$

$$\text{(3.30)} \qquad \operatorname{Cov}[\theta_n(Y_0, Y_1), \theta_n(Y_j, Y_{j+1})] \\ \to \operatorname{Cov}[\theta\{\omega(X_0), \omega(X_1)\}, \theta\{\omega(X_j), \omega(X_{j+1})\}].$$

Similarly, from (3.29), (3.30) holds for $j \ge 2$.

From (3.23) and (3.24) we see for each $n \ge 1$

$$\text{(3.31)} \qquad E|\phi_{n2}(X_0, X_1)|^{2+\delta} \le c_1 < \infty \quad \text{and} \\ E|\phi_{n3}(X_0, X_1)|^{2+\delta} \le c_2 < \infty, \qquad 0 < c_1, c_2 \in \mathbf{R}.$$

Applying [8, Theorem 17.22] to the sequence $\{X_j\}$ with $t = 1$, $\tau = j - 1$, $j \ge 2$, $\xi = \phi_{n2}(X_0, X_1)$, $\eta = \phi_{n3}(X_j, X_{j+1})$, and from (3.31), we obtain

$$\text{(3.32)} \qquad |\operatorname{Cov}[\phi_{n2}(X_0, X_1), \phi_{n3}(X_j, X_{j+1})]| \le C\alpha_X(j-1)^{\delta/(2+\delta)},$$

where $0 < C \in \mathbf{R}$ depends only on c_1, c_2 and δ. Thus from the conditions satisfied by α_X in (3.22), (3.29), (3.32) and the D.C.T. for counting measure, we see that the second term on the r.h.s. of (3.27) converges to the second term on the r.h.s. of (3.26). Hence (3.27) and the convergence of (3.28) to the appropriate limit imply that (3.25) holds.

THEOREM 3.4. *Under* (3.22)–(3.24) *and the assumption* $\tau^2 > 0$,

$$n^{-1/2} \sum_{j=1}^{n} \{\xi_{j,n} - E\xi_{j,n}\} \to \mathcal{N}(0, \tau^2)$$

in distribution.

PROOF. In view of [12, Theorem 2.1] and [13] we shall first show that

$$\text{(3.33)} \qquad \sup_{a,n} E|S_n(a,b)|^{2+\eta} = O(b^{1+\eta/2+\eta_1}) \qquad \text{as } b \to \infty,$$

where $S_n(a,b) = \sum_{j=a+1}^{a+b} \xi_{j,n}$, $0 \le a$, $1 \le b \le n-a$, and the set process ξ is given by (3.22). Let $\eta = \delta$ which is as in (3.24) and $\eta_1 = 1 + \delta/2$. Then

$$E\left|\frac{S_n(a,b)}{b}\right|^{2+\delta} \le E|\theta_n(Y_{0,n}, Y_{1,n})|^{2+\delta}$$
$$\le C\left\{(1-\gamma_n)E|h\{\omega(X_0)\}|^{2+\delta} + \gamma_n E\int |h\{\omega(X_0,z)\}|^{2+\delta}\, dL_n(z)\right\}.$$

The above inequality follows from the definition of $Y_{j,n}$'s and the conditions satisfied by X_j's and $v_{j,n}$'s, (0.1), (3.23) and the Jensen inequality. From (3.24), (3.33) is satisfied. From Lemma 3.2 and the fact that $l(k,u) \le 16\alpha(k)$, $0 \le k < \infty$, u real, we have $l(k,u) = o(k^{-\delta})$ as $k \to \infty$, where $\delta = 2\eta_1/\eta$, is satisfied. From Lemma 3.3 and $\tau^2 > 0$, $\sigma_n^2 \to \infty$ as $n \to \infty$ and $\underline{\lim}_n\, \sigma_n^2/n > 0$ are satisfied. Since $\{Y_j, 0 \le j \le n\}$ is stationary, $\tilde{c}_n(j) = \sup|\operatorname{Cov}(\xi_{d,n}, \xi_{m,n})|$, $0 \le j \le n$, can be written as $\tilde{c}_n(j) = \sup|\operatorname{Cov}[\xi_{1,n}, \xi_{|d-m|+1,n}]|$, $0 \le j \le n$, where sup is taken over $\{d, m\colon |d-m| \ge j\}$. From this, (3.10), the same argument as in (3.31) to (3.32) and (3.24), we get

$$\tilde{c}(j) \le 8C\alpha_X(j-1)^{\delta/(2+\delta)}, \quad \text{for } j \ge 2. \tag{3.34}$$

From the conditions satisfied by α_X in (3.22), (3.24) and (3.34), $\sum_{j=0}^{\infty} \tilde{c}(j) < \infty$ is satisfied and hence the C.L.T. holds for ξ.

NOTATION. Next, set $u(s) = E\cos[s\varepsilon_1]$ and $v(s) = E\sin[s\varepsilon_1]$, $s \in \mathbb{R}$. Also, by the random elements X_n and Y_n satisfying $X_n(s) = Y_n(s) + \overline{o}_p(1)$ we shall mean $\sup_{s\in\mathscr{S}} |X_n(s) - Y_n(s)| = o_p(1)$.

THEOREM 3.5. *Let the assumptions of Theorem* 3.1 *hold. Also let*

(a) $$\int |f(x-u) - f(x)|\, dx < C|u|, \qquad u \in \mathbb{R},\ \textit{for some } 0 < C \in \mathbb{R},$$

(b) $$\operatorname{Var}[\cos(s\varepsilon_1)] > 0, \quad \operatorname{Var}[\sin(s\varepsilon_1)] > 0,$$
$$\operatorname{Var}[u(s)\sin(s\varepsilon_1) - v(s)\cos(s\varepsilon_1)] > 0, \qquad s \in \mathscr{S},$$

and

(c) $$\sup_n E|Z_n|^{2+\alpha} < \infty, \qquad 0 < E|\varepsilon_1|^{2+\alpha} < \infty, \qquad \alpha > 0,$$

hold. Then

$$n^{1/2}[\hat{\rho}_n(\cdot) - \rho + \mu_n(\cdot)] \tag{3.35}$$

converges weakly in $C(\mathscr{S})$ *to a Gaussian process with mean* 0 *and covariance* $(EX_0^2)^{-1}[|\phi(s)||\phi(t)|]^{-2}(st)^{-1}h(t,s)$, *where*

$$2h(t,s) = u(s-t)[u(s)u(t) + v(s)v(t)] + u(s+t)[v(s)v(t) - u(s)u(t)]$$
$$+ v(s-t)[v(s)u(t) - u(s)v(t)] - v(s+t)[u(s)v(t) + v(s)u(t)]$$

and

$$\mu_n(s) = s^{-1}\left[\frac{\partial}{\partial t} m_n(t,s)|_{t=\bar{\rho}(s)}\right]^{-1}$$

(3.36)
$$\cdot\Big[E\cos[s(\varepsilon_1+v_1-\rho v_0)]Ev_0\sin[s(\varepsilon_1+v_1-\rho v_0)] - E\sin[s(\varepsilon_1+v_1-\rho v_0)]Ev_0\cos[s(\varepsilon_1+v_1-\rho v_0)]\Big].$$

PROOF. From (3.1), (3.3) and (3.5)–(3.7) we get

(3.37)
$$\begin{aligned}
&n^{1/2}(\hat{\rho}_n(s)-\rho)s\frac{\partial}{\partial t}m_n(t,s)|_{t=\bar{\rho}(s)}\\
&= -\bar{n}^{1/2}\sum_{j=1}^{n} Y_{j-1}\{U_n(\rho,s)\sin[s[Y_j-\rho Y_{j-1}]) - V_n(\rho,s)\cos(s[Y_j-\rho Y_{j-1}])\} + \bar{o}_p(1)\\
&= -n^{-1/2}\sum_{j=1}^{n} Y_{j-1}\Big[\{E\cos[s(\varepsilon_1+v_1-\rho v_0)]\}\sin[s(\varepsilon_j+v_j-\rho v_{j-1})]\\
&\qquad - \{E\sin[s(\varepsilon_1+v_1-\rho v_0)]\}\cos[s(\varepsilon_j+v_j-\rho v_{j-1})]\Big]\\
&- n^{-1}\sum_{j=1}^{n} Y_{j-1}\Big[\{\sin[s(\varepsilon_j+v_j-\rho v_{j-1})] - E\sin[s(\varepsilon_1+v_1-\rho v_0)]\}\cdot\\
&\qquad \cdot\sqrt{n}\{U_n(\rho,s) - E\cos[(\varepsilon_1+v_1-\rho v_0)]\}\Big]\\
&+ n^{-1}\sum_{j=1}^{n} Y_{j-1}\Big[\{\cos[s(\varepsilon_j+v_j-\rho v_{j-1})] - E\cos[s(\varepsilon_1+v_1-\rho v_0)]\}\\
&\qquad \cdot\sqrt{n}\{V_n(\rho,s) - E\sin[s(\varepsilon_1+v_1-\rho v_0)]\}\Big]\\
&- \left\{n^{-1}\sum_{j=1}^{n} Y_{j-1}\right\}\Big[E\sin[s(\varepsilon_1+v_1-\rho v_0)]\sqrt{n}\{U_n(\rho,s) - E\cos[s(\varepsilon_1+v_1-\rho v_0)]\}\\
&\qquad - E\cos[s(\varepsilon_1+v_1-\rho v_0)]\sqrt{n}\{V_n(\rho,s) - E\sin[s(\varepsilon_1+v_1-\rho v_0)]\}\Big]\\
&+ \bar{o}_p(1).
\end{aligned}$$

We shall now proceed to prove that the sup norm of the second, third and fourth terms in (3.37) converge to zero in probability. To achieve this we *shall prove*

(3.38)
$$\begin{aligned}
&\sqrt{n}\{U_n(\rho,s) - E\cos[s(\varepsilon_1+v_1-\rho v_0)]\},\\
&\sqrt{n}\{V_n(\rho,s) - E\sin[s(\varepsilon_1+v_1-\rho v_0)]\},\quad s\in\mathscr{S},
\end{aligned}$$

https://doi.org/10.1017/S1446788700035709 Published online by Cambridge University Press

converges weakly in $C(\mathscr{S})$ to a Gaussian process,

$$\text{(3.39i)}\quad n^{-1}\sum_{j=1}^{n} Y_{j-1}\{\sin[s(\varepsilon_j+v_j-\rho v_{j-1})]-E\sin[s(\varepsilon_1+v_1-\rho v_0)]\} = \bar{o}_p(1)$$

and

$$\text{(3.39ii)}\quad n^{-1}\sum_{j=1}^{n} Y_{j-1}\{\cos[s(\varepsilon_j+v_j-\rho v_{j-1})]-E\cos[s(\varepsilon_1+v_1-\rho v_0)]\} = \bar{o}_p(1).$$

In view of (3.37)–(3.39), to study the weak convergence of $n^{1/2}[\hat{\rho}_n-\rho+\mu_n)]$ it suffices to study the weak convergence of the first term in (3.37) when centered.

PROOF OF (3.38). Denote

$$\xi_{j,n}(s) = \cos[s(\varepsilon_j + v_j - \rho v_{j-1})] - E\cos[(\varepsilon_1 + v_1 - \rho v_0)] \quad \text{for all } s \in \mathscr{S}.$$

Since $Y_j - \rho Y_{j-1}$ and $Y_k - \rho Y_{k-1}$ are independent for all $|j-k| \geq 2$ we see that the set process ξ given by $\xi_{j,n}(s)$ as above, is strongly α-mixing with $\alpha(k) = 0$, for $k \geq 2$. Also, $s \in \mathscr{S}$, because $\gamma_n \to 0$, and so

$$\text{(3.40)}\qquad n^{-1}\sigma_n^2 = E\xi_{1,n}^2(s) + 2(n-1)n^{-1}E\xi_{1,n}(s)\xi_{2,n}(s) \to \operatorname{Var}[\cos(s\varepsilon_0)],$$

which is positive because of the assumption (b). The remaining conditions of [12, Theorem 2.1] and [13] are trivially satisfied.

Hence from [2, page 49] the finite dimensional distributions of

$$\text{(3.41)}\qquad n^{-1/2}\sum_{j=1}^{n}\xi_{j,n}(s)$$

converge to that of a Gaussian process. Also for any $s, t \in \mathscr{S}$

$$\text{(3.42)}\quad E\left|n^{1/2}\sum_{j=1}^{n}\xi_{j,n}(s) - n^{-1/2}\sum_{j=1}^{n}\xi_{j,n}(t)\right|^2 = \operatorname{Var}[\xi_{1,n}(s) - \xi_{1,n}(t)] + \frac{(n-1)}{n}\operatorname{Cov}[\xi_{1,n}(s) - \xi_{1,n}(t), \xi_{2,n}(s) - \xi_{2,n}(t)].$$

From the Cauchy-Schwartz inequality and the Lipschitz property of the cosine function the r.h.s. of (3.42) is dominated by $C|t-s|^2$, where $0 < C \in \mathbf{R}$.

(3.43) Thus from [2, Theorem 12.3], the process $n^{-1/2}\sum_{j=1}^{n}\xi_{j,n}(s)$ is tight, and its weak convergence to a Gaussian limit in $C(\mathscr{S})$-space follows from [2, Theorem 8.1].

The weak convergence of the second process in (3.38) to a Gaussian limit in $C(\mathscr{S})$-space follows similarly.

PROOF OF (3.39). The l.h.s. of (3.39i) without the sup can be dominated by

$$(3.44)\qquad 2n^{-1}\sum_{j=1}^{n}|v_{j-1}| + \left|n^{-1}\sum_{j=1}^{n}X_{j-1}\{\sin[s(\varepsilon_j + v_j - \rho v_{j-1})] - \sin[s\varepsilon_j]\}\right| + \left|n^{-1}\sum_{j=1}^{n}X_{j-1}\sin[s\varepsilon_j]\right| + \left|n^{-1}\sum_{j=1}^{n}X_{j-1}\right|.$$

That the first term in (3.44) goes to zero in probability follows from the assumption (c) and $\gamma_n = o(1)$. From the Lipschitz property of the sine function, the Cauchy-Schwartz inequality, the Stationary Ergodic Theorem, assumption (c) and $\gamma_n = o(1)$, the sup norm of the second term in (3.44) converges to zero in probability. That the sup norm of the third term in (3.44) converges to zero a.s. follows from [7, Lemma 3.1]. The last term in (3.44) converges a.s. to 0 by the Stationary Ergodic Theorem. The proof of (3.39ii) is similar.

It remains to study the weak convergence of the first term in (3.37) when centered. To that effect let

$$\begin{aligned}\xi_{j,n}(s) = Y_{j-1}\Big[&\{E\cos[s(\varepsilon_1 + v_1 - \rho v_0)]\}\sin[s(\varepsilon_j + v_j - \rho v_{j-1})] \\ &- \{E\sin[s(\varepsilon_1 + v_1 - \rho v_0)]\}\cos[s(\varepsilon_j + v_j - \rho v_{j-1})]\Big] \\ &- s\frac{\partial}{\partial t}m_n(t,s)\Big|_{t=\overline{\rho}(s)}\mu_n(s).\end{aligned}$$

We shall first prove the finite dimensional distributions convergence of $n^{-1/2}\sum_{j=1}^{n}\xi_{j,n}(\cdot)$, using Theorem 3.4. Put

$$\begin{aligned}\theta_n(x,y) = x[&\{E\cos[s(\varepsilon_1 + v_1 - \rho v_0)]\}\sin[s(y-\rho x)] \\ &- \{E\sin[s(\varepsilon_1 + v_1 - \rho v_0)]\}\cos[s(y-\rho x)]],\end{aligned}$$

and

$$\theta(x,y) = x\{u(s)\sin[s(y-\rho x)] - v(s)\cos[s(y-\rho x)]\}, \qquad x, y \in \mathbf{R},$$

and take $h \equiv x$, X_j, Y_j, v_j as in the model assumptions with $\omega(x,y) = x + y$, $x, y \in \mathbf{R}$, in Theorem 3.4. Since $X_j = \sum_{k=0}^{\infty}\rho^k\varepsilon_{j-k}$ a.s., using assumptions (a) and (c) and [10, Theorem 2.1] with $\delta = 2$, $A(k) = \rho^k$, gives $\{X_j\}$ to be strongly α-mixing with

$$(3.45)\qquad \alpha(k) \le C_\rho|\rho|^{2k/3}, \quad \text{for } k \ge 2, \text{ for some } C_\rho > 0.$$

By assumption (b) and (c), $\tau^2 = EX_0^2\operatorname{Var}[u(s)\sin(s\varepsilon_0) - v(s)\cos(s\varepsilon_0)] > 0$ for each $s \in \mathscr{S}$. Thus all the conditions of Theorem 3.4 are satisfied. Hence the C.L.T. holds for ξ as defined above, for each $s \in \mathscr{S}$. Now using the argument as in (3.41) we get the required finite dimensional distributions convergence. Since for all i and j with $|i-j| \geq 2$, ε_j, v_j and v_{j-1} are independent of $\{(v_k, Y_k), k = i-1, i\}$, we have

$$\text{(3.46)} \qquad \operatorname{Cov}[\xi_{i,n}(t), \xi_{j,n}(s)] = 0 \quad \text{for all } s, t \in \mathscr{S}.$$

Using (3.46), the same argument as in equations (3.42) and (3.43), we get $n^{-1/2}\sum_{j=1}^{n}\xi_{j,n}(\cdot)$ converges weakly in $C(\mathscr{S})$-space to a Gaussian process with mean 0 and covariance $EX_0^2 h(t,s)$. Thus from (3.3)–(3.5), (3.8), [2, Theorem 4.1] and (3.37), we get (3.35).

Remark. From (3.8), the assumption $\sqrt{n}\gamma_n = O(1)$ and simple computations using (0.1), we can see that $\sqrt{n}\mu_n$ in Theorem 3.5 can be replaced by

$$\text{(3.47)} \quad \begin{aligned} \nu_n(s) = & -n^{1/2}s^{-1}|\phi_{\varepsilon_1}(s)|^{-2}[EX_0^2]^{-1} \\ & \cdot \Big[E\cos[s(\varepsilon_1 + v_1 - \rho v_0)]Ev_0\sin[s(\varepsilon_1 + v_1 - \rho v_0] \\ & \qquad - E\sin[s(\varepsilon_1 + v_1 - \rho v_0)]Ev_0\cos[s(\varepsilon_1 + v_1 - \rho v_0)]\Big]. \end{aligned}$$

Note that $\nu_n(s)$ represents the asymptotic bias of $n^{1/2}(\hat{\rho}_n(s) - \rho)$. Consider the following assumptions:

(a) $\sqrt{n}\gamma_n = o(1)$;
(b) $\sqrt{n}\gamma_n = O(1)$ and $Z_n \to 0$ in probability;
(c) $\sqrt{n}\gamma_n \to \gamma$ and $Z_n \to Z$ in probability.

Using (0.1) and the continuity of the sine and cosine functions, one concludes that under (a) or (b), $\sup_{s\in\mathscr{S}}|\nu_n(s)| \to 0$ and hence $\sqrt{n}\mu_n$ in Theorem 3.5 can be replaced by 0. Using the Lipschitz property of the sine and the cosine functions, the condition (c) implies that $\sup_{s\in\mathscr{S}}|\nu_n(s) - \mu(s)| \to 0$, where

$$\begin{aligned} \mu(s) = & -s^{-1}|\phi_{\varepsilon_1}(s)|^{-2}[EX_0^2]^{-1}\left[\left\{\gamma E\int z\sin[s(\varepsilon_1 - \rho z)]dL(z)\right\}\right. \\ & \cdot\left[u(s) + 2\gamma E\int\cos[s(\varepsilon_1 + 2^{-1}(1-\rho)z)]\cos[s2^{-1}(1+\rho)z]dL(z)\right] \\ & - \left\{\gamma E\int z\cos[s(\varepsilon_1 - \rho z)]dL(z)\right\} \\ & \left.\cdot\left[v(s) + 2\gamma E\int\sin[s(\varepsilon_1 + 2^{-1}(1-\rho)z)]\cos[s2^{-1}(1+\rho)z]dL(z)\right]\right]. \end{aligned}$$

Consequently $\sqrt{n}\mu_n(s)$ in Theorem 3.5 can be replaced by $\mu(s)$.

REMARK. If f is a double exponential or $\mathcal{N}(0, \sigma^2)$ density, Z_n is such that $\overline{\lim}_n E|Z_n|^3 < \infty$ and $\gamma_n = o(1)$, then simple calculations show that all the conditions of Theorem 3.5 are satisfied.

Acknowledgements

I wish to express my thanks to Professor Hira Lal Koul for his guidance and Professor James Hannan for his constructive criticism.

References

[1] R. B. Ash, *Real analysis and probability* (Academic Press, New York, 1972).
[2] P. Billingsley, *Convergence of probability measure* (Wiley New York, 1968).
[3] K. L. Chung, *A course in probability theory*, 2nd ed. (Academic Press, New York, 1974).
[4] S. Csörgö, 'The theory of functional least squares', *J. Austral. Math. Soc. Ser. A* **34** (1983), 336–355.
[5] L. Denby and D. Martin, 'Robust estimation of the first order autoregressive parameter', *J. Amer. Statist. Assoc.* **74** (365) (1979), 140–146.
[6] P. Gaenssler and W. Stute, *On uniform convergence of measures with applications to uniform convergence of empirical distributions*, (Lecture Notes in Math. 566, 1976).
[7] C. R. Heathcote and A. H. Welsh, 'The robust estimation of auto-regressive processes by functional least squares', *J. Appl. Prob.* **20** (1983), 737–753.
[8] I. A. Ibragimov and Y. V. Linnik, *Independent and stationary sequences of random variables* (Wolters-Noordhoff, Groningen, 1971).
[9] D. Martin and V. J. Yohai, 'Influence functionals for time series', *Ann. Statist.* **14** (3) (1986), 781–818.
[10] T. D. Pham and L. T. Tran, 'Some mixing properties of time series models', *Stochastic Processes Appl.* **19** (1985), 297–303.
[11] W. Stute and G. Schumann, 'A general Glivenko-Cantelli theorem for stationary sequences of random observations', *Scand. J. Statist.* **7** (1980), 102–104.
[12] C. S. Withers, 'Central limit theorems for dependent variables. I', *Z. Wahrsch. Verw. Gebiete* **57** (1981), 509–534.
[13] C. S. Withers, 'Corrigendum to central limit theorems for dependent variables. I', *Z. Wahrsch. Verw. Gebiete* **63** (1983), 555.

Department of Mathematics
The University of Alabama
P.O. Box 870350
Tuscaloosa, Alabama 35487-0350
U.S.A.