YOKOHAMA MATHEMATICAL
JOURNAL VOL. 39, 1991

ASYMPTOTIC NORMALITY OF RANK NEAREST NEIGHBOR REGRESSION FUNCTION ESTIMATORS UNDER STRONG MIXING[1]

By

K. L. MEHRA, Y. S. RAMAKRISHNAIAH[2] and M. SUDHAKARA RAO[2]

(Received May 25, 1990)

Summary. Let $\{(X_n, Y_n): n=1, 2, \cdots\}$ be a strictly stationary strong mixing sequence of random vectors in R^{d+p} and denote by $r_\phi(x_0)=E[\phi(Y)|X=x_0]$, where ϕ is a real Borel function defined on $R^p, p\geqq 1$. In this paper, we prove for the above sequence, the asymptotic normality of the rank nearest neighbor kernel estimators of $r_\phi(x_0)$, studied by Yang [11], Stute [10] and Yoshihara [13].

1. Introduction. Let $\{(X_n, Y_n): n=1, 2, \cdots\}$ be a strictly stationary sequence of $(d+p)$-dimensional random vectors (r. v.'s) with a continuous distribution (d. f.) $H(x, y)=P(X\leqq x, Y\leqq y)$ and marginal d. f.'s F and G respectively. Let ϕ be a real Borel function on R^p, with $E|\phi(Y)|<\infty$ and $r_\phi(x)=E[\phi(Y)|X=x]$ denote the regular conditional expectation of $\phi(Y)$, given $X=x$ assuming that the latter quantity exists and is finite. Following the ideas of Stone [9], Yang [11] proposed a kernel estimator for $r_\phi(x_0)$ of the rank nearest neighbor (RNN) type for the case $p=d=1$. A natural extension of Yang's estimator for $p\geqq 1$, $d\geqq 1$ is given by

$$r_{n,\phi}(x_0)=(na_n^d t_n(x_0))^{-1}\sum_{i=1}^{n}\phi(Y_i)K_{n,x_0}(X_i), \tag{1.1}$$

where $K_{n,x_0}(x)=K(a_n^{-1}(F_{n1}(x_1)-F_{n1}(x_{01})), \cdots, a_n^{-1}(F_{nd}(x_d)-F_{nd}(x_{0d})))$, with K denoting a d-dimensional suitable kernel function, $x=(x_1, x_2, \cdots, x_d)'$, $x_0=(x_{01}, x_{02}, \cdots, x_{0d})'$, F_{nj} denoting the standard empirical d. f. based on $\{X_{ij}, 1\leqq i\leqq n, 1\leqq j\leqq d\}$, $t_n(x_0)=(na_n^d)^{-1}\sum_{i=1}^n K_{n,x_0}(X_i)$ and $\{a_n\}$ a bandwidth sequence with $a_n\to 0$, as $n\to\infty$. From the nonparametric standpoint, the RNN estimator (1.1) is preferable to its Nadaraya-Watson (NW) conterpart obtained by replacing

[1] The above research was supported in part by a CRF-NSERC Grant from the University of Alberta and the Grant No. A-3061 from the National Science and Engineering Research Council of Canada.

[2] On leave from Osmania University, Hyderabad, India.

AMS 1980 Subject Classification. 62J02, 62E20, 60G05.

Key words and Phrases. Rank Nearest Neighbor regression, strong mixing.

$F_{nj}(x_j)$ with x_j, $1 \leqq j \leqq d$ in (1.1). Since the asymptotic variance of the above proposed RNN estimator equals $f(x_0)$ times the variance of the NW estimator (cf. Robinson [7] and Theorem 3.1 below), in terms of the asymptotic relative efficiency the RNN estimator would be superior to the NW estimator for most values of $\boldsymbol{x}_0$ (See Remark 4.2 below).

For the case $p=1$, $d=1$, $\phi(Y)=Y$ and when $\{(X_i, Y_i)\}$ is a stationary independent sequence, Stute [10] had shown that the estimator (1.1), suitably normalized, is asymptotically normal. This result was recently extended to $*$-mixing and ϕ-mixing sequences by Yoshihara [13], who utilized for this purpose his celebrated "approximation" lemma for absolutely regular processes (See Lemma 3.1 of [13]). However, his results do not cover the important and more general case of strong mixing sequences.

The objective of the present paper is to establish the above result to cover stationary strong mixing vector sequences $\{(X_i, Y_i)\}$. It should be pointed out in this connection that the above referred "approximation" technique of Yoshihara [13], coupled with Stute's [10] method, does not seem to extend to the case of strong mixing sequences. We are, however, able to circumvent this ostensible difficulty by using the weak convergence of multidimensional empirical processes for strong mixing sequences coupled with some approximations using "truncation" arguments. We also make use of certain probability bounds on the oscillations of empirical processes, which along with some notation, assumptions and preliminaries are given in Section 2. The main result is proved in Section 3. Finally, some useful remarks are presented in the concluding Section 4.

2. Notations, assumptions and a preliminary result.

2.1. *Notations and definitions.* The rank nearest neighbor (RNN) estimator of the conditional d.f. $m(\boldsymbol{y}|\boldsymbol{x}_0)=P(\boldsymbol{Y} \leqq \boldsymbol{y}|\boldsymbol{X}=\boldsymbol{x}_0)$, $\boldsymbol{y} \in R^p$ for fixed $\boldsymbol{x}_0 \in R^d$, (cf. Yang [11] for $p=1$, $d=1$) is given by

$$m_n(\boldsymbol{y}|\boldsymbol{x}_0)=(n a_n^d t_n(x_0))^{-1} \sum_{i=1}^{n} I_{[Y_i \leq y]} K_{n,x_0}(X_i), \tag{2.1}$$

where K, t_n and a_n are as defined in (1.1). If we replace the indicator function $I_{[Y \leq y]}$ in (2.1) with any real Borel function $\phi(Y)$, we obtain a RNN estimator $r_{n,\phi}(\boldsymbol{x}_0)$ defined by (1.1) of $r_\phi(\boldsymbol{x}_0)=E[\phi(\boldsymbol{Y})|\boldsymbol{X}=\boldsymbol{x}_0]$.

Let $\{\xi_{ni}: 1 \leqq i \leqq n,\ n \geqq 1\}$ be a triangular sequence of random vectors defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$. The sequence $\{\xi_{ni}\}$ is said to be *strong mixing* (s.m.), if there is a function $\alpha(\cdot)$ defined on non-negative integers with $\alpha(N) \downarrow 0$ as $N \to \infty$ and

$$\sup|P(A \cap B)-P(A)P(B)| \leqq \alpha(N) \tag{2.2}$$

for all $A \in \mathcal{F}_{1,k}^n$ and $B \in \mathcal{F}_{k+N,\infty}^n$, where k and n are positive integers, $\mathcal{F}_{a,b}^n$ denotes the sub-σ-algebra generated by the r.v.'s $\{\xi_{ni}: a \leqq i \leqq b\}$ and the sup in (2.2) is taken over all $n \geqq 1$, $k \geqq 1$. The s.m. condition (2.2) was introduced by Rosenblatt [8] and is more general than certain other restrictive types of mixing conditions, such as those of $*$-mixing, ϕ-mixing and absolute regularity (see Ibragimov and Linnik [4]).

2.2. *Assumptions.*

A.I The sequence $\{(X_i, Y_i)\}$ is strong mixing and for some δ, $0<\delta<1$,

$$\sum_{j=1}^{\infty} j^k \alpha^{\delta}(j) < \infty .$$

A.II(i) $H(\boldsymbol{x}, \boldsymbol{y})$ has bounded density $h(\boldsymbol{x}, \boldsymbol{y})$ with marginals $f(\boldsymbol{x})$ and $g(\boldsymbol{y})$ such that $f(x)$ is bounded away from zero in some open neighborhood (nghd) N_0 of $\boldsymbol{x}_0 \in \boldsymbol{R}^d$;

(ii) the joint d.f. $F_j(\boldsymbol{x}_1, \boldsymbol{x}_2)$ of $(\boldsymbol{X}_1, \boldsymbol{X}_{1+j})$ has a continuous density $f_j(\boldsymbol{x}_1, \boldsymbol{x}_2)$ in an open nghd N_0^* of $(\boldsymbol{x}_0, \boldsymbol{x}_0)$, $\boldsymbol{x}_0 \in \boldsymbol{R}^d$.

A.III For each fixed $\boldsymbol{y} \in \boldsymbol{R}^p$, $m(\boldsymbol{y} | \boldsymbol{F}^{-1}(\boldsymbol{t}))$, where $\boldsymbol{F}^{-1}(\boldsymbol{t}) = (F_1^{-1}(t_1), F_2^{-1}(t_2), \cdots, F_d^{-1}(t_d))$ with $0<t_j<1$ for $j=1, 2, \cdots, d$, is twice continuously differentiable in $\boldsymbol{t} \in N_0^{**} = N_{01}^{**} \times \cdots \times N_{0d}^{**}$, where N_{0j}^* is an open nghd of $F(x_{0j})$ for $j=1, 2, \cdots, d$, and

$$\max_{1 \leq i, j \leq d} \sup_{\boldsymbol{t} \in N_0^{**}} \sup_{\boldsymbol{y}} \left| \frac{\partial^2}{\partial t_i \partial t_j} m(\boldsymbol{y} | \boldsymbol{F}^{-1}(\boldsymbol{t})) \right| < \infty .$$

A.IV K is any probability kernel function on $\boldsymbol{R}^d$, twice continuously differentiable and vanishing outside a compact interval, (say) $[-1, 1]^d$ and satisfying $\int u_j K(\boldsymbol{u}) d\boldsymbol{u} = 0$ for $1 \leqq j \leqq d$.

A.V The bandwidth sequence $\{a_n\}$ satisfies $0<a_n \downarrow 0$, $n a_n^{d+2+\delta} \to \infty$ and $n a_n^{d+4} \to \tau$ as $n \to \infty$, where $\tau \geqq 0$ is a constant and δ $(0<\delta<1)$ is as given in A.I.

The symbols such as $C_1, C_2, \cdots$ (i.e., C's with subscripts) that appear throughout denote generic constants.

2.3. *A preliminary result.* We need the result of Lemma 2.1 below pertaining to the oscillations of the (univariate) empirical processes for strong mixing sequences. The results can be deduced from Mehra and Rao [5]. We sketch the proof giving only necessary details: Let $\{U_i : i \geqq 1\}$ be a mixing sequence of uniform $[0, 1]$ r.v.'s and consider the empirical process $\{V_n(t): 0 \leqq t \leqq 1\}$, where $V_n(t) = n^{-1/2} \sum_{i=1}^{n} (I_{[U_i \leq t]} - t)$. We now state

Lemma 2.1. *Let* $\{U_i: i \geqq 1\}$ *be a s.m. sequence of uniform r.v.'s. satisfying A.I with* $k=2$. *Then for given,* c, $\lambda>0$ *and* $0 \leqq s, t \leqq 1$), *we have for sufficiently large* n

$$P(\sup_{s\le t\le s+ca_n}|V_n(t)-V_n(s)|\ge\lambda a_n^{(1-2^{-1}\delta)/2})\le C_1\lambda^{-4},$$

where C_1 *is a constant, depending only on* c, α, δ.

Proof. The proof follows on the lines of the arguments of Billingsley [2], Theorem 22.1). We have from Lemma 2.6 (i) of Mehra and Rao [5], with $C_i\equiv 1$, $q\equiv 1$ there,

$$E|V_n(t)-V_n(s)|^4\le C_1[(t-s)^{2-\delta}+n^{-1}(t-s)^{1-\delta}]$$
$$\le 2C_1\varepsilon^{-1}(t-s)^{2-\delta},\quad 0\le s,\ t\le 1,$$

for any arbitrary $\varepsilon>0$ and $|t-s|>\varepsilon n^{-1}$. Consequently, arguing as in Billingsley ([2], p. 199), it can be shown, for a suitably selected small p and a positive integer m that

$$P(\sup_{s\le t\le s+mp}|V_n(t)-V_n(s)|\ge\lambda a_n^{(1-2^{-1}\delta)/2})$$
$$\le 2\varepsilon^{-1}C_1(\lambda a_n^{(1-2^{-1}\delta)/2})^{-4}(mp)^{2-\delta}.$$

which completes the proof (with $mp=ca_n$). □

3. The main result. In this section we shall prove the asymptotic normality of the RNN estimator $r_{n,\phi}(\boldsymbol{x}_0)$ defined by (1.1) of the regression function $r_\phi(\boldsymbol{x}_0)$, $\boldsymbol{x}_0\in\boldsymbol{R}^d$.

Let $H_n(\boldsymbol{x},\boldsymbol{y})=n^{-1}\sum_{i=1}^n I_{[X_i\le x, Y_i\le y]}$ for $\boldsymbol{x}\in\boldsymbol{R}^d$, $\boldsymbol{y}\in\boldsymbol{R}^p$ and for $\boldsymbol{x}_0\in\boldsymbol{R}^d$, set

$$\nu_{n,\phi}(\boldsymbol{x}_0)=(na_n^d)^{-1/2}\sum_{i=1}^n[\phi(\boldsymbol{Y}_i)-r_\phi(\boldsymbol{x}_0)]K_{n,x_0}(\boldsymbol{X}_i), \tag{3.1}$$

and

$$\beta_{n,\phi}(\boldsymbol{x}_0)=(na_n^d)^{1/2}[r_{n,\phi}(\boldsymbol{x}_0)-r_\phi(\boldsymbol{x}_0)]. \tag{3.2}$$

Then, from (3.1) and (3.2), $\beta_{n,\phi}(\boldsymbol{x}_0)=\nu_{n,\phi}(\boldsymbol{x}_0)/t_n(\boldsymbol{x}_0)$. By showing $t_n(\boldsymbol{x}_0)\xrightarrow[p]{}1$ as $n\to\infty$, we shall establish the asymptotic normality of $\beta_{n,\phi}(\boldsymbol{x}_0)$ via that of $\nu_{n,\phi}(\boldsymbol{x}_0)$. Towards this end, by the assumption A.IV and Taylor's expansion, we have the followlng decomposition of $\nu_{n,\phi}(\boldsymbol{x}_0)$:

$$\begin{aligned}\nu_{n,\phi}(\boldsymbol{x}_0)&=(n/a_n^d)^{1/2}\int_{A_n}[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]K_{(n),x_0}(\boldsymbol{x})dH_n(\boldsymbol{x},\boldsymbol{y})\\&+a_n^{-(d+2)}\int_{A_n}[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]\sum[U_{nj}(x_j)-U_{nj}(x_{0j})]K_{(n),x_0}^{(j)}(\boldsymbol{x})dH_n(\boldsymbol{x},\boldsymbol{y})\\&+\frac{1}{2}(na_n^{(d+4)})^{-1/2}\int_{A_n}[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]\sum_{j,j'}[U_{nj}(x_j)-U_n(x_{0j})]\\&\times[U_{nj}(x_j')-U_{nj}(x_{0j'})]K^{(j,j')}(\Delta_{n,x})dH_n(\boldsymbol{x},\boldsymbol{y})\\&=J_{n1}+J_{n2}+J_{n3}\ \text{(say)},\end{aligned} \tag{3.3}$$

where

$$A_n=\cap_{j=1}^{d}\{x_j:\ |F_{nj}(x_j)-F_{nj}(x_{0j})|\leqq a_n\},$$

$$K^{(j)}(\boldsymbol{x})=\frac{\partial}{\partial x_j}K(\boldsymbol{x}),\qquad K^{(j,j')}(\boldsymbol{x})=\frac{\partial^2}{\partial x_j x_{j'}}K(\boldsymbol{x}),$$

$$K_{(n),x_0}(\boldsymbol{x})=K((F_1(x_1)-F_1(x_{01})/a_n,\ \cdots,\ (F_d(x_d)-F_d(x_{0d}))/a_n),$$

$$U_{nj}(x_j)=\sqrt{n}\,[F_{nj}(x_j)-F_j(x_j)]\quad\text{and}\quad \Delta_{n,x}=(\Delta_{n,x_1}^{(1)},\ \cdots\ \Delta_{n,x_d}^{(d)})',$$

$a_n\Delta_{n,x_j}^{(j)}$ lying between $F_{nj}(x_j)-F_{nj}(x_{0j})$ and $F_j(x_j)-F_j(x_{0j})$ for $1\leqq j\leqq d$. The integral sign in (3.3) and below should be understood as single or multiple integral depending on the context.

We first consider the asymptotic behaviour of J_{nj}, $j=2, 3$:

Lemma 3.1. *If the sequence* $\{(\boldsymbol{X}_i, Y_i)\}$ *is s.m. satisfying A.I with* $k=2(d+p)$, *A.II to A.V, then* $E|\phi(Y_1)|^{2+\delta_1}<\infty$ *for some* $\delta_1>2\delta/(2+\delta)$ *implies* $J_{nj}\underset{p}{\rightarrow}0$, *for* $j=2$ *and* 3, *as* $n\rightarrow\infty$, *where* J_{nj}*'s are as defined in* (3.3).

Proof. We first deal with J_{n3}. Since K vanishes outside $[-1, 1]$, the expansion (3.3) holds with integration restricted to the set A_n and, further, since $\max_{1\leq j\leq d}\sup_{x_j\in R}|U_{nj}(x_j)|=O_p(1)$, as $n\rightarrow\infty$, in view of Theorem 3.2 of Mehra and Rao [5], on the preceding set $|F_j(x_j)-F_j(x_{0j})|\leqq c_j a_n$, $1\leqq j\leqq d$, for some constants c_j's and sufficiently large n (see Lemma 3.6 of Yoshihara [13]). Consequently,

$$(3.4)\qquad |J_{n3}|\underset{p}{\leqq}\frac{1}{2}(na_n^{d+4})^{-1/2}\max_{1\leq j\leq d}\sup_{x\in A_n^*}|U_{nj}(x_j)-U_{nj}(x_{0j})|^2$$

$$\times\int_{A_n^*}|\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)|\sum_{j,j'}|K^{(j,j')}(\Delta_{n,x})|\,dH_n(\boldsymbol{x},\boldsymbol{y}),$$

where $A_n^*=\cap_{j=1}^{d}\{x_j:\ [|F_j(x)-F_j(x_{0j})|\leqq c_j a_n]\}$. Further, since $K^{(j,j')}$ is bounded, the integral on the RHS of (3.4) is $O_p(1)$ as $n\rightarrow\infty$ by the law of large numbers (which does hold for strong mixing sequences and the assumed conditions here). Also since $\{V_{nj}\circ F\}\underset{\mathcal{L}}{=}\{U_{nj}\}$, we have from Lemma 2.1

$$(3.5)\qquad \max_{1\leq j\leq d}\sup_{x\in A_n^*}|U_{nj}(x_j)-U_{nj}(x_{0j})|=O_p(a_n^{(2-\delta)/4}),\quad n\rightarrow\infty,$$

for the s.m. sequence satisfying A.I with $k=2(1+p)$. Thus $J_{n3}\underset{p}{\rightarrow}0$ as $n\rightarrow\infty$, follows from (3.4) and (3.5) in view of Assumption A.V. Next we show that $J_{n2}\underset{p}{\rightarrow}0$ as $n\rightarrow\infty$. This result is crucial in establishing the main result (cf. Lemma 2 of Stute [10] and Lemma 4.2 of Yoshihara [13]). We have

$$(3.6)\qquad J_{n2}=a_n^{-(d+2)/2}\int_{A_n\times R^p}[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]\sum_j[U_{nj}(x_j)-U_{nj}(x_{0j})]$$

$$\times K_{(n),x_0}^{(j)}(\boldsymbol{x})d[H_n(\boldsymbol{x},\boldsymbol{y})-H(\boldsymbol{x},\boldsymbol{y})]$$

$$+a_n^{-(d+2)/2}\int_{A_n^*\times R^p}[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]\sum_j[U_{nj}(x_j)-U_{nj}(x_{0j})]$$
$$\times K_{(n),\boldsymbol{x}_0}^{(j)}(\boldsymbol{x})dH(\boldsymbol{x},\boldsymbol{y}),$$
$$=\xi_n+\eta_n \text{ (say)}.$$

Since $H(\boldsymbol{x},\boldsymbol{y})=G(\boldsymbol{y}|\boldsymbol{x})F(\boldsymbol{x})$ and K vanishes outside $[-1, 1]$, we have as for (3.4)

$$|\eta_a|=a_n^{-(d+2)/2}\left|\int_{A_n}E[(\phi(Y)-r_\phi(\boldsymbol{x}_0))|X=\boldsymbol{x}]\sum_j[U_{nj}(x_j)-U_{nj}(x_{0j})]K_{(n),\boldsymbol{x}_0}^{(j)}(\boldsymbol{x})dF(\boldsymbol{x})\right|$$

$$\underset{p}{\leqq} a_n^{-(d+2)/2}\int_{A_n^*}|r_\phi(\boldsymbol{x})-r_\phi(\boldsymbol{x}_0)|\sum_j|U_{nj}(x_j)-U_{nj}(x_{0j})|\,|K_{(n),\boldsymbol{x}_0}^{(j)}(\boldsymbol{x})|dF(\boldsymbol{x}).$$

Now making the transformation $F_j(x_j)-F_j(x_{0j})=a_nt_j$, $1\leqq j\leqq d$, using the continuous differentiability of $r_\phi\circ\boldsymbol{F}^{-1}(\boldsymbol{t})$ in a nghd N_0^{**} (Assumption A.III) and Taylor's expansion, it is easy to verify in view of (3.5) that, under the stated mixing condition,

(3.7) $$\eta_n=O_p(a_n^{(d+1-\delta/2)/2}) \quad \text{as} \quad n\to\infty.$$

In dealing with the first term ξ_n in (3.6), the cases $d=1$ and $d>1$ need separate treatment. We shall first treat the simpler $d>1$ case: For some constant C_1,

(3.8) $$|\xi_n|\leqq\frac{d}{a_n^{(d+2)/2}}\max_{1\leq j\leq d}\sup_{\boldsymbol{x}\in A_n^*}|U_{nj}(x_j)-U_{nj}(x_{0j})|$$
$$\times\int|\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)|\max_{1\leq j\leq d}|K_{(n),\boldsymbol{x}_0}^{(j)}(\boldsymbol{x})|d(H_n+H)(\boldsymbol{x},\boldsymbol{y})$$
$$\leqq C_1a_n^{(d-1-(\delta/2))/2}(\xi_{n1}^*+\xi_{n2}^*) \text{ (say)},$$

where, as $n\to\infty$

(3.8a) $$\xi_{n1}^*=\frac{1}{na_n^d}\sum_{i=1}^n|\phi(Y_i)-r_\phi(\boldsymbol{x}_0)|\max_{1\leq j\leq d}|K_{(n),x_0}^{(j)}(X_i)|$$
$$\underset{p}{\longrightarrow}(f(\boldsymbol{x}_0)/\Pi_{j=1}^d f_j(x_{0j}))\int E[|\phi(Y)-r_\phi(\boldsymbol{x}_0)|/X=\boldsymbol{x}_0]\max_{1\leq j\leq d}|K^{(j)}(\boldsymbol{t})|dt$$
$$=\lim_{n\to\infty}\xi_{n2}^*<\infty;$$

(3.8) and (3.8a) imply

(3.9) $$|\xi_n|\leqq O(a_n^{(d-1-(\delta/2))/2})$$
$$\overset{p}{\to}0$$

as $n\to\infty$, if $d\geqq2$. Next for the case $d=1$, we write

(3.10) $$\xi_n=-a_n^{-3/2}r_\phi(x_0)\int_{A_n\times R^p}[U_n(x)-U_n(x_0)]K'_{(n),x_0}(x)d(H_n-H)(x,\boldsymbol{y})$$

$$+a_n^{-3/2}\int_{A_n\times R^p}\phi(\boldsymbol{y})[U_n(x)-U_n(x_0)]K'_{(n),x_0}(x)d(H_n-H)(x,\boldsymbol{y})$$
$$=\xi_{n1}+\xi_{n2},\ \text{(say)}.$$

We will show that $\xi_{n2}\xrightarrow[p]{}0$ as $n\to\infty$ and the same holds for ξ_{n1} by parallel arguments: Let $B_n=\{\boldsymbol{y}: |\phi(\boldsymbol{y})|\leqq b_n\}$, where $b_n\uparrow\infty$ as $n\to\infty$ and B_n^c denote the complement of B_n. Then we have

$$\begin{aligned}(3.11)\qquad \xi_{n2}&=(na_n^3)^{-1/2}\int_{A_n^*\times B_n}\phi(\boldsymbol{y})[U_n(x)-U_n(x_0)]K'_{(n),x_0}(x)dW_n(x,\boldsymbol{y})\\&+a_n^{-3/2}\int_{A_n^*\times B_n^c}\phi(\boldsymbol{y})[U_n(x)-U_n(x_0)]K'_{(n),x_0}(x)dH_n(x,\boldsymbol{y})\\&-a_n^{-3/2}\int_{A_n^*\times B_n^c}\phi(\boldsymbol{y})[U_n(x)-U_n(x_0)]K'_{(n),x_0}(x)dH(x,\boldsymbol{y})\\&=\xi_{n21}+\xi_{n22}+\xi_{n23}\quad\text{(say)},\end{aligned}$$

where $W_n(x,\boldsymbol{y})=\sqrt{n}\,[H_n(x,\boldsymbol{y})-H(x,\boldsymbol{y})]$ and A_n^* is as defined in (3.4). Now by the boundedness of K' and $g_x(\boldsymbol{y})$ and the assumption $E|\phi(Y)|^{2+\delta_1}<\infty$, we have, as $n\to\infty$,

$$(na_n)^{-1}\sum_{i=1}^n|\phi(Y_i)|^{(2+\delta_1)/2}|K'_{(n),x_0}(X_i)|\xrightarrow[p]{}E[|\phi(Y_1)|^{(2+\delta_1)/2}|x_0]\int|K'(u)|du.$$

In view of this and (3.5), we get

$$\begin{aligned}(3.12)\qquad |\xi_{n22}|&\leqq_p Ca_n^{-3/2}a_n^{(1/2)(1-(\delta/2))}\int_{A_n^*\cap B_n^c}|\phi(\boldsymbol{y})|\,|K'_{(n),x_0}(x)|\,dH_n(x,\boldsymbol{y})\\&\leqq_p Ca_n^{-\delta/4}\cdot b_n^{-\delta_1/2}\cdot\left[a_n^{-1}\int|\phi(\boldsymbol{y})|^{(2+\delta_1)/2}|K'_{(n),x_0}(x)|\,dH_n(x,\boldsymbol{y})\right]\\&=O_p(a_n^{-\delta/4}\cdot b_n^{-\delta_1/2}).\end{aligned}$$

Similarly, we can get

$$(3.13)\qquad |\xi_{n23}|=O_p(a_n^{-\delta/4}\cdot b_n^{-\delta_1/2}).$$

Finally, we note that the multi-dimensional empirical process $\{W_n(x,\boldsymbol{y})\}$ defined in (3.11) converges weakly to an a.s. Gaussian process under strong mixing conditions A.I with $k=2(1+p)$ (cf. Yoshihara [12], Mehra et al. [6],) so that $\sup_{x,y}|W_n(x,\boldsymbol{y})|=O_p(1)$. Since K' is bounded, we thus have from (3.5) and (3.11) that

$$\begin{aligned}(3.14)\qquad |\xi_{n21}|&\leqq C(na_n^3)^{-1/2}\sup_{x\in A_n^*}|U_n(x)-U_n(x_0)|\cdot\sup_{x,y}|W_n(x,\boldsymbol{y})|\cdot b_n\\&=O_p((na_n^{3+\delta})^{-1/2}\cdot a_n^{(1+\delta/2)/2}\cdot b_n).\end{aligned}$$

By letting $b_n=a_n^{-(1+\delta/2)/2}$, it follows from (3.12) to (3.14) that $\xi_{n2}\underset{p}{\to}0$ as $n\to\infty$, in view of $na_n^{3+\delta}\to\infty$, the fact that $\delta_1>2\delta/(2+\delta)$ and the Assumption A. V. This coupled with (3.10) and (3.9) that $\xi_n\underset{p}{\to}0$ as $n\to\infty$, for all $d\geqq 1$. Consequently, the proof of the lemma is complete in view of (3.6) and (3.7). □

Lemma 3.2. *Under the assumptions of Lemma* 3.1, *the sequence* $\{\nu_{n,\phi}(\boldsymbol{x}_0)\}$ *is expressible as*

$$\nu_{n,\phi}(\boldsymbol{x}_0)=T_{n,\phi}(\boldsymbol{x}_0)+\psi_n,$$

where

$$T_{n,\phi}(\boldsymbol{x}_0)=(n/a_n^d)^{1/2}\int[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})d(H_n-H)(\boldsymbol{x},\boldsymbol{y}) \tag{3.15}$$

and

$$\lim_{n\to\infty}\psi_n=\tau^{1/2}\sum_{j=1}^d\frac{\partial^2}{\partial^2 x_j}r_\phi(\boldsymbol{x})\Big|_{\boldsymbol{x}=\boldsymbol{x}_0}\Big(\int t_j^2K(\boldsymbol{t})d\boldsymbol{t}\Big)=b_\phi(\boldsymbol{x}_0)$$

Proof. From (3.3) the term J_{n1} can be written as

$$J_{n1}=(n/a_n^d)^{1/2}\int[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})d(H_n-H)(\boldsymbol{x},\boldsymbol{y})+\psi_{n1}+\psi_{n2}, \tag{3.16}$$

where setting

$$\bar{A}_n=\cap_{j=1}^d\{x_j:\ |F_j(x_j)-F_j(x_{0j})|\leqq a_n\},$$

$$\psi_{n1}=-(n/a_n^d)^{1/2}\int_{A_n^c\cap\bar{A}_n}[\varphi(\boldsymbol{y})-r_\varphi(\boldsymbol{x}_0)]K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})dH_n(\boldsymbol{x},\boldsymbol{y}) \tag{3.17}$$

and

$$\begin{aligned}\psi_{n2}&=(n/a_n^d)^{1/2}\int[\phi(\boldsymbol{y})-r_\phi(\boldsymbol{x}_0)]K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})dH(\boldsymbol{x},\boldsymbol{y})\\&=(n/a_n^d)^{1/2}\int[r_\phi(\boldsymbol{x})-r_\phi(\boldsymbol{x}_0)]K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})dF(\boldsymbol{x})\\&\longrightarrow b_\phi(\boldsymbol{x}_0)\quad\text{as}\quad n\to\infty,\end{aligned} \tag{3.18}$$

by using Taylor's expansion of $r_\phi(\boldsymbol{x})$ around $\boldsymbol{x}_0$ and transformations as for (3.7), and the Assumptions A. III, A. IV and A. V. Now we deal with the term ψ_{n1}: For this first note that, for sufficiently large n, on the set A_n for some j $(1\leqq j\leqq d)$

$$\begin{aligned}|F_j(x_j)-F_j(x_{0j})|&\geqq||F_{nj}(x_j)-F_{nj}(x_{0j})|-n^{-1/2}|U_n(x_j)-U_n(x_{0j})||\\&\underset{p}{>}a_n(1-\tau_n),\end{aligned} \tag{3.19}$$

the last inequality following since by Lemma 2.1, uniformly in $\boldsymbol{x}$ (also $\boldsymbol{x}_0$),

$$n^{-1/2}|U_n(x_j)-U_n(x_{0j})|\underset{p}{\leqq}a_n\tau_n,\qquad\text{with}\quad\tau_n=[c/(na_n^2)^{1/2}],$$

as $n\to\infty$, for some constant $c>0$. In view of (3.17) and (3.19)

$$|\psi_{n1}| \underset{p}{\leqq} (n/a_n^d)^{1/2}\int_{D_n(\boldsymbol{x}_0)} |\varphi(\boldsymbol{y})-r_{\varphi}(\boldsymbol{x}_0)|\,|K_{(n),\boldsymbol{x}_0}(\boldsymbol{x})|\,dH_n(\boldsymbol{x},\boldsymbol{y}) \tag{3.20}$$

$$\underset{p}{\to} 0, \quad \text{as} \quad n\to\infty,$$

where

$$D_n(\boldsymbol{x}_0)=\cup_{j=1}^d\{x_j:\ a_n(1-\tau_n)<|F_j(x_j)-F_j(x_{0j})|\leqq 1\},$$

the last convergence following again by the law of large numbers and by noting that the expectation and variance of the preceding expression converge to zero, as $n\to\infty$. To convince oneself of the last assertion, one simply needs to make the transformations $F_j(x_j)-F_j(x_{0j})=a_n t_j$, $1\leqq j\leqq d$, in the above-referred expressions for the mean and variance and use the same reasoning as for (3.7) and (3.8a) coupled with the fact that, in view of assumption A.I and A.II, the set $\cup_{j=1}^d\{\boldsymbol{t}:(1-\tau_n)<t_j\leqq 1\}$ converges to a Lebesgue null set for all values of c, as $n\to\infty$. The proof of the lemma is complete in view of (3.16), (3.18) and (3.20). □

Lemma 3.3. *Under the assumptions of Lemma* 3.1, $t_n(\boldsymbol{x}_0)\to 1$ *as* $n\to\infty$, *where* $t_n(\boldsymbol{x}_0)$ *is defined in* (1.1).

Proof. The proof follows on the same lines as that of Theorem 3.1 below, based on the decomposition (3.5) and Lemma 3.1 with $\phi(\boldsymbol{y})\equiv 1$. □

Now we state the main result of our paper.

Theorem 3.1. *Suppose that the stationary sequence* $\{\boldsymbol{X}_i, \boldsymbol{Y}_i\}$ *is s.m. satisfying A.I with* $k=2(d+p)$, *A.II–A.V. Further if* $E|\phi(\boldsymbol{Y})|^{2+\delta_1}<\infty$ *for some* $\delta_1>2\delta/(2+\delta)$, *where* δ *is defined by A.I, then for any fixed* $\boldsymbol{x}_0\in\boldsymbol{R}^d$

(i) $$\lim_{n\to\infty} var(T_{n,\phi}(\boldsymbol{x}_0))=\sigma_\phi^2(\boldsymbol{x}_0),$$

where $\sigma_\phi^2(\boldsymbol{x}_0)=var(\phi(\boldsymbol{Y}_1)|\boldsymbol{X}_1=\boldsymbol{x}_0)\int K^2(\boldsymbol{u})d\boldsymbol{u}$,

(ii) $$\beta_{n,\phi}(\boldsymbol{x}_0)=(na_n^d)^{1/2}[r_{n,\phi}(\boldsymbol{x}_0)-r_\phi(\boldsymbol{x}_0)]\xrightarrow[\mathcal{L}]{}\beta_\phi(\boldsymbol{x}_0),$$

where $\beta_\phi(\boldsymbol{x}_0)$ *is a* $N(b_\phi(\boldsymbol{x}_0), \sigma_\phi^2(\boldsymbol{x}_0))$ *random variable with* $b_\phi(\boldsymbol{x}_0)$ *as defined in Lemma* 3.2.

(iii) *Let* $r_\phi^{(M)}=(r_\phi(\boldsymbol{x}_{01}), r_\phi(\boldsymbol{x}_{02}), \cdots, r_\phi(\boldsymbol{x}_{0M}))'$ *for a positive integer M and fixed points* $\boldsymbol{x}_{01}, \boldsymbol{x}_{02}, \cdots, \boldsymbol{x}_{0M}$ *in* $\boldsymbol{R}^d$ *and* $r_{n,\phi}^{(M)}$ *denote the corresponding vector of RNN estimators. Then, as* $n\to\infty$, *the joint asymptotic distribution of* $(na_n^d)^{1/2}(r_{n,\phi}^{(M)}-r_\phi^{(M)})$ *is* $N(\boldsymbol{B}_\phi, \Sigma_\phi^{(M)})$, *where* $\Sigma_\phi^{(M)}=diag(\sigma_\phi^2(\boldsymbol{x}_{01}), \cdots, \sigma_\phi^2(\boldsymbol{x}_{0M}))$ *with* $\sigma_\phi^2(\boldsymbol{x}_{0j})$ *as defined in* (i) *for* $1\leqq j\leqq M$ *and* $\boldsymbol{B}_\phi=(b_\phi(\boldsymbol{x}_{01}), b_\phi(\boldsymbol{x}_{02}), \cdots, b_\phi(\boldsymbol{x}_{0d}))'$.

Proof. We note from (3.15) that

$$T_{n,\phi}(\boldsymbol{x}_0)=n^{-1/2}\sum_{i=1}^{n}[Z_{ni}-E(Z_{ni})],$$

where

$$Z_{ni}=a_n^{-d/2}[\phi(Y_i)-r_\phi(\boldsymbol{x}_0)]K(a_n^{-1}(F_1(X_{i1})-F_1(x_{01})),\cdots,a_n^{-1}(F_d(X_{id})-F_d(x_{0d})))$$

clearly $T_{n,\phi}(\boldsymbol{x}_0)$ is a normalized version of N-W type estimator of $r_\phi(\boldsymbol{x}_0)$-infact it is the usual N-W estimator of $E[\phi(Y)|F_1(X_1)=F_1(x_{01}),\cdots,F_d(X_d)=F_d(x_{0d})]$ which is, under the assumed conditions, the same as $E[\phi(Y)|\boldsymbol{X}=\boldsymbol{x}_0]$. Thus the asymptotic normality of $T_{n,\phi}(x_0)$ with the asymptotic variance as given in (i) follows from Theorem 5.2 of Robinson [7]. The result of (ii) now follows from (3.3), and Lemmas 3.1, 3.2 and 3.3. The joint asymptotic normality of part (iii) also follows from Theorem 5.2 of [7]. The conditions of Theorem 5.2 of [7] are easily verified under the assumed conditions of our theorem. □

It should be noted that when $\tau=0$, i.e. when $na_n^{d+4}\to 0$, the bias $b_\phi(\boldsymbol{x}_0)=0$ in Lemma 3.2 and consequently in Theorem 3.1.

4. Concluding remarks.

1. It should be noted that Theorem 3.1 also yields the asymptotic normality, as $n\to\infty$, of $m_n(\boldsymbol{y}|\boldsymbol{x}_0)$ defined by (2.1):

$$(na_n)^{1/2}[m_n(\boldsymbol{y}|\boldsymbol{x}_0)-m(\boldsymbol{y}|\boldsymbol{x}_0)]\xrightarrow[\mathcal{L}]{} N(b_\phi(\boldsymbol{x}_0),\ \tilde{\sigma}^2_{x_0}),$$

where $\tilde{\sigma}^2_{x_0}=m(\boldsymbol{y}|\boldsymbol{x}_0)[1-m(\boldsymbol{y}|\boldsymbol{x}_0)]\left(\int K^2(\boldsymbol{u})d\boldsymbol{u}\right)$, provided the conditions of Theorem 3.1 hold.

2. According to Theorem 5.1 of Robinson [7], the asymptotic variance of N-W estimator of $r_\phi(x_0)$ is given by $\sigma_\phi^{*2}(x_0)=\sigma_\phi^2(x_0)/f(x_0)$ with $\sigma_\phi^2(x_0)$ as defined in Theorem 3.1 (i) above, the result being the same as in the iid case. Thus from Theorem 3.1, it is clear that the asymptotic relative efficiency of N-W estimator relative to the *RNN* estimator of $r_\phi(x_0)$ in the s. m. case is given by $[\sigma_\phi^2(\boldsymbol{x}_0)/\sigma_\phi^{*2}(\boldsymbol{x}_0)]=f(\boldsymbol{x}_0)$, which is usually small, and less than 1, for a large class of densities f, especially when x_0 is in the tails of the distribution F. Accordingly, the *RNN* estimator should be preferable from the A.R.E. point of view. However, finite sample comparisons should be studied for finer conclusions.

3. Finally, we would like to point out that a *RNN* estimator of conditional density $f(\boldsymbol{y}|\boldsymbol{x}_0)=m'(\boldsymbol{y}|\boldsymbol{x}_0)$ can be defined on the same lines as (1.1):

$$g_n(\boldsymbol{y}|\boldsymbol{x}_0)=(na_n^{d+p})^{-1}\sum_{i=1}^{n}K_n(\boldsymbol{X}_i,\ Y_i;\ \boldsymbol{x}_0,\ \boldsymbol{y}),\tag{4.1}$$

where $\{a_n\}$ denotes a bandwidth sequence and K_n is given by

$$K_n(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{x}_0, \boldsymbol{y}) = K(a_n^{-1}(F_n(x_1) - F_n(x_{01})), \cdots, a_n^{-1}(F_n(x_d) - F_n(x_{0d})), a_n^{-1}(\boldsymbol{z} - \boldsymbol{y}))$$

with K being a $d+p$ dimensional probability kernel. The asymptotic normality of (5.1) can be established using the same method of proof employed in Section 3. Note that g_n can be suitably normalized to make it a probability density and the normalizing factor converges to 1 as $n \to \infty$. For simplicity we state the asymptotic normality of (4.1) when $p=1$, $d=1$.

Theorem 4.1. *Suppose the probability kernel $K(u, v)$ is twice differentiable with bounded and continuous partial derivatives of order 2 and that $\iint tK(s, t)dt = 0$. Further assume that A.I, A.II, A.III hold. If the bandwidth sequence $\{a_n\}$ satisfies $na_n^4 \to \infty$, $na_n^5 \to 0$ as $n \to 0$, then*

(i) $$g_n(y \mid x_0) \to g(y \mid x_0) \ a.s.,$$

and

(ii) $$n^{1/2} a_n (g_n(y \mid x_0) - g(y \mid x_0)) \to N(0, \sigma_0^2),$$

where $\sigma_0^2 = g(y \mid x_0) \iint K^2(u, v)dudv$.

References

[1] Bickel, P. J. and Wichura, M. J.: Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, (1971), 1656-1670.

[2] Billingsley, P.: *Convergence of Probability Measures.* John Wiley, New York, 1968.

[3] Doukhan, P. and Portal, F.: Principe d'invariance pour la fonction de répartition empirique dans un cadre multidimensionnel et melangéant. *Probab. Math. Statist.,* **8** (1987), 117-132.

[4] Ibragimov, I. A. and Linnik, Yu. V.: *Independent and Stationary Sequences of Random Variables.* Wolters-Noordoff Publishers, 1971.

[5] Mehra, K. L. and Rao, M. S.: Weak convergence of generalized empirical processes relative to d_q-metrics under strong mixing. *Ann. Probab.* **3**, (1975), 979-991.

[6] Mehra, K. L., Rao, M. S. and Reddy, M. K.: Weak convergence in d_q-metrics of multidimensional empirical processes for strong mixing sequences. Communicated to *Sankhyā*, A, (1987).

[7] Robinson, P. M.: Nonparametric estimators for time series. *J. Time Series Analysis,* **4** (1983), 185-207.

[8] Rosenblatt, M.: A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA,* **42** (1956), 43-47.

[9] Stone, C.: Consistent nonparametric regression (with discussion). *Ann. Statist.,* **5** (1977), 595-645.

[10] Stute, W.: Asymptotic normality of nearest neighbor regression function estimates. *Ann. Statist.,* **12** (1984), 917-926.

[11] Yang, S.: Linear functions of concomitants of order statistics with applications to nonparametric estimation of regression functions. *J. Amer. Statist. Assoc.*, **76** (1981), 658-662.

[12] Yoshihara, K. I.: Weak convergence of multidimensional empirical processes for strong mixing sequences of stochastic vectors. *Z. Wahrsch. Verw. Gebiete.*, **33** (1975), 133-137.

[13] Yoshihara, K. I.: Asymptotic normality of nearest neighbor regression function estimates based on some dependent observations. *Yokohama Math. J.*, **36** (1988), 55-68.

K. L. Mehra
Department of Statistics
and Applied Probability
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

Y. S. Ramakrishnaiah
Department of Statistics
Osmania University
Hyderabad-50007 (A. P.)
INDIA

M. Sudhakara Rao
Department of Statistics
Osmania University
Hyderabad-50007 (A. P.)
INDIA