ASYMPTOTIC NORMALITY OF NEAREST NEIGHBOR
REGRESSION FUNCTION ESTIMATES BASED ON
NONSTATIONARY DEPENDENT OBSERVATIONS

Michel Harel and Madan L. Puri
Université de Limoges, U.R.A. C.N.R.S. 745, Toulouse, France
and Indiana University, Bloomington, IN 47405

SYNOPTIC ABSTRACT

In this paper the convergence of the regression function estimators and the central limit theorem for these estimators are proved for the case when the underlying sequence of random variables is dependent and nonstationary.

Key Words and Phrases: regression function; nearest neighbor regression function estimates; absolute regularity; ϕ -mixing; strong mixing; convergences in law, in probability and with probability equal to one.

1. INTRODUCTION

Let $\{\mathbf{Z}_i = (X_i, Y_i), i \geq 1\}$ be a sequence of random variables with continuous d.f.'s (distribution functions) $H_i(\mathbf{z}), i \geq 1, \mathbf{z} \in \mathbb{R}^2$, and we assume that $H_i(\mathbf{z})$ admits a strictly positive density $h_i(\mathbf{z}), i \geq 1$.

Assume that the process satisfies the absolute regularity condition:

$$\max_{j \geq 1} E \left\{ \sup_{A \in \sigma(\mathbf{Z}_i, i \geq j+m)} |P(A) - \sigma(\mathbf{Z}_i, 1 \leq i \leq j) - P(A)| \right\} = \beta(m) \downarrow 0 \quad (1)$$

as $m \rightarrow \infty$.

Here $\sigma(\mathbf{Z}_i, 1 \leq i \leq j)$ and $\sigma(\mathbf{Z}_i, i \geq j+m)$ are the σ -field generated by $(\mathbf{Z}_1, \dots, \mathbf{Z}_j)$ and $(\mathbf{Z}_{j+m}, \mathbf{Z}_{j+m+1}, \dots)$ respectively. Also recall that $\{\mathbf{Z}_i\}$ satisfies the strong mixing condition if

$$\max_{j \geq 1} [\sup \{ |P(A \cap B) - P(A)P(B)|; A \in \sigma(\mathbf{Z}_i, 1 \leq i \leq j), B \in \sigma(\mathbf{Z}_i, i \geq j+m) \}] = \alpha(m) \downarrow 0 \text{ as } m \rightarrow \infty$$

and $\{\mathbf{Z}_i\}$ satisfies the ϕ -mixing condition if

$$\max_{j \geq 1} [\sup \{ |P(A|B) - P(A)|; B \in \sigma(\mathbf{Z}_i, 1 \leq i \leq j), A \in \sigma(\mathbf{Z}_i, i \geq j+m) \}] = \phi(m) \downarrow 0 \text{ as } m \rightarrow \infty$$

Since $\alpha(m) \leq \beta(m) \leq \phi(m)$, it follows that if $\{\mathbf{Z}_i\}$ is absolutely regular, then it is also strong mixing and if $\{\mathbf{Z}_i\}$ is ϕ -mixing, it is also absolutely regular.

Suppose that H_i converges to a distribution function H which admits a density h and H have the two marginals F and G with f and g as densities.

For a random vector (X, Y) in $\mathbb{R}^2$ with finite expectation $E|Y|$ and which admits H (the distribution limit of H_i) as distribution function, the regression function is defined by $m(x) = E(Y|X = x), x \in \mathbb{R}$ of Y on X and is a.s. in x uniquely defined in view of the equation

$$m(X) = E(Y|X).$$

Consider the nearest neighborhood estimate

$$\tilde{m}_n(x_0) = (na_n)^{-1} \sum_{i=1}^n Y_i K \left(\frac{\tilde{F}_n(x_0) - \tilde{F}_n(X_i)}{a_n} \right) \quad (2)$$

of $m(x_0)$ where K is an appropriate kernel function, a_n is a sequence of bandwidths such that

$$a_n \rightarrow 0 \text{ as } n \rightarrow \infty \quad (3)$$

and

$$na_n^4 \rightarrow \infty \text{ as } n \rightarrow \infty \quad \text{and} \quad na_n^5 \rightarrow 0 \text{ as } n \rightarrow \infty \quad (4)$$

where $\hat{F}_n$ is the empirical distribution function of $X_1, \dots, X_n$.

With K and a_n as defined above, we also consider the estimate

$$m_n^*(x_0) = \frac{\sum_{i=1}^n Y_i K((x_0 - X_i)/a_n)}{\sum_{i=1}^n K((x_0 - X_i)/a_n)} \quad (5)$$

When the sequence $\{\mathbf{Z}_i\}$ is stationary, several authors have proved different kinds of convergence of $m_n(x_0)$ or $m_n^*(x)$ to $m(x_0)$ and also the central limit theorem for $m_n(x_0)$ a.s. in x .

The central limit theorem was proved by Stute (1984) when the random variables are independent and by Yoshihara (1988) when the random variables are ϕ -mixing. Our aim is to extend these results when the random variables are absolutely regular as well as nonstationary which gives a larger class of applications. These results are very useful for geometrically ergodic Markov processes or ARMA processes for which the initial measure is not necessary to be the invariant measure. The regression function is of course invariant and under suitable conditions, any initial measure converges to the invariant measure. We estimate the regression function by the estimators $\tilde{m}_n$ and m_n^* defined below in (2) and (5). Their applications are detailed in section 3.

We make the following assumptions:

K is a nonnegative bounded integrable function on $\mathbb{R}$ whose radial majorant ψ is integrable.

$$\psi(x) = \sup_{|y| \geq |x|} K(y), \quad \int \psi(x) dx < \infty \quad (6)$$

and

$$\sup_{i \geq 1} E(|Y_i|^{5+\delta}) < \infty \quad \text{for some } \delta > 0 \quad (7)$$

or

$$\sup_{i \geq 1} |Y_i| \leq c < \infty. \quad (8)$$

The main results which allow such a generalization are given in the lemmas of the Appendix.

For the particular case where $Y_i \equiv X_{i+1}$, the estimators $\tilde{m}_n$ and m_n^* have the following expressions

$$\tilde{m}_n(x_0) = (na_n)^{-1} \sum_{i=1}^n X_{i+1} K \left(\frac{\tilde{F}_n(x_0) - \tilde{F}_n(X_i)}{a_n} \right) \quad (9)$$

$$m_n^*(x_0) = \frac{\sum_{i=1}^n X_{i+1} K((x_0 - X_i)/a_n)}{\sum_{i=1}^n K((x_0 - X_i)/a_n)} \quad (10)$$

2. ASYMPTOTICALLY UNBIASED ESTIMATOR OF $m(x_0)$.

Let $K(x)$ be a twice continuously differentiable kernel function such that

$$K(u) = 0 \text{ for } |u| \geq 1, \quad \int K(u) du = 1 \text{ and } \int uK(u) du = 0. \quad (11)$$

Let $H_{i,j}$, $F_{i,j}$ and $G_{i,j}$ be the d.f.'s respectively of $(\mathbf{Z}_i, \mathbf{Z}_j)$, (X_i, X_j) and (Y_i, Y_j) which admit respectively the densities $h_{i,j}$, $f_{i,j}$ and $g_{i,j}$.

We say that a sequence of d.f.'s $\{H_\ell^*\}$ on $\mathbb{R}^4$ admits F_ℓ^* , G_ℓ^* and H as marginals on $\mathbb{R}^2$ and F and G on $\mathbb{R}$ if we have

$$\lim_{\substack{y \rightarrow +\infty \\ v \rightarrow +\infty}} H_\ell^*(x, y, u, v) = F_\ell^*(x, u)$$

$$\lim_{\substack{x \rightarrow +\infty \\ u \rightarrow +\infty}} H_\ell^*(x, y, u, v) = G_\ell^*(y, v)$$

$$\lim_{\substack{u \rightarrow +\infty \\ v \rightarrow +\infty}} H_\ell^*(x, y, u, v) = H(x, y)$$

$$\lim_{\substack{x \rightarrow +\infty \\ y \rightarrow +\infty}} H_\ell^*(x, y, u, v) = H(u, v)$$

and the marginals of H are F and G .

THEOREM 12. Suppose the sequence $\{Z_i\}$ is absolutely regular with rates satisfying

$$\beta(m) = O(\rho^m) \quad \text{for some } 0 < \rho < 1 \quad (13)$$

Furthermore, assume that for any $\ell > 1$, there exists a continuous d.f. H_ℓ^* on $\mathbb{R}^4$ with marginals F_ℓ^*, G_ℓ^* and H on $\mathbb{R}^2$ and F and G on $\mathbb{R}$ which admits densities respectively f_ℓ^*, g_ℓ^*, h, f and g

$$\|H_{i,j} - H_{j-i}^*\| = O(\rho_0^i) \quad 1 \leq i < j \leq n, n \geq 1 \quad (14)$$

for some $0 < \rho_0 < 1$ where $\|\cdot\|$ denotes the norm of total variation. Then we have

$$m_n^*(x) \xrightarrow{a.s.} m(x) \quad \text{as } n \rightarrow \infty \quad (15)$$

if

$$\sum_{n=1}^{\infty} n^{1-\gamma} \exp(-n^\gamma a_n) < \infty \quad \text{for some } \gamma \ (0 < \gamma < 1) \quad (16)$$

and if (3) and (8) hold, where $\xrightarrow{a.s.}$ means convergence with probability one.

$$m_n^*(x) \xrightarrow{p} m(x) \quad \text{as } n \rightarrow \infty \quad (17)$$

if (3), (4) and (7) hold, where $\xrightarrow{p}$ means convergence in probability.

$$\tilde{m}_n(x) \xrightarrow{p} m(x) \quad \text{as } n \rightarrow \infty \quad (18)$$

if (3), (4) and (7) hold.

PROOF. Let us denote

$$Z_n(x) = (na_n)^{-1} \sum_{i=1}^{i=n} Y_i K((x - X_i)/a_n) \quad (19)$$

and

$$W_n(x) = (na_n)^{-1} \sum_{i=1}^n K((x - X_i)/a_n). \quad (20)$$

If we show that under the conditions of Theorem 12, $Z_n(x) \xrightarrow{a.s.} m(x)f(x)$ (resp. $Z_n(x) \xrightarrow{p} m(x)f(x)$) for almost all x , then we have also shown that

$W_n(x) \xrightarrow{a.s.} f(x)$ ($W_n(x) \xrightarrow{p} f(x)$). Since $m_n^*(x) = \frac{Z_n(x)}{W_n(x)}$ and since for almost all x , $f(x) > 0$, we can conclude that $m_n^*(x) \xrightarrow{a.s.} m(x)$ (resp. $m_n^*(x) \xrightarrow{p} m(x)$) for almost all x .

We can split the proof in two parts in view of the fact that

$$|Z_n(x) - m(x)f(x)| \leq |Z_n(x) - E(Z_n(x))| + |E(Z_n(x)) - m(x)f(x)| \quad (21)$$

First we define the following functions:

$$\begin{aligned} \phi_i(x, s) &= E(s^{-1} Y_i K((X_i - x)/s)) \quad i \geq 1 \\ \phi(x, s) &= E(s^{-1} Y K((X - x)/s)). \end{aligned}$$

Then, we have

$$\phi_i(x, s) = E(s^{-1} m_i(X_i) K((X_i - x)/s)) = \int s^{-1} K((y - x)/s) m_i(y) f_i(y) dy$$

where $m_i(X_i) = E(Y_i | X_i)$, and

$$\phi(x, s) = \int s^{-1} K((y - x)/s) m(y) dy.$$

By definition, we have

$$\begin{aligned} |E(Z_n(x)) - m(x)f(x)| &\leq |n^{-1} \sum_{i=1}^n \phi_i(x, a_n) \\ &\quad - \phi(x, a_n)| + |\phi(x, a_n) - m(x)f(x)| \\ &= |n^{-1} \sum_{i=1}^n \int K(u) m_i(u + xa_n) f_i(u + xa_n) du \\ &\quad - \int K(u) m(u + xa_n) f(u + xa_n) du| \\ &\quad + |\int K(u) m(u + xa_n) f(u + xa_n) du - m(x)f(x)| \\ &= I_1 + I_2. \end{aligned} \quad (22)$$

It follows that

$$\begin{aligned} I_1 &= |n^{-1} \sum_{i=1}^n \int K(u) \int y p_i(u + xa_n, y) dy f_i(u + xa_n) du \\ &\quad - \int K(u) \int y p(u + xa_n, y) dy f(u + xa_n) du| \end{aligned} \quad (23)$$

where $p_i(x, y) = \frac{h_i(x, y)}{f_i(x)}$, $p(x, y) = \frac{h(x, y)}{f(x)}$, where h_i is the density function of H_i .

$$I_1 = |n^{-1} \sum_{i=1}^n \int \int K(u) y h_i(u + x a_n, y) du dy - \int \int K(u) y h(u + x a_n, y) du dy|. \quad (24)$$

From condition (7), for any $\epsilon > 0$, we can find an $M > 0$ sufficiently large such that for any $v \in \mathbb{R}$, we have

$$\sup_{i \geq 1} \int I_{[|y| \geq M]} y p_i(v, y) dy < \epsilon/3 \quad \text{and} \quad \int I_{[|y| \geq M]} y p(v, y) dy < \epsilon/3. \quad (25)$$

From (24) and (25), it follows that

$$\begin{aligned} I_1 &\leq 2\epsilon/3 + |n^{-1} \sum_{i=1}^n \int \int K(u) y I_{[|y| \leq M]} h_i(u + x a_n, y) du dy \\ &\quad - \int \int K(u) y I_{[|y| \leq M]} h(u + x a_n, y) du dy| \end{aligned} \quad (26)$$

and from condition (14), there exists an N_0 sufficiently large such that $\forall n \geq N_0$

$$\begin{aligned} n^{-1} \sum_{i=1}^n \int \int K(u) y I_{[|y| \leq M]} h_i(u + x a_n, y) du dy \\ - \int \int K(u) y I_{[|y| \leq M]} h(u + x a_n, y) du dy < \epsilon/3 \end{aligned}$$

which implies

$$I_1 \leq \epsilon/2. \quad (27)$$

Then using Stein (1970, Theorem 2, pp. 62-63) and proceeding as in Devroye and Wagner (1980), we deduce from (11) and (6) that $\phi(x, s) \rightarrow m(x)f(x)$ as $s \rightarrow 0$ for almost all x . This implies that for any $\epsilon > 0$ that there exists an N_1 sufficiently large such that $\forall n \geq N_1$, we have

$$I_2 < \epsilon/2. \quad (28)$$

From (27) and (28), we find that $\forall n \geq N_0 \vee N_1$

$$|E(Z_n(x)) - m(x)f(x)| < \epsilon \quad \text{for almost all } x. \quad (29)$$

This proves (15), (resp. (17)) if we can show that for almost all x $Z_n(x) \xrightarrow{a.s.} E(Z_n(x))$ with conditions (3), (8) and (16) (resp. $Z_n(x) \xrightarrow{a.s.} E(Z_n(x))$ with conditions (3), (4) and (7)).

Clearly

$$Z_n(x) - E(Z_n(x)) = n^{-1} \sum_{i=1}^n \{T_i(a_n) - E(T_i(a_n))\} \quad (30)$$

where

$$T_i(s) = s^{-1} Y_i K((x - X_i)/s).$$

If $b = \sup_{u \in \mathbb{R}} K(u)$ and $c' = \sup_{i \geq 1} \sup_{n \geq 1} \phi_i(x, a_n)$ (c' is finite because of condition (8)), then we notice that

$$|T_i(a_n)| \leq bc/a_n \quad (31)$$

$$E(T_i^2(a_n)) \leq bcc'/a_n \quad (32)$$

If $U_1, U_2, \dots, U_n$ are independent random variables with $|U_i| \leq m$, $E(U_i) = 0$ and $E(U_i^2) \leq \sigma_i^2$, then an inequality due to Bennett (1962, p. 39) states that

$$P[|n^{-1} \sum_{i=1}^n U_i| \geq \epsilon] \leq 2 \exp\{-n\epsilon^2/2(\sigma^2 + 2m\epsilon)\} \quad (33)$$

where $\sigma^2 = n^{-1} \sum_{i=1}^n \sigma_i^2$.

Put $k = k_n = [n^{1-\gamma}] + 1$, and write

$$Z_n(x) - E(Z_n(x)) = \sum_{j=1}^k V_{n,j}(x) \quad (34)$$

where

$$V_{n,j}(x) = \sum_{p=0}^{\ell_j} T_{j+pk}(a_n) - E(T_{j+pk}(a_n)) \quad (35)$$

and ℓ_j is the largest integer such that $j + \ell_j k \leq n$. Then

$$\begin{aligned} P[|Z_n(x) - E(Z_n(x))| \geq \epsilon] &\leq P[n^{-1} \sum_{j=1}^k |V_{n,j}(x)| \geq \epsilon] \\ &\leq \sum_{j=1}^k P[|V_{n,j}(x)| \geq \epsilon n k^{-1}]. \end{aligned} \quad (36)$$

For any j , $1 \leq j \leq k$, define

$$B_j = \left\{ (y_1, \dots, y_{\ell_j}) : \left| \sum_{p=1}^{\ell_j} y_p \right| \geq \epsilon n k^{-1} \right\}$$

and put

$$g(y_1, \dots, y_{\ell_j}) = \begin{cases} 1 & \text{if } (y_1, \dots, y_{\ell_j}) \in B_j \\ 0 & \text{otherwise.} \end{cases}$$

By using Lemma 2.1 in Harel and Puri (1989) and (33)

$$\begin{aligned} P[|V_{in,j}(x)| \geq \epsilon n] &= E g(T_j(a_n), \dots, T_{j+\ell_j k}(a_n)) \\ &\leq P\left[\sum_{p=0}^{\ell_j} \tilde{T}_{j+pk}(a_n) \geq n\epsilon k^{-1}\right] + 2\ell_j \beta(k) \\ &\leq 2 \exp\{-\ell_j \epsilon^2 / (2bcc' / a_n + 4bce / a_n)\} + 2\ell_j \beta(k) \\ &\leq 2 \exp\{-\alpha \ell_j a_n\} + 2\ell_j \beta(k) \quad (\alpha = \epsilon^2 / (2bcc' + 4bce)) \end{aligned} \quad (37)$$

where $\tilde{T}_1(a_n), \dots, \tilde{T}_n(a_n)$ are independent and $\tilde{T}_i(a_n)$ is distributed as $T_i(a_n)$.

From (36) and (37), we deduce

$$P[|Z_n(x) - E(Z_n(x))| \geq \epsilon] \leq 2n^{1-\gamma} [\exp\{-\alpha \ell_j a_n\} + n^\gamma \beta(n^{1-\gamma}) + 1]. \quad (38)$$

From the Borel–Cantelli Lemma, conditions (4) and (16) imply that $Z_n(x) \xrightarrow{w.p.1} E(Z_n(x))$ for almost all x , and (15) is proved.

Now we prove (17). From Lemma 64 (see Section 4, Auxiliary Results), we obtain

$$\begin{aligned} E[Z_n(x) - E(Z_n(x))]^2 &= E\left[n^{-1} \sum_{i=1}^n \{T_i(a_n) - E(T_i(a_n))\}\right]^2 \\ &\leq n^{-1} \left\{ \sup_{i \leq n} (E|T_i(a_n)|^{2+\delta})^{2/(2+\delta)} \right\} \sum_{i=1}^n (\alpha(i))^{\delta/(2+\delta)}. \end{aligned} \quad (39)$$

From condition (7), we have

$$\sup_{n \in \mathbb{N}} \sup_{i \leq n} (E|T_i(a_n)|^{2+\delta})^{2/(2+\delta)} \leq M_0 < \infty$$

and from condition (13), we have

$$\sum_{i=1} (\alpha(i))^{\delta/(2+\delta)} = M_1 < \infty.$$

It follows that

$$E[Z_n(x) - E(Z_n(x))]^2 \leq M_0 M_1 n^{-1}.$$

By using the Markov inequality, we deduce that for any $\epsilon > 0$

$$P[|Z_n(x) - E(Z_n(x))| > \epsilon] \leq \epsilon^{-2} n^{-1} M_0 M_1 \quad (40)$$

which implies that $Z_n(x) \xrightarrow{P} E(Z_n(x))$ for almost all x and (17) is proved.

It remains to prove (18). It is easy to see that (17) is also true for a more general problem: estimation of a regression $E(\tilde{Y}|\tilde{X})$ from $\{(\tilde{X}_{ni}, \tilde{Y}_{ni}), 1 \leq i \leq n, n \geq 1\}$ when these pairs satisfy similar conditions of dependence, integration and convergence as in Theorem 12.

Denote by F_n the d.f. defined by $F_n = n^{-1} \sum_{i=1}^n F_i$. From the definition of $\tilde{m}_n(x_0)$ in (9) we can write

$$\begin{aligned} \tilde{m}_n(x_0) &= (na_n)^{-1} \sum_{i=1}^n Y_i K((F_n(x_0) - F_n(X_i))/a_n) \\ &\quad + a_n^{-2} n^{-1} \sum_{i=1}^n Y_i \{\tilde{F}_n(x_0) - \tilde{F}_n(X_i) - \bar{F}_n(x_0) + \bar{F}_n(X_i)\} K'(\mu_n) \\ &= J_1 + J_2 \end{aligned} \quad (41)$$

where μ_n is between $((F_n(x_0) - F_n(X_i))/a_n)$ and $((\tilde{F}_n(x_0) - \tilde{F}_n(X_i))/a_n)$. From the boundedness of the empirical process $n^{\frac{1}{2}}(\tilde{F}_n - F_n)$ and the condition $a_n^{-2} n^{-\frac{1}{2}} \rightarrow 0$, it follows that

$$a_n^{-2} \left\{ \tilde{F}_n(x_0) - \tilde{F}_n(X_i) - F_n(x_0) + F_n(X_i) \right\} \xrightarrow{P} 0. \quad (42)$$

From conditions (7) and (13), we deduce similarly as in the proof of (40) that

$$\frac{1}{n} \sum_{i=1}^n Y_i \xrightarrow{P} C \quad (43)$$

where C is a finite constant.

Hence $J_2 \xrightarrow{p} 0$ follows from (42), (43) and the fact that K' is bounded.

From (17), we deduce that

$$J_1 \xrightarrow{p} m(x) \quad (44)$$

by putting $\tilde{X}_{ni} = F_n(X_i)$ and $\tilde{Y}_{ni} = Y_i$. From (41)–(44), we deduce (18) of Theorem 12 and the proof is completed.

3. WEAK CONVERGENCE OF THE NEAREST NEIGHBOR REGRESSION FUNCTION ESTIMATES.

3.1 Main Theorem.

For any sequence of d.f.'s $\{\tilde{F}_\ell^*\}$ on $\mathbb{R}^1$ with marginals F_ℓ^* , G_ℓ^* and H on $\mathbb{R}^2$ and F and G on $\mathbb{R}$, we denote

$$\sigma_0^2 = \text{Var}(Y|X = x_0) \int_{-1}^1 K^2(u) du \quad (45)$$

where (X, Y) is a random vector which admits H as distribution function.

THEOREM 46. Suppose that the sequence $\{\mathbf{Z}_i\}$ is absolutely regular with rates (13). Furthermore, assume that for any $\ell > 1$, there exists a continuous d.f. H_ℓ^* on $\mathbb{R}^1$ admitting a density h_ℓ^* , marginals F_ℓ^* , G_ℓ^* and H on $\mathbb{R}^2$ and F and G on $\mathbb{R}$ such that (14) and (45) exist for which there exists a sequence $\tilde{\mathbf{Z}}_i = \{(\tilde{X}_i, \tilde{Y}_i)\}$ of stationary random variables such that $\{\tilde{\mathbf{Z}}_i\}$ is absolutely regular with rates (13) and $(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{Z}}_j)$ has H_{j-i}^* as d.f. ($i < j+1$) such that $m \circ F^{-1}$ is twice continuously differentiable in a neighborhood of x_0 and f (the density of F) is continuous in a neighborhood of x_0 ($0 < F(x_0) < 1$). Suppose also that the conditions (3), (4), (6) and (7) are satisfied. Then, $(na_n)^{\frac{1}{2}} \{\hat{m}_n(x_0) - m(x_0)\}$ converges to the normal law with mean 0 and variance σ_0^2 for almost all $x_0 \in \mathbb{R}$ such that $f(x_0) > 0$.

Remark 47. In Yoshihara (1988) the condition (4) was replaced by the weaker condition $na_n^3 \rightarrow \infty$ and $na_n^5 \rightarrow 0$, but we have to strengthen his condition to obtain the generalization to the absolutely regular and non-stationary case (see proof of Lemma 78).

3.2 Applications.

Consider a sequence $\{X_i\}_{i \geq 1}$ of real random variables which is a Markov process with transition probability $P(x; A)$ where $x \in \mathbb{R}$ and $A \in \mathcal{B}$, $\mathcal{B}$ is the Borel σ -field of $\mathbb{R}$.

Recall that the Markov process is geometrically ergodic if it is ergodic and if there exists $0 < \rho < 1$ such that

$$\|P^n(x; \cdot) - \mu(\cdot)\| = O(\rho^n) \quad \text{for all a.s. } x \in \mathbb{R} \quad (48)$$

where μ is the invariant measure and P^n the n -step transition probability.

We say that the process $\{X_i\}_{i \geq 1}$ has ν for the initial measure if the law of probability of X_1 is defined by ν and for any $i > 1$, the law of probability P_i of X_i is defined by νP^{i-1} . For any probability measure ν and any transition probability Q we denote by $Q \otimes \nu$ the probability measure defined on $\mathbb{R}^2$ by

$$Q \otimes \nu(A \times B) = \int_B Q(x; A) \nu(dx) \quad \text{for any } A \times B \in \mathcal{B} \times \mathcal{B}. \quad (49)$$

The Markov process is called strongly aperiodic if for any $x \in \mathbb{R}$, the transition probability $P(x; \cdot)$ is equivalent to the Lebesgue measure. The Markov process is called Harris recurrent if there exists a σ -finite measure ν on $\mathbb{R}$ with $\nu(\mathbb{R}) > 0$ such that $\nu(A) > 0$ implies $P(x; X_i \in A \text{ i.o.}) = 1$ for all $x \in \mathbb{R}$.

THEOREM 50. Let $\{X_i\}_{i \geq 1}$ be a Markov process which is strongly aperiodic, Harris recurrent and geometrically ergodic. Suppose that the invariant measure μ has a continuous density f and the transition probability P has a continuous density $h(\cdot; \cdot)$ and furthermore

$$\int |x|^{5+\delta} \mu(dx) < +\infty \quad (51)$$

Then for any initial measure, the conclusions of Theorems 12 and 46 hold for the nonparametric estimators $\tilde{m}_n$ and m_n^* defined in (9) and (10).

Remark 52. This theorem implies that for any initial measure admitting a strictly positive density, we can estimate the unique regression function $m(x) = \int y P(x; dy)$ associated by $\tilde{m}_n(x)$ or $m_n^*(x)$ and also apply the

central limit theorem to $\tilde{m}_n(x)$ to study the confidence region based on this statistic.

Proof. From Theorems 12 and 46, we have only to prove (13) and (14). First, we prove (13). From Davydov (1973), and the condition of strong aperiodicity, we have

$$\begin{aligned}\beta(m) &= \sup_n \int P_n(dx) \|P^m(x; \cdot) - P_{m+n}(\cdot)\| \\ &\leq \sup_n \int P_n(dx) \|P^m(x; \cdot) - \mu(\cdot)\| + \|P_{m+n}(\cdot) - \mu(\cdot)\|.\end{aligned}$$

As the process is geometrically ergodic, we can find $0 < \rho < 1$ such that

$$\|P^m(x; \cdot) - \mu(\cdot)\| = O(\rho^m) \quad \text{for all a.s. } x \in \mathbb{R}. \quad (53)$$

From Theorem 12 of Nummelin and Tuominen (1982), we deduce

$$\beta(m) = O(\rho^m) \quad (54)$$

which is (13).

Now, we prove (14). We have from (53)

$$\begin{aligned}\|P^m \otimes P_n - P^m \otimes \mu\| &= 2 \sup_{A \times B \in \mathcal{B} \times \mathcal{B}} \left| \int_B P^m(x; A) P_n(dx) \right. \\ &\quad \left. - \int_B P^m(x; A) \mu(dx) \right| \\ &\leq 2 \sup_{A \times B \in \mathcal{B} \times \mathcal{B}} P^m(x; A) \left| \int_B P_n(dx) - \mu(dx) \right| \\ &\leq 2 \|P_n - \mu\| = O(\rho^n)\end{aligned}$$

that is

$$\|P^m \otimes P_n - P^m \otimes \mu\| = O(\rho^n). \quad (55)$$

Thus the conclusions of Theorems 12 and 46 hold.

Example 56. We consider an ARMA process

$$X_i = aX_{i-1} + b\varepsilon_i + \varepsilon_{i-1}, \quad i \in \mathbb{N}^* \quad (57)$$

where X_0 admits a strictly positive density, $\{\varepsilon_i\}$ is a sequence of i.i.d. real valued random variables with strictly positive density such that $E(\varepsilon_i) = 0$, and a and b are real numbers such that $|a| < 1$. Then the conditions of Theorem 50 are satisfied for the process defined in (57), because we have here a particular case of a Markov process satisfying our conditions. The law of the process on which observations are taken is defined by the initial measure (i.e. the measure which defines the law of X_0) and the transitional measure (defined from the formula (61)). We can estimate the regression function by the estimators $\tilde{m}_n$ and m_n^* for any initial measure of X_0 which admits a strictly positive density.

THEOREM 58. Let $\{X_i\}_{i \geq 1}$ be a Markov process which is aperiodic and Doeblin recurrent. Suppose that the invariant measure μ has a continuous density function f and the transition probability P has a continuous density $h(\cdot; \cdot)$ and furthermore (51) is satisfied. Then for any initial measure, the conclusions of Theorems 12 and 46 hold for the non-parametric estimators $\tilde{m}_n$ and m_n^* defined in (9) and (10).

Proof. From Theorem (4.1) of Davydov (1973) the process $\{X_i\}$ is geometrically φ mixing which implies geometrically absolute regularity. The proof is now similar to Theorem 50.

Example 59. We consider the process

$$X_i = f(X_{i-1}) + \varepsilon_{i-1}, \quad i \in \mathbb{N}^* \quad (60)$$

where X_0 admits a strictly positive density and $\{\varepsilon_i\}_{i \geq 1}$ is a sequence of i.i.d. real random variables with strictly positive density such that $E(\varepsilon_i) = 0$ and f is a bounded function. That is a particular case of Markov process satisfying the conditions of Theorem 58.

Now more generally, consider the model

$$Y_i = \psi(X_i) + e_i, \quad i \geq 1 \quad (61)$$

where X_i denotes a random variable of observed values, ψ is a continuous known function such that (7) is satisfied, e_i is a white noise of identically absolutely regular random variables with a geometrical rate and strictly

positive density and Y_i is a predictor variable. If the sequence $\{X_i\}_{i \geq 1}$ of random variables satisfies the conditions of Theorems 50 or 58, then for any initial measure of $\{(X_i, Y_i)\}_{i \geq 1}$, the conclusions of Theorems 12 and 46 hold for the nonparametric estimators $\tilde{m}_n$ and m_n^* defined in (2) and (5) and particularly when the process $\{X_i\}$ is a process of the Example 56 or the Example 59 defined in (57) or (60).

4. AUXILIARY RESULTS.

Let $\{X_{ni}, 1 \leq i \leq n, n \geq 1\}$ be a nonstationary sequence of r.v.'s. Let $p \geq 1$ and $1 \leq i_1 < i_2 < \dots < i_p \leq n$ be arbitrary integers. For any j ($1 \leq j \leq p-1$), $P_{j,n}^{i_1, \dots, i_p}$ is the probability measure defined by

$$\begin{aligned} & P_{j,n}^{i_1, \dots, i_p}(A^{(j)} \times B^{(p-j)}) \\ &= P[(X_{ni_1}, \dots, X_{ni_j}) \in A^{(j)}] \times P[(X_{ni_{j+1}}, \dots, X_{ni_p}) \in B^{(p-j)}] \end{aligned}$$

and $P_{o,n}^{i_1, \dots, i_p}$ is the probability measure defined by

$$P_{o,n}^{i_1, \dots, i_p}(A^{(p)}) = P[X_{ni_1}, \dots, X_{ni_p} \in A^{(p)}]$$

for any $A^{(j)} \in \sigma(X_{ni_1}, \dots, X_{ni_j})$ ($1 \leq j \leq p$) and any $B^{(p-j)} \in \sigma(X_{ni_{j+1}}, \dots, X_{ni_p})$ ($1 \leq j \leq p-1$).

We state the following two lemmas, the proof of the first one follows essentially by proceeding as in Lemma 1 of Yoshihara (1976) and the second lemma is from Doukhan and Portal (1987, Proposition 2.8).

LEMMA 62. For every $p \geq 1$ and $(i_1, \dots, i_p)$ such that $i_1 < i_2 < \dots < i_p$ and any j ($0 \leq j \leq p-1$), let $h(x_1, \dots, x_p)$ be a Borel function such that

$$\int_{\mathbb{R}^p} |h(x_1, \dots, x_p)|^{1+\delta} dP_{j,n}^{(i_1, \dots, i_p)} \leq M$$

for some $\delta > 0$. Then

$$\begin{aligned} & \left| \int_{\mathbb{R}^p} h(x_1, \dots, x_p) dP_{o,n}^{(i_1, \dots, i_p)} - \int_{\mathbb{R}^p} h(x_1, \dots, x_p) dP_{j,n}^{(i_1, \dots, i_p)} \right| \\ & \leq 4M^{1/(2+\delta)} \beta^{\delta/(1+\delta)} (i_{j+1} - i_j). \end{aligned} \tag{63}$$

As a special case, if $h(x_1, \dots, x_p)$ is bounded, say, $|h(x_1, \dots, x_p)| \leq M$, then we can replace the right side of (63) by $2M\beta(i_{j+1} - i_j)$.

LEMMA 64. Let $\{X_{ni}\}$ be strong mixing. Let Z be $\sigma(X_{ni}, 1 \leq i \leq j)$ -measurable ($1 \leq j \leq n$) and V be $\sigma(X_{ni}, i \geq j + m)$ -measurable. If $E(|Z|^p) < \infty$, $E(|V|^q) < \infty$ and $r^{-1} + p^{-1} + q^{-1} = 1$ ($r, p, q > 0$) then

$$|E(ZV) - E(Z)E(V)| \leq M(\alpha(m))^{1/r} \{E|Z|^p\}^{1/p} \{E|V|^q\}^{1/q} \quad (65)$$

where M is some constant > 0 .

5. PROOF OF THEOREM 46.

From now on, the letter C , with or without subscript, will denote some positive quantity.

Writing $K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right)$ as

$$K \left\{ \left(\bar{F}_n(x_0) - \bar{F}_n(X_i) + \tilde{F}_n(x_0) - \tilde{F}_n(X_i) - \bar{F}_n(x_0) + \bar{F}_n(X_i) \right) / a_n \right\}$$

and using Taylor's expansion, we get

$$\begin{aligned} \tilde{m}_n(x_0) &= (na_n)^{-1} \sum_{i=1}^n Y_i K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \\ &\quad + a_n^{-2} n^{-1} \sum_{i=1}^n Y_i \{ \tilde{F}_n(x_0) - \tilde{F}_n(X_i) - \bar{F}_n(x_0) + \bar{F}_n(X_i) \} \\ &\quad \times K' \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \\ &\quad + a_n^{-3} n^{-1} \sum_{i=1}^n Y_i \{ \tilde{F}_n(x_0) - \tilde{F}_n(X_i) - \bar{F}_n(x_0) + \bar{F}_n(X_i) \} K''(\Delta_i) \\ &= I_1 + I_2 + I_3 \end{aligned} \quad (66)$$

where Δ_i is some random number between $a_n^{-1} \{ \tilde{F}_n(x_0) - \tilde{F}_n(X_i) \}$ and $a_n^{-1} \{ \bar{F}_n(x_0) - \bar{F}_n(X_i) \}$.

We need the following lemmas.

LEMMA 67. Under the conditions of Theorem 46,

$$(na_n)^{\frac{1}{2}} I_3 \xrightarrow{p} 0 \text{ as } n \rightarrow \infty. \quad (68)$$

PROOF. Since K vanishes outside $(-1, 1)$, the above expansion of $m_n(x_0)$ holds true with integration restricted to those x for which $|\tilde{F}_n(x_0) - \tilde{F}_n(x)| < a_n$. Let $b_n = (na_n^3)^{\frac{1}{3}}$. Then, by Lemmas 123 and 126

$$\begin{aligned} & P\left[\sup_{x: |\tilde{F}_n(x_0) - \tilde{F}_n(x)| \leq C_1 a_n} (na_n^3)^{\frac{1}{3}} |\tilde{F}_n(x_0) - \tilde{F}_n(x) - \bar{F}_n(x_0) + \bar{F}_n(x)| \geq b_n\right] \\ &= P\left[\sup_{x: |\tilde{F}_n(x_0) - \tilde{F}_n(x)| \leq C_1 a_n} |W_n(\tilde{F}_n(x_0)) - W_n(\tilde{F}_n(x))| \geq b_n a_n^{\frac{1}{2}}\right] \\ &\leq C b_n^{-4} = C(na_n^3)^{-2/3} \end{aligned} \quad (69)$$

where C is some constant > 0 .

On the other hand, as $\{Y_i\}$ is strong mixing with rates (13) and condition (7) is satisfied, we obtain

$$E\left\{n^{-1}\left(\sum_{i=1}^n Y_i - E(Y_i)\right)\right\}^2 \leq C_2 n^{-1} \quad (70)$$

where C_2 is some constant > 0 .

From the Markov inequality, we deduce that for any $\epsilon > 0$

which implies that $n^{-1} \sum_{i=1}^n Y_i - E(n^{-1} \sum_{i=1}^n Y_i) \xrightarrow{p} 0$ and we have $E(n^{-1} \sum_{i=1}^n Y_i) \rightarrow E(Y) < \infty$. Thus

$$n^{-1} \sum_{i=1}^n Y_i \xrightarrow{p} E(Y) < \infty. \quad (72)$$

Thus, for an arbitrary number $\epsilon > 0$, and any positive number η , we can find n sufficiently large such that

$$P(B_n(\eta)) \geq 1 - \epsilon$$

where

$$B_n(\eta) = \left\{ \omega, E(Y) - \eta \leq n^{-1} \sum_{i=1}^n Y_i \leq E(Y) + \eta \right\}.$$

From (69) and the fact that K'' is bounded, upon observing that $na_n^3 \rightarrow \infty$ we can find an integer n sufficiently large such that on the set $B_n(\eta)$ for any $\delta > 0$

$$(na_n)^{\frac{1}{2}} I_3 < \delta. \quad (73)$$

As δ is arbitrary, (68) follows.

Denote $m_i(x) = E(Y_i|X_i = x)$ and $\bar{m}_n(x) = n^{-1} \sum_{i=1}^n m_i(x)$, $x \in \mathbb{R}$.

LEMMA 74. Under the conditions of Theorem 46, $(na_n)^{\frac{1}{2}} I_2$ is asymptotically equivalent to

$$-(a_n^{-\frac{1}{2}} \bar{m}_n(x_0) n^{-\frac{1}{2}}) \times \left\{ \left(\sum_{i=1}^n K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \right) - \left(\sum_{i=1}^n E \left(K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \right) \right) \right\}. \quad (75)$$

PROOF. Define $\{Z_n, n \geq 1\}$ by

$$Z_n = n^{-1} a^{-\frac{3}{2}} \sum_{i=1}^n (Y_i - m_i(X_i)) (\tilde{G}_n(x_0) - \tilde{G}_n(X_i)) K' \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right)$$

where

$$\tilde{G}_n(x) = n^{-\frac{1}{2}} \sum_{i=1}^n (I_{[X_i \leq x]} - F_i(x)), \quad x \in \mathbb{R}. \quad (76)$$

Proceeding as in Lemma 64 of Yoshihara (1988), we prove that

$$Z_n \xrightarrow{p} 0 \quad \text{as } n \rightarrow \infty.$$

Now consider the function

$$k_n(x, y) = \bar{m}_n(x) K' \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(x)}{a_n} \right) (I_{[y \leq x_0]} - I_{[y \leq x]})$$

and the random variable

$$T_n = \int k_n(x, y) \tilde{G}_n(dx) \tilde{G}_n(dy).$$

Then, using the method of the proof of Lemma 2.2 in Harel and Puri (1989) and using (13), we obtain

$$E(T_n^2) = O(1) \quad \text{as } n \rightarrow \infty$$

which implies

$$a_n^{-\frac{3}{2}} \int k_n(x, y) \tilde{G}_n(dy) \{ \tilde{F}_n(dx) - F_n(dx) \} \xrightarrow{p} 0$$

since $na_n^3 \rightarrow \infty$.

Thus, we obtain that $(na_n)^{\frac{1}{2}} I_2$ is asymptotically equivalent to

$$a_n^{-3/2} \int m_n(x) (\tilde{G}_n(x_0) - \tilde{G}_n(x)) K' \left(\frac{F_n(x_0) - F_n(x)}{a_n} \right) \bar{F}_n(dx).$$

Now we prove that

$$\begin{aligned} I_1 &= a_n^{-3/2} \int |m_n(x) - m_n(x_0)| \left| \tilde{G}_n(x_0) - \tilde{G}_n(x) \right| \\ &\quad \times \left| K' \left(\frac{F_n(x_0) - F_n(x)}{a_n} \right) \right| \bar{F}_n(dx) \xrightarrow{p} 0 \end{aligned} \quad (77)$$

for a.s. $x_0 \in \mathcal{R}$.

From $F_n^{-1}(F_n(x_0)) = x_0$ for a.s. $x_0 \in \mathcal{R}$ and any $n \in \mathcal{N}^*$, we obtain

$$\lim_{n \rightarrow \infty} \int_{-1}^1 |m_n(F_n^{-1}(s) - ua_n) - m_n(F_n^{-1}(s))| |K'(u)| du = 0$$

for a.s. $s \in \mathcal{R}$. Put

$$e_n(x_0) = \left(a_n^{-1} \int_0^1 |m_n(F_n^{-1}(u)) - m_n(x_0)| \left| K' \left(\frac{F_n(x_0) - u}{a_n} \right) \right| du \right)^{-\frac{1}{6}}$$

Then, by Lemma 126, we have that for any $\varepsilon > 0$,

$$\begin{aligned} P(I_1 > \varepsilon) &\leq P \left(a_n^{-\frac{3}{2}} \sup_{|F_n(x_0) - u| \leq a_n} \left| \tilde{G}_n(F_n(x_0)) - \tilde{G}_n(u) \right| \right. \\ &\quad \times \left. \int_0^1 |m_n(F_n^{-1}(u)) - m_n(x_0)| \left| K' \left(\frac{F_n(x_0) - u}{a_n} \right) \right| du > \varepsilon \right) \\ &\leq P \left(\sup_{|F_n(x_0) - u| \leq a_n} \left| \tilde{G}_n(F_n(x_0)) - \tilde{G}_n(u) \right| \geq c_n(x_0) a_n^{\frac{1}{2}} \right) \\ &\leq C e_n^{-4}(x_0) = o(1). \end{aligned}$$

Hence (77) is proved and it follows that $(na_n)^{\frac{1}{2}}I_2$ is asymptotically equivalent to

$$\begin{aligned} & a_n^{-3/2}m(x_0) \int [\tilde{G}_n(x_0) - \tilde{G}_n(x)] K' \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(x)}{a_n} \right) \bar{F}_n(dx) \\ &= -a_n^{-3/2}m(x_0) \int [\tilde{G}_n(x)] K' \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(x)}{a_n} \right) \bar{F}_n(dx) \\ &= -a_n^{-1/2}m(x_0) \int K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(x)}{a_n} \right) \tilde{G}_n(dx) \end{aligned}$$

which leads to (75).

LEMMA 78. Under the conditions of Theorem 46

$$\begin{aligned} I_4 &= (a_n n)^{-\frac{1}{2}} \sum_{i=1}^n (Y_i - m_i(x_0)) K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \\ &\quad - (a_n n)^{-\frac{1}{2}} \sum_{i=1}^n E \{ (Y_i - m_i(x_0)) K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \} \quad (79) \end{aligned}$$

has asymptotically (as $n \rightarrow \infty$) the normal distribution with mean 0 and variance σ_0^2 (denoted as $\mathcal{N}(0, \sigma_0^2)$).

PROOF. Let

$$\begin{aligned} \tilde{B}_{ni} &= a_n^{-\frac{1}{2}} (\tilde{Y}_i - m(x_0)) K \left(\frac{F(x_0) - F(\tilde{X}_i)}{a_n} \right) \\ &\quad - a_n^{-\frac{1}{2}} E \left\{ (\tilde{Y}_i - m(x_0)) K \left(\frac{F(x_0) - F(\tilde{X}_i)}{a_n} \right) \right\} \quad i = 1, \dots, n, \end{aligned}$$

$$\begin{aligned} B_{ni}^* &= a_n^{-\frac{1}{2}} (Y_i - m_i(x_0)) K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \\ &\quad - a_n^{-\frac{1}{2}} E \left\{ (Y_i - m_i(x_0)) K \left(\frac{\bar{F}_n(x_0) - \bar{F}_n(X_i)}{a_n} \right) \right\}, \quad i = 1, \dots, n, \end{aligned}$$

and

$$\begin{aligned} B_{ni} &= a_n^{-\frac{1}{2}} (Y_i - m(x_0)) K \left(\frac{F(x_0) - F(X_i)}{a_n} \right) \\ &\quad - a_n^{-\frac{1}{2}} E \left\{ (Y_i - m(x_0)) K \left(\frac{F(x_0) - F(X_i)}{a_n} \right) \right\}, \quad i = 1, \dots, n. \end{aligned}$$

Then, we have:

$$I_4 = n^{-\frac{1}{2}} \sum_{i=1}^n B_{ni}^* = n^{-\frac{1}{2}} \sum_{i=1}^n B_{ni} + n^{-\frac{1}{2}} \sum_{i=1}^n (B_{ni}^* - B_{ni}).$$

From (14), it follows that

$$\sup_{x \in \mathbb{R}} |F(x) - \bar{F}_n(x)| = O(n^{-1}) \text{ which along with (7) implies that } E|B_{ni}^* - B_{ni}| = O(a_n^{-\frac{3}{2}} n^{-1})$$

or

$$E|n^{-\frac{1}{2}} \sum_{i=1}^n (B_{ni} - B_{ni}^*)| = O(a_n^{-\frac{3}{2}} n^{-\frac{1}{2}})$$

We deduce that

$$n^{-\frac{1}{2}} \sum_{i=1}^n (B_{ni} - B_{ni}^*) \xrightarrow{p} 0.$$

Now, we have to show that

$$n^{-\frac{1}{2}} \sum_{i=1}^n B_{ni} \text{ converges in law to } \mathcal{N}(0, \sigma_0^2).$$

Let us now denote

$$\begin{aligned} \varphi(x, y) &= (y - m(x_0))K\left(\frac{F(x_0) - F(x)}{a_n}\right) \\ &\quad - \int (z - m(x_0))K\left(\frac{F(x_0) - F(u)}{a_n}\right) H(du, dz). \end{aligned}$$

For any $M > 0$, denote also by

$$\begin{aligned} B_{ni}^M &= a_n^{-\frac{1}{2}} (Y_i I_{[|Y_i| \leq M]} - m(x_0))K\left(\frac{F(x_0) - F(X_i)}{a_n}\right) \\ &\quad - a_n^{-\frac{1}{2}} E\left\{ (Y_i I_{[|Y_i| \leq M]} - m(x_0))K\left(\frac{F(x_0) - F(X_i)}{a_n}\right) \right\} \end{aligned}$$

$$\bar{B}_{ni}^M = B_{ni} - B_{ni}^M$$

$$\begin{aligned} \varphi^M(x, y) &= (I_{[|y| \leq M]} y - m(x_0))K\left(\frac{F(x_0) - F(x)}{a_n}\right) \\ &\quad - \int (I_{[|z| \leq M]} z - m(x_0))K\left(\frac{F(x_0) - F(u)}{a_n}\right) H(du, dz) \end{aligned}$$

$$\bar{\varphi}^M(x, y) = g(x, y) - \varphi^M(x, y)$$

$$D_n^{M, \ell, p}(i) = 2 a_n^{-1} \int \varphi_\ell(x, y) \varphi_p(u, z) H_i^*(dx, dy, du, dz), \quad 0 \leq \ell, p \leq 1, i \geq 1$$

$$D_n^{M, \ell, p}(0) = \int \varphi_\ell(x, y) \varphi_p(x, y) H(dx, dy), \quad 0 \leq \ell, p \leq 1$$

$$\text{where } \varphi_0(x, y) = \varphi^M(x, y), \varphi_1(x, y) = \bar{\varphi}^M(x, y)$$

$$D_n(i) = \sum_{0 \leq \ell, p \leq 1} D_n^{M, \ell, p}(i), \quad i \geq 0.$$

We then have

$$\begin{aligned} & |n^{-1} \sum_{1 \leq i, j \leq n} E(B_{ni} B_{nj}) - n^{-1} \sum_{i=0}^n (n-i) D_n(i)| \\ &= |n^{-1} \sum_{1 \leq i, j \leq n} \{E(B_{ni}^M) + E(\bar{B}_{ni}^M)\} \{E(B_{nj}^M) + E(\bar{B}_{nj}^M)\} \\ &\quad - n^{-1} \sum_{0 \leq \ell, p \leq 1} \sum_{i=1}^n (n-i) D_n^{M, \ell, p}(i)|. \end{aligned}$$

From condition (7), we have

$$\sup_{n \geq 1} \max_{1 \leq i \leq n} E(\bar{B}_{ni}^M)^2 = O(M^{-2\delta}) \quad \text{and} \quad \int (\bar{g}^M(x, y))^2 H(dx, dy) = O(M^{-2\delta}). \quad (80)$$

We can write

$$\begin{aligned} & |n^{-1} \sum_{1 \leq i < j \leq n} E(B_{ni}^M B_{nj}^M) - n^{-1} \sum_{i=1}^n (n-i) D_n^{M, 0, 0}(i)| \\ &\leq n^{-1} \left| \sum_{1 \leq j-i \leq k} E(B_{ni}^M B_{nj}^M) - \sum_{i=1}^k (n-i) D_n^{M, 0, 0}(i) \right| \\ &\quad + n^{-1} \left| \sum_{k < j-i \leq n} E(B_{ni}^M B_{nj}^M) - \sum_{i=k+1}^n (n-i) D_n^{M, 0, 0}(i) \right| \\ &= I_n^{(1)} + J_n^{(1)}. \end{aligned}$$

From condition (14) and the fact that K' is bounded, if we take $M = a_n^{-1}$,

$$I_n^{(1)} = k O(a_n^{-3} n^{-1})$$

and from condition (13)

$$J_n^{(1)} = O(\rho_1^k) \quad \text{where} \quad \rho_1 = \rho^{\delta/(2+\delta)}.$$

If we take $k = a_n^{-1}$, then

$$I_n^{(1)} + J_n^{(1)} = O(a_n^{-4} n^{-1}) + O(a_n^{-1} \rho_1^{a_n^{-1}}) = O(a_n^{-4} n^{-1}). \quad (81)$$

We deduce similarly that

$$|n^{-1} \sum_{i=1}^n E(B_{ni}^M)^2 - D_n^{M,0,0}(0)| = O(n^{-1} a_n^{-4}). \quad (82)$$

We now write

$$\begin{aligned} & |n^{-1} \sum_{1 \leq i < j \leq n} E(B_{ni}^M B_{nj}^M) - n^{-1} \sum_{i=1}^n (n-i) D_n^{M,1,0}(i)| \\ & \leq |n^{-1} \sum_{1 \leq j-i \leq k} E(B_{ni}^M B_{nj}^M) - \sum_{i=1}^k (n-i) D_n^{M,1,0}(i)| \\ & \quad + |n^{-1} \sum_{k < j-i \leq n} E(B_{ni}^M B_{nj}^M) - \sum_{i=k+1}^n (n-i) D_n^{M,1,0}(i)| \\ & = I_n^{(2)} + J_n^{(2)}. \end{aligned} \quad (83)$$

From (80), we have

$$I_n^{(2)} = O(k a_n^{\delta-2}),$$

and from condition (13)

$$J_n^{(2)} = O(\rho_1^k).$$

If we take $k = a_n^{-1}$ and noting that $M = a_n^{-1}$, we have

$$I_n^{(2)} + J_n^{(2)} = O(a_n^{2\delta-4}) + O(a_n^{-1} \rho_1^{a_n^{-1}}) = O(a_n^{\delta-3}). \quad (84)$$

Similarly, we can show that

$$\left| n^{-1} \sum_{i=1}^n E(B_{ni}^M B_{ni}^M) - D_n^{M,1,0}(0) \right| = O(a_n^{\delta-3}), \quad (85)$$

$$\left| n^{-1} \sum_{1 \leq i, j \leq n} E(B_{ni}^M \bar{B}_{nj}^M) - n^{-1} \sum_{i=0}^n (n-i) D_n^{M,0,1}(i) \right| = O(a_n^{\delta-3}), \quad (86)$$

and

$$\left| n^{-1} \sum_{1 \leq i, j \leq n} E(\bar{B}_{ni}^M \bar{B}_{nj}^M) - n^{-1} \sum_{i=0}^n (n-i) D_n^{M,1,1}(i) \right| = O(a_n^{2\delta-2}). \quad (87)$$

We deduce from (81), (82), (84)—(87) that

$$\left| n^{-1} \sum_{1 \leq i, j \leq n} E(B_{ni} B_{nj}) - n^{-1} \sum_{i=0}^n (n-i) D_n(i) \right| = O(a_n^{\delta-3}).$$

Hence

$$\lim_{n \rightarrow \infty} \left| (n)^{-1} \sum_{1 \leq i, j \leq n} E(B_{ni} B_{nj}) - (n)^{-1} \sum_{i=0}^{n-1} (n-i) D_n(i) \right| = 0 \quad (88)$$

which implies that

$$\lim_{n \rightarrow \infty} \left| E \left(n^{-\frac{1}{2}} \sum_{i=1}^n B_{ni} \right)^2 - E \left(n^{-\frac{1}{2}} \sum_{i=1}^n \tilde{B}_{ni} \right)^2 \right| = 0.$$

Now, we prove that

$$\lim_{n \rightarrow \infty} E(n^{-\frac{1}{2}} \sum_{i=1}^n \tilde{B}_{ni})^2 = \sigma_0^2 \quad (89)$$

where σ_0^2 is defined by (45).

We have

$$\begin{aligned} E(\tilde{B}_{ni})^2 &= a_n^{-1} \int |y - m(x_0)|^2 K^2 \left(\frac{F(x_0) - F(x)}{a_n} \right) H(dx, dy) \\ &\quad - a_n^{-1} \left\{ \int (y - m(x_0)) K \left(\frac{F(x_0) - F(x)}{a_n} \right) H(dx, dy) \right\}^2 \\ &= a_n^{-1} \int \psi(x) K^2 \left(\frac{F(x_0) - F(x)}{a_n} \right) F(dx) \\ &\quad - a_n^{-1} \left\{ \int_0^1 \varphi(x) K \left(\frac{F(x_0) - F(x)}{a_n} \right) F(dx) \right\}^2 \\ &= \int_{-1}^{+1} \psi(F^{-1}(F(x_0) - a_n u)) K^2(u) du \\ &\quad - a_n \left\{ \int \varphi(F^{-1}(F(x_0) - a_n u)) K(u) du \right\}^2 \end{aligned} \quad (90)$$

where $\psi(x) = E\{|\tilde{Y}_1 - m(x_0)|^2 | \tilde{X}_1 = x\}$ and $\varphi(x) = E\{\tilde{Y}_1 - m(x_0) | \tilde{X}_1 = x\}$.

From (90), we easily deduce that

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n E(\tilde{B}_{ni})^2 = \sigma_0^2. \quad (91)$$

Now, we can write

$$\begin{aligned} E \left(n^{-\frac{1}{2}} \sum_{i=1}^n \tilde{B}_{ni} \right)^2 &= n^{-1} \sum_{i=1}^n E(\tilde{B}_{ni})^2 + n^{-1} \sum_{1 \leq i \neq j \leq n} E(\tilde{B}_{ni} \tilde{B}_{nj}) \quad (92) \\ &= n^{-1} \sum_{i=1}^n E(\tilde{B}_{ni})^2 + 2n^{-1} \sum_{\substack{1 \leq i < j \leq n \\ j-i \leq k}} E(\tilde{B}_{ni} \tilde{B}_{nj}) + 2n^{-1} \sum_{\substack{1 \leq i < j \leq n \\ j-i > k}} E(\tilde{B}_{ni} \tilde{B}_{nj}). \end{aligned}$$

Then, we have

$$\begin{aligned} E(\tilde{B}_{ni} \tilde{B}_{nj}) &= a_n^{-1} \int \varphi(x, y) \varphi(u, z) H_{j-i}^*(dx, dy, du, dz) \quad (93) \\ &\leq C a_n^{-1} \left\{ \int \varphi_{i,j}^*(x, u) K \left(\frac{F(x_0) - F(x)}{a_n} \right) K \left(\frac{F(x_0) - F(u)}{a_n} \right) F_{j-i}^*(dx, du) \right\} \\ &= C a_n \left\{ \int \varphi_{i,j}^*(F^{-1}(F(x_0) - a_n v), F^{-1}(F(x_0) - a_n w)) \times \right. \\ &\quad \left. \times \frac{f_{j-i}^*(F^{-1}(F(x_0) - a_n v), F^{-1}(F(x_0) - a_n w))}{f \circ F^{-1}(F(x_0) - a_n v) f \circ F^{-1}(F(x_0) - a_n w)} K(v) K(w) dv dw \right\} \\ &\leq C a_n. \end{aligned}$$

This follows from (7) and the fact that f is continuous in a neighborhood of x_0 with $f(x_0) > 0$, where

$$\varphi_{i,j}^*(x, u) = E \left\{ (\tilde{Y}_j - m(x_0))(\tilde{Y}_i - m(x_0)) | \tilde{X}_i = x, \tilde{X}_j = u \right\}$$

and f_{j-i}^* is the density of F_{j-i}^* .

For any $\delta > 0$,

$$E |\tilde{B}_{ni}|^{2+\delta} \leq C a_n^{-\frac{2+\delta}{2}} \int |y - m(x_0)|^{2+\delta} \left| K \left(\frac{F(x_0) - F(x)}{a_n} \right) \right|^{2+\delta} H(dx, dy)$$

$$\begin{aligned}
&= C a_n^{-\frac{2+\delta}{2}} \int \psi^*(x) \left| K \left(\frac{F(x_0) - F(x)}{a_n} \right) \right|^{2+\delta} F(dx) \\
&\leq C a_n^{-\frac{\delta}{2}} \int_{-1}^{+1} \psi^*(F^{-1}(F(x_0) - a_n u)) |K(u)|^{2+\delta} du \\
&\leq C a_n^{-\frac{\delta}{2}}
\end{aligned} \tag{94}$$

where $\psi^*(x) = E\{|Y_1 - m(x_0)|^{2+\delta} | X_1 = x\}$.

From (93) and (94), it follows that

$$E \left(n^{-1} \sum_{1 \leq i \neq j \leq n} \tilde{B}_{ni} \tilde{B}_{nj} \right) \leq 2C k a_n + 2 \left(C a_n^{-\frac{\delta}{2}} \right)^{2/(2+\delta)} \sum_{i=k+1}^n [\beta(i)]^{\delta/(2+\delta)}. \tag{95}$$

If we take $k = [a_n^{-\frac{1}{2}}]$, we get (89) from (13) and (91).

By using (94) and Lemma 129 which generalizes Lemma 3.4 of Yoshihara (1988) from the ϕ mixing case to the strong mixing case, we obtain the following inequality which generalizes Lemma 5.1 of Yoshihara (1988) to the strong mixing case.

$$E \left| \sum_{i=1}^m B_{ni} \right|^{2+\delta} \leq C m^{\frac{2+\delta}{2}} \sup_{1 \leq i \leq n} E|B_{ni}|^{2+\delta}, \quad m \leq n. \tag{96}$$

Now we can proceed as in Yoshihara (1988) (proof of Theorem 2). Let $p = [n^{2/3}]$, $q = [n^{1/3}]$ and $k = [n/(p+q)]$. Put

$$\eta_j = \sum_{i=(j-1)(p+q)+1}^{(j-1)(p+q)+p} B_{ni}, \quad j = 1, \dots, k.$$

$$\theta_j = \sum_{i=(j-1)(p+q)+p+1}^{(j-1)(p+q)+p} B_{ni}, \quad j = 1, \dots, k.$$

$$\theta_{k+1} = \sum_{i=k(p+q)+1}^n B_{ni}.$$

Then, we have

$$n^{-\frac{1}{2}} \sum_{i=1}^n B_{ni} = n^{-\frac{1}{2}} \sum_{j=1}^k \eta_j + n^{-\frac{1}{2}} \sum_{j=1}^{k+1} \theta_j.$$

From (96), we deduce that

$$n^{-\frac{1}{2}} \sum_{j=1}^{k+1} \theta_j \xrightarrow{p} 0.$$

To prove Lemma 78, it remains now to show that $n^{-\frac{1}{2}} \sum_{j=1}^k \eta_j$ converges in law to $\mathcal{N}(0, \sigma_0^2)$ random variable.

From Lemma 62 we obtain

$$\left| E \left\{ \exp(it n^{-\frac{1}{2}} \sum_{j=1}^k \eta_j) \right\} - \prod_{j=1}^k \left[E \left\{ \exp(it n^{-\frac{1}{2}} \eta_j) \right\} \right] \right| \leq Ck\beta(q)$$

where C is some constant > 0 . Hence it is enough to prove that

$$\prod_{j=1}^k \left[E \left\{ \exp(it n^{-\frac{1}{2}} \eta_j) \right\} \right] \quad \text{converges to}$$

Using (96), we obtain

From the fact that $na_n^3 \rightarrow \infty$ and

$$\lim_{n \rightarrow \infty} \frac{kE(\eta_j)^2}{n} = \sigma_0^2 \quad (\text{from (89)})$$

we get (97). The proof follows.

LEMMA 98. Under the conditions of Theorem 46

$$(na_n^{-1})^{\frac{1}{2}} \left[\left\{ n^{-1} \sum_{i=1}^n \int y K \left(\frac{F_n(x_0) - F_n(x)}{a_n} \right) H_i(dx, dy) \right\} - m(x_0) \right]$$

as $n \rightarrow 0$.

PROOF. We have

$$\begin{aligned}
 D_n &= (na_n^{-1})^{\frac{1}{2}} \left[n^{-1} \sum_{i=1}^n \int y K \left(\frac{F_n(x_0) - F_n(x)}{a_n} \right) H_i(dx, dy) \right. \\
 &\quad \left. - n^{-1} \sum_{i=1}^n \int y K \left(\frac{F(x_0) - F(x)}{a_n} \right) H_i(dx, dy) \right] \\
 &\quad + (na_n^{-1})^{\frac{1}{2}} \left[n^{-1} \sum_{i=1}^n \int y K \left(\frac{F(x_0) - F(x)}{a_n} \right) H_i(dx, dy) \right. \\
 &\quad \left. - n^{-1} \sum_{i=1}^n \int y K \left(\frac{F(x_0) - F(x)}{a_n} \right) H(dx, dy) \right] \\
 &\quad + (na_n^{-1})^{\frac{1}{2}} \left[n^{-1} \sum_{i=1}^n \int y K \left(\frac{F(x_0) - F(x)}{a_n} \right) H(dx, dy) - m(x_0) \right] \\
 &= H_n^{(1)} + H_n^{(2)} + H_n^{(3)}.
 \end{aligned} \tag{100}$$

From (7) and (14), it follows that

$$H_n^{(1)} = O(a_n^{-\frac{3}{2}} n^{-\frac{1}{2}}) \tag{101}$$

Denote

$$\begin{aligned}
 \ell(x, y) &= y K \left(\frac{F(x_0) - F(x)}{a_n} \right), \\
 \ell^{(M)}(x, y) &= y I_{[|y| \leq M]} K \left(\frac{F(x_0) - F(x)}{a_n} \right),
 \end{aligned}$$

and

$$\bar{\ell}^{(M)}(x, y) = \ell(x, y) - \ell^{(M)}(x, y) \quad \text{for any } M > 0.$$

We then have

$$\begin{aligned}
 H_n^{(2)} &= (na_n)^{\frac{1}{2}} \left[n^{-1} \sum_{i=1}^n \int \ell^{(M)}(x, y) (H_i(dx, dy) - H(dx, dy)) \right. \\
 &\quad \left. + n^{-1} \sum_{i=1}^n \left(\int \bar{\ell}^{(M)}(x, y) H_i(dx, dy) - \int \bar{\ell}^{(M)}(x, y) H(dx, dy) \right) \right].
 \end{aligned}$$

From (7) and (14), if we take $M = n^{\frac{1}{4}}$, we get

$$H_n^{(2)} = O(n^{-\frac{1}{4}} a_n^{-\frac{1}{2}}) + O\left(n^{-\frac{1}{8} - \frac{\delta}{4}} a_n^{-\frac{1}{2}}\right) \tag{102}$$

From Stute (1984, Corollary, p. 919), it follows that

$$H_n^{(3)} \longrightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (103)$$

(101)—(103) yield (99). The proof of the Theorem 46 now follows from Lemmas 67, 74, 78 and 98.

6. APPENDIX.

Let X_{ni} , $1 \leq i \leq n$, $n \geq 1$ be real-valued r.v.s (random variables) with continuous d.f.s (distribution functions) $F_{ni}(x)$, $x \in \mathbb{R}$ and let

$$\hat{G}_n(t) = \sum_{i=1}^n I_{[F_n(X_{ni}) \leq t]} - F_{ni}(\bar{F}_n^{-1}(t)), \quad t \in [0, 1] \quad (104)$$

where $\bar{F}_n = n^{-1} \sum_{i=1}^n F_{ni}$.

LEMMA 105. If the sequence $\{X_{ni}\}$ is strong mixing with rates satisfying

$$\sum_{i=0}^{+\infty} i^2 (\alpha(i))^{\delta/(2+\delta)} < \infty \quad \text{for some } \delta > 0, \quad (106)$$

$$E[|\hat{G}_n(t) - \hat{G}_n(s)|]^4 \leq C[n^2(t-s)^{\frac{4}{2+\delta}} + n(t-s)^{\frac{2}{2+\delta}}], \quad 0 \leq s \leq t \leq 1. \quad (107)$$

PROOF. Denote $A_{n,i} = I_{[s \leq F_n(X_{ni}) \leq t]} - \{F_{ni}(\bar{F}_n^{-1}(t)) - F_{ni}(\bar{F}_n^{-1}(s))\}$. We have

$$\begin{aligned} & E[|\hat{G}_n(t) - \hat{G}_n(s)|]^4 \\ & \leq \sum_{1 \leq i, j, k, \ell \leq n} |E(A_{n,i} A_{n,j} A_{n,k} A_{n,\ell})| \\ & \leq 4! \left(\sum_{i=0}^{n-1} \sum_{j=1}^{n-i} \sum_{k \leq i} \sum_{\ell \leq i} |E(A_{n,j} A_{n,j+i} A_{n,j+i+k} A_{n,j+i+k+\ell})| \right. \\ & \quad + \sum_{i=0}^{n-1} \sum_{j=i+1}^n \sum_{k \leq i} \sum_{\ell \leq i} |E(A_{n,j-i-k-\ell} A_{n,j-i-k} A_{n,j-i} A_{n,j})| \\ & \quad \left. + \sum_{i=0}^{n-1} \sum_{j=1}^{n-i} \sum_{k \leq i} \sum_{\ell \leq i} |E(A_{n,j-\ell} A_{n,j} A_{n,j+i} A_{n,j+i+k})| \right) \\ & = 4!(K_1 + K_2 + K_3). \end{aligned} \quad (108)$$

Denote by $B_{ni} = F_{ni}(F_n^{-1}(t)) - F_{ni}(F_n^{-1}(s))$.

Then, from Lemma 64, it follows that

$$\begin{aligned}
K_1 &\leq \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \sum_{j=1}^{n-i} (B_{nj})^{1/(2+\delta)} (B_{n,j+i})^{1/(2+\delta)} \\
&\leq \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \left(\sum_{j=1}^{n-i} B_{nj} \right)^{1/(2+\delta)} \left(\sum_{j=1}^{n-i} B_{n,j+i} \right)^{1/(2+\delta)} \\
&\leq \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \left(\sum_{j=1}^n B_{nj} \right)^{1/(2+\delta)} \left(\sum_{j=1}^n B_{nj} \right)^{1/(2+\delta)} \\
&\leq \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \left(\sum_{j=1}^n B_{nj} \right)^{2/(2+\delta)} \\
&= \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \left(n \left(n^{-1} \sum_{j=1}^n B_{nj} \right) \right)^{2/(2+\delta)} \\
&\leq \left(\sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \right) n(t-s)^{2/(2+\delta)} \\
&\leq C n(t-s)^{2/(2+\delta)}. \tag{109}
\end{aligned}$$

Similarly

$$K_2 \leq C n(t-s)^{2/(2+\delta)}. \tag{110}$$

From Lemma 64, we also obtain

$$\begin{aligned}
K_3 &\leq \sum_{i=0}^{n-1} \sum_{j=1}^{n-i} \sum_{k \leq i} \sum_{\ell \leq i} |E(B_{n,j-\ell} B_{n,j}) E(B_{n,j+i} B_{n,j+i+k})| \\
&\quad + \sum_{i=0}^{n-1} i^2 (\alpha(i))^{\delta/(2+\delta)} \sum_{j=1}^{n-i} (B_{nj})^{1/(2+\delta)} (B_{n,j+i})^{1/(2+\delta)} \\
&\leq \sum_{i=0}^{n-1} \sum_{j=1}^{n-i} |E(B_{n,j} B_{n,j+i})| \sum_{k=0}^{n-1} \sum_{\ell=1}^{n-k} |E(B_{n,\ell} B_{n,\ell+k})| + C n(t-s)^{2/(2+\delta)} \\
&\leq \left\{ \sum_{i=0}^{n-1} (\alpha(i))^{\delta/(2+\delta)} \sum_{j=1}^{n-i} (B_{nj})^{1/(2+\delta)} (B_{n,j+i})^{1/(2+\delta)} \right\}^2 + C n(t-s)^{2/(2+\delta)} \\
&\leq C \left\{ n^2 (t-s)^{4/(2+\delta)} + n(t-s)^{2/(2+\delta)} \right\}. \tag{111}
\end{aligned}$$

From (108)–(111) we obtain (107) and Lemma 105 is proved.

Denote $W_n(t) = n^{-\frac{1}{2}} \hat{G}_n(t)$, $0 \leq t \leq 1$ and let $F_{n,i,j}$ be the d.f. of (X_{ni}, X_{nj}) .

LEMMA 112. If the sequence $\{X_{ni}\}$ is strong mixing with rates (106), then $W_n(t)$ converges in law in Skorohod topology to a Gaussian process $W_0(t)$ with trajectories a.s. continuous with mean 0 and variance $\sigma^2(t)$. Furthermore, assume that for any $\ell > 1$, there exists a continuous d.f. F_ℓ^* on $\mathbb{R}^*$ with marginal F such that

$$\|F_{n,i,j} - F_{j-i}^*\| = O(\rho^i) \quad 1 \leq i \leq n, n \geq 1 \quad \text{for some } 0 < \rho_0 < 1 \quad (113)$$

$$\sigma^2(t) = t(1-t) + 2 \sum_{i=1}^{\infty} \{F_i^*(F^{-1}(t), F^{-1}(t)) - t^2\}. \quad (114)$$

PROOF. The process W_n defines a probability measure P_n on D_1 which is the space of functions on $[0, 1]$ that are right continuous and have left-hand limits. From 15.1 of Billingsley (1968) we have to prove that (i) the finite dimensional distribution of P_n converges in law to a normal law and (ii) P_n is tight.

First we prove (i) which is equivalent to proving that $\sum_{\ell=1}^p \lambda_\ell W_n(s_\ell)$ converges in law to a normal distribution for any $p \in \mathbb{N}^*$, for any $s_\ell \in [0, 1]$ and any $\lambda_\ell \in \mathbb{R}$ ($1 \leq \ell \leq p$). Without loss of generality, we can take $p = 2$ and suppose that $s_1 < s_2$.

We then have

$$\begin{aligned} \sum_{\ell=1}^2 \lambda_\ell W_n(s_\ell) &= n^{-\frac{1}{2}} \left[\sum_{i=1}^n \lambda_1 (I_{[F_n(X_{ni}) \leq s_1]} - F_{ni}(\hat{F}_n^{-1}(s_1))) \right. \\ &\quad \left. + \lambda_2 (I_{[F_n(X_{ni}) \leq s_2]} - F_{ni}(\hat{F}_n^{-1}(s_2))) \right]. \end{aligned}$$

From Corollary 1 of Withers (1975) and condition (106), we have only to prove that $E(\sum_{\ell=1}^2 \lambda_\ell W_n(s_\ell))^2$ converges to some constant C .

We define $C_{ni}(s_\ell)$ by $C_{ni}(s_\ell) = I_{\{F_n(X_{ni}) \leq s_\ell\}} - F_{ni}(F_n^{-1}(s_\ell))$, $\ell = 1, 2$.
Then

$$\begin{aligned}
 & E \left(\sum_{\ell=1}^2 \lambda_\ell W_n(s_\ell) \right)^2 \\
 &= n^{-1} \left\{ \lambda_1^2 \left[\sum_{i=1}^n \sum_{j=1}^n E(C_{ni}(s_1)C_{nj}(s_1)) \right] \right. \\
 &\quad \left. + 2\lambda_1\lambda_2 \left[\sum_{i=1}^n \sum_{j=1}^n E(C_{ni}(s_1)C_{nj}(s_2)) \right] + \lambda_2^2 \left[\sum_{i=1}^n \sum_{j=1}^n E(C_{ni}(s_2)C_{nj}(s_2)) \right] \right\} \\
 &= L_1 + L_2 + L_3. \tag{115}
 \end{aligned}$$

Let us now denote

$$\begin{aligned}
 C(1) &= s_1(1 - s_1) \\
 C(i) &= 2(F_i^*(F^{-1}(s_1), F^{-1}(s_1)) - s_1^2).
 \end{aligned}$$

Then, we have

$$\begin{aligned}
 & |n^{-1} \sum_{i=1}^n \sum_{j=1}^n E(C_{ni}(s_1)C_{nj}(s_1)) - \sum_{i=1}^{+\infty} C(i)| \\
 &\leq |n^{-1} \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} E(C_{nj}(s_1)C_{n,j+i}(s_1)) - n^{-1} \sum_{i=1}^n (n-i)C(i)| \\
 &\quad + \sum_{i=n+1}^{+\infty} |C(i)| + \sum_{i=1}^n \sum_{k=i}^{\infty} |C(k)| = |A_n| + B_n + C_n. \tag{116}
 \end{aligned}$$

From the inequality (65) we deduce $|C(i)| \leq (\alpha(i))^{\delta/(2+\delta)}$ which implies from condition (106) that $B_n \rightarrow 0$ and $C_n \rightarrow 0$ (as $n \rightarrow \infty$).

We now write

$$\begin{aligned}
 & |E(C_{nj}(s_1)C_{n,j+i}(s_1)) - (1/2)C(i)| \\
 &\leq |F_{n,j,j+i}(\bar{F}_n^{-1}(s_1), \bar{F}_n^{-1}(s_1)) - F_{nj}(\bar{F}_n^{-1}(s_1))F_{n,j+i}(\bar{F}_n^{-1}(s_1))| \\
 &\quad - F_{j-i}^*(\bar{F}_n^{-1}(s_1), \bar{F}_n^{-1}(s_1)) + (F(\bar{F}_n^{-1}(s_1))^2) + |F_{j-i}^*(\bar{F}_n^{-1}(s_1), \bar{F}_n^{-1}(s_1))| \\
 &\quad - F(\bar{F}_n^{-1}(s_1))^2 - F_{j-i}^*(F^{-1}(s_1), F^{-1}(s_1)) + s_1^2|. \tag{117}
 \end{aligned}$$

From conditions (106) and (113) and from (117), we easily deduce that $|A_n| \rightarrow 0$ as $n \rightarrow \infty$. Thus

$$\lambda_1^2(s_1(t - s_1)) + 2 \sum_{i=1}^{+\infty} \{F_i^*(F^{-1}(s_1), F^{-1}(s_1)) - s_1^2\}. \quad (118)$$

Similarly

$$L_2 \longrightarrow 2 \lambda_1 \lambda_2(s_1 \wedge s_2 - s_1 s_2) + 2 \sum_{i=1}^{+\infty} \{F_i^*(F^{-1}(s_1), F^{-1}(s_2)) - s_1 s_2\} \quad (119)$$

and

$$L_3 \longrightarrow \lambda_2^2(s_1(1 - s_2)) + 2 \sum_{i=1}^{+\infty} \{F_i^*(F^{-1}(s_2), F^{-1}(s_2)) - s_2^2\}. \quad (120)$$

From (115) and (118)–(120), we deduce that

$$E \left(\sum_{i=1}^2 \lambda_i W_n(s_i) \right)^2 \longrightarrow C \quad (\text{as } n \rightarrow \infty)$$

and (i) is proved.

Now, we prove (ii). From Theorem 15.5 of Billingsley (1968), we have to verify that $\forall \epsilon > 0, \exists \eta > 0, (0 < \eta < 1)$ and an integer N_0 such that $\forall n \geq N_0$

$$P \left[\sup_{|s-t| < \eta} |W_n(s) - W_n(t)| \geq \epsilon \right] \leq \epsilon. \quad (121)$$

From Lemma 105, we deduce that if $t \geq s$ and $t - s \geq n^{-1}$, then

$$E(W_n(t) - W_n(s))^4 \leq C(t - s)^{4/(2+\delta)} \quad (122)$$

where $4/(2+\delta) > 1$, which implies (121) (by using Proposition 2.2 in Harel and Puri (1987)) and (ii) is proved.

Put $G_n^*(t) = n^{-1} \sum_{i=1}^n I_{[F_n(X_{ni}) \leq t]}$.

LEMMA 123. Let $\epsilon > 0$ be arbitrarily. Then there is a constant C , such that, up to an event of probability less than or equal to ϵ , we have

$$|t - s| \leq C a_n \quad (124)$$

whenever

$$|G_n^*(t) - G_n^*(s)| \leq a_n. \quad (125)$$

PROOF. From Theorems 6.1 and 14.3 of Billingsley (1968) and the fact that the sequence of probability measures P_n defined by W_n is tight (see proof of (ii) in Lemma 112), we deduce that there exists some finite number M such for all n sufficiently large,

$$P(A_n(M)) \geq 1 - \epsilon$$

where $A_n(M) = \{\sup_{0 \leq t \leq 1} |n^{-1}G_n^*(t)| \leq C_1 n^{-\frac{1}{2}}\}$.

From (125), on the set $A_n(M)$ we have

$$\begin{aligned} |t - s| &\leq |t - G_n^*(t)| + |G_n^*(t) - G_n^*(s)| + |G_n^*(s) - s| \\ &= |n^{-1}G_n^*(t)| + |G_n^*(t) - G_n^*(s)| + |n^{-1}G_n^*(s)| \\ &\leq a_n + 2C_1 n^{-\frac{1}{2}} \leq C a_n \end{aligned}$$

and the proof is completed.

LEMMA 126. If the sequence $\{X_{ni}\}$ is strong mixing with rates (106), then for any $\lambda > 0$,

$$P \left[\sup_{s \leq t \leq s + C_1 a_n} |W_n(t) - W_n(s)| \geq \lambda a_n^{\frac{1}{2}} \right] \leq C' \lambda^{-4}, \quad (127)$$

where $W_n = n^{-\frac{1}{2}} \hat{G}_n$.

PROOF. From Lemma 105, if $t - s \geq \epsilon n^{-1}$, then $E|W_n(t) - W_n(s)|^4 \leq C' \epsilon^{-1} (t - s)^{4/(2+\delta)}$ for all n . Hence following the method of the proof of (22.20) in Billingsley (1968), we get

$$P \left[\sup_{s \leq t \leq s + mp} |W_n(t) - W_n(s)| \geq \lambda a_n^{\frac{1}{2}} \right] \leq C' \epsilon^{-1} \lambda^{-4} (mp)^{4/(2+\delta)}. \quad (128)$$

Now putting $mp = C_1 a_n$, we have (127) from (128).

The following lemma is Theorem 1 of Yokoyama (1980).

LEMMA 129. Let $\{Y_j\}$ be a strong mixing sequence with $E(Y_j) = 0$, $j \geq 1$, and $\sup_{j \geq 1} E|Y_j|^{r+\delta} < \infty$ for some $r > 2$ and $\delta > 0$. If

$$\sum_{i=0}^{\infty} (i+1)^{\frac{r}{2}-1} [\alpha(i)]^{\delta/(r+\delta)} < \infty \quad (130)$$

then

$$E \left| \sum_{i=1}^n Y_i \right|^r \leq C' n^{\frac{r}{2}}, \quad n \geq 1. \quad (131)$$

ACKNOWLEDGEMENTS.

The authors would like to express sincere thanks to the referees for highly valuable comments which indeed were very helpful in revising the manuscript. Research of Madan L. Puri was supported by the Office of Naval Research Contract N00014-91-J-1020.

REFERENCES

- Bennett, G. (1962). Probability inequalities for the sum of independent random variables. Journal of the American Statistical Association 57, 33-45.
- Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.
- Davydov, Yu.A. (1973). Mixing conditions for Markov chains. Theory of Probability Appl., 18, 312-328.
- Devroye, L.P. and Wagner, T.J. (1980). On the L_1 convergence of kernel estimators of regression functions with applications in discrimination. Zeitschrift Wahrscheinlichkeitstheorie verw. Gebiete 51, 15-25.
- Doukhan, P. and Portal, F. (1987). Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et mélangeant. Prob. and Math. Stat. Vol. 8, Fa. 2, 117-132.
- Harel, M. and Puri, M.L. (1987). Weak convergence of weighted multivariate empirical processes under mixing conditions. In Mathematical Statistics and Probability Theory, Vol. A: Theoretical Aspects (M.L. Puri, P. Révész and W. Wertz, Eds.). Reidel Publication Company Holland, 1987, pp. 121-141.
- Harel, M. and Puri, M.L. (1989). Limiting behavior of U -statistics, V -statistics and one sample rank order statistics for nonstationary absolutely regular processes. Journal of Multivariate Analysis 30, 181-204.
- Harel, M. and Puri, M.L. (1991). Théorème central limite de l'estimateur de densité et de l'erreur quadratique intégrée pour des variables aléatoires non stationnaires. Compte Rendus A. Sciences, t. 312, Série I, 145-148.

Nummelin, E., and Tuominen, P. (1982). Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. Stochastic Processes and Applications, 12, 187-202.

Stein, E.M. (1970). Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton, New Jersey.

Stute, W. (1984). Asymptotic normality on nearest neighbor regression function estimates. Annals of Statistics 12, 917-926.

Withers, C.S. (1975). Convergence of empirical processes of mixing rv's on $[0, 1]$. Annals of Statistics 3, 1101-1108.

Yokoyama, R. (1980). Moment bounds for stationary mixing sequences. Zeitschrift Wahrscheinlichkeitstheorie verw. Gebiete 52, 45-57.

Yoshihara, K. (1976). Limiting behavior of U -statistics for stationary absolutely regular processes. Zeitschrift Wahrscheinlichkeitstheorie verw. Gebiete 35, 237-252.

Yoshihara, K. (1988). Asymptotic normality of nearest neighbor regression function estimates based on some dependent observations. Yokohama Mathematical Journal 36, 55-68.