Chin. Ann. of Math.
6B (2) 1985

STRONG CONVERGENCE OF KERNEL ESTIMATES OF NONPARAMETRIC REGRESSION FUNCTIONS

ZHAO LINCHENG(赵林城)* FANG ZHAOBEN(方兆本)*

Abstract

Let $(X, Y), (X_1, Y_1), \cdots, (X_n, Y_n)$ be i. i. d. random vectors taking values in $R_d \times R$ with $E(|Y|)<\infty$. To estimate the regression function $m(x)=E(Y|X=x)$, we use the kernel estimate $m_n(x)=\sum_{i=1}^{n} K\left(\frac{X_i-x}{h_n}\right)Y_i \Big/ \sum_{j=1}^{n} K\left(\frac{X_j-x}{h_n}\right)$, where $K(x)$ is a kernel function and h_n a window width. In this paper, we establish the strong consistency of $m_n(x)$ when $E(|Y|^p)<\infty$ for some $p>1$ or $E\{\exp(t|Y|^\lambda)\}<\infty$ for some $\lambda>0$ and $t>0$. It is remakable that other conditions imposed here are independent of the distribution of (X, Y).

§1. Introduction

Let (X, Y), (X_1, Y_1), $\cdots$, (X_n, Y_n) be independent identically distributed $R^d \times R$ valued random vectors with $E|Y|<\infty$. The regression function

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

is estimated by

$$m_n(x)=\sum_{i=1}^{n} W_{ni}(x)Y_i, \tag{1}$$

where $W_{ni}(x)$ is a Borel measurable function of x and $X_1, X_2, \cdots, X_n$. The so-called kernel estimate can be obtained by putting

$$W_{ni}(x)=K\left(\frac{X_i-x}{h}\right)\Big/\sum_{j=1}^{n} K\left(\frac{X_j-x}{h}\right), \tag{2}$$

where $K(x)$ is a given kernel density on R^d and $h=h_n$ is a positive number depending upon n only; for definition, we treat 0/0 in (2) as 0.

The fundamental problem of large-sample nonparametric regression theory is to find the conditions under which $m_n(x)$ is a strong consistent estimate of $m(x)$. The first general result in this direction belongs to Devroye (1981[1]). He proved the strong consistency of $m_n(x)$ under a series of restrictions including the crucial one that Y is bounded. Recently we succeeded in getting rid of this excessively

Manscript received February 9, 1983.

* Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, China.

stringent condition, and establish the strong consistency of $m_n(x)$ under more natural and reasonable conditions.

In this paper, either the ordinary Euclidean norm or the maximum component norm $\|x\|=\max\limits_{1\leq i\leq d}|x^{(i)}|$ can be taken as the norm of $x=(x^{(1)}, \cdots, \lambda^{(d)})$. Such as in [1], the kernel $K(x)$ satisfies

(i) there exist positive numbers, r, $\tilde{c}_1$, $\tilde{c}_2$ such that

$$\tilde{c}_1 I(\|x\|\leq r)\leq K(x)\leq \tilde{c}_2 I(\|x\|\leq r), \tag{3}$$

where $I(A)$ is the indicator function of set A. We estabish the following

Theorem 1. *Suppose that $E|Y|^p<\infty$ for some $p>1$,* kernel K *satisfies conditi on* (3), $\alpha\in\left(\frac{1}{p}, 1\right)$ *is a constant,* $\lim\limits_{n\to\infty} h_n=0$, *and* $\inf\limits_n \{h_n^d/n^{\alpha-1}\}>0$. *Then we have*

$$m_n(X)\to m(X) \quad \text{a. s.} \quad as\ n\to\infty. \tag{4}$$

Theorem 2. *Suppose that $E\{\exp(t|Y|^\lambda)\}<\infty$ for some $\lambda>0$ and $t>0$,* kernel *K satisfies condition* (3), $\alpha>\frac{1}{\lambda}$ *is a constant,* $\lim\limits_{n\to\infty} h_n=0$, *and* $\inf\limits_n \{nh_n^d/(\log n)^{1+\alpha}\}>0$, *then* (4) *is true.*

It is remarkable that the conditions of these theorems impose no specific restrictions on the distribution μ of X.

§2. Proof of the Theorems

For simplicity, we use the following symbols in this paper:

$c>0$ denotes a contant;

$c(x)>0$ denotes a constant depending upon x;

$c(x, \Delta)>0$ denotes a constant depending upon x and Δ, where $\Delta=(X_1, X_2\cdots)$. (These constants can be assumed to be different values in their appearance, even within the same expression.)

μ denotes the distribution of X; F denotes the support of μ;

S_ρ-the closed sphere of radius ρ centered at x.

Lemma 1. *Suppose that* $\int|f(x)|^p\mu(dx)<\infty$ *for some* $p>0$, *then*

$$\lim_{\rho\to 0}\int_{S_\rho}|f(u)-f(x)|^p\mu(du)/\mu(S_\rho)=0, \quad for \quad a.\ e.\ x(\mu). \tag{5}$$

We emphasize that (5) is true for both the norms mentioned ealier. Refer to [2], p. 191, example 20.

Lemma 2. *Let $h=h_n$ be a sequence of positive numbers with* $\lim\limits_{n\to\infty} h=0$. *For all $c>0$, there exists a nonnegative function g with $g(x)<\infty$ such that*

$$h^d/\mu(S_{ch})\to g(x) \quad as \quad n\to\infty, \text{ a. e. } x(\mu). \tag{6}$$

Refer to the proof of [1], Lemma 2.2.

Lemma 3.[3] *Let r. v.* $Y \sim B(n, p)$, $0<p<1$, *for any* $\varepsilon>0$, *we have*

$$P\left(\frac{Y}{n}-p\leqq-\varepsilon\right)\leqslant\exp(-n\varepsilon^2/(2p+\varepsilon)). \tag{7}$$

There exists a similar result for $P\left(\frac{Y}{n}-p\geqslant\varepsilon\right)$.

Proof of the Theorems. In the following we shall make repeatedly use of Lemma 1 and Lemma 2. On each special occasion of its use, there is an exceptive set on which the related formula may not be true, these exceptive sets sum up to a μ-null set. For simplicity of writing we suppose that this set is empty, this can be done without loss of generality. Take $x\in F$. Put $A_i=\{\|X_i-x\|\leqslant rh\}$,

$$N=\sum_{j=1}^{n} I(A_j),\ p_n=\mu(S_{rh}).$$

By (3)

$$J_n(x)\triangleq\left|\sum_{i=1}^{n} W_{ni}(x)(m(X_i)-m(x))\right|\leqslant\tilde{c}_2/\tilde{c}_1\cdot N^{-1}\sum_{i=1}^{n}|m(X_i)-m(x)|I(A_i). \tag{8}$$

Write $g(u)=|m(u)-m(x)|$. Let

$$U_n(x)=N^{-1}\sum_{i=1}^{n} I(A_i)|m(X_i)-m(x)|=N^{-1}\sum_{i=1}^{n} I(A_i)g(X_i), \tag{9}$$

$$\tilde{E}U_n(x)=E\{U_n(x)|I(A_1),\ \cdots,\ I(A_n)\}. \tag{10}$$

Using lemma 1, we have

$$\tilde{E}U_n(x)=N^{-1}\sum_{i=1}^{n} I(A_i)\int_{S_{rh}}|m(u)-m(x)|\mu(du)/\mu(S_{rh})$$

$$=\int_{S_{rh}}|m(u)-m(x)|\mu(du)/\mu(S_{rh})\to 0,\ n\to\infty. \tag{11}$$

We proceed to show that

$$\lim_{n\to\infty} U_n(x)=0\quad \text{a. s.} \tag{12}$$

Write $g_j=g(X_j)$, $\log_2 n=\log\log n$, and take $d_n=\log n\log_2 n$, $c_j=j^{1/p}$ (in Theorem. 1) or $2(1/t\log j)^{1/\lambda}$ (in Theorem. 2).

Let

$$\left.\begin{aligned} &g'_j=g_jI(g_j>c_j),\ U'_n=N^{-1}\sum_{j=1}^{n} I(A_j)g'_j,\\ &g''_{nj}=g_jI(N^{-1}I(A_j)g_j\leqslant d_n^{-1}),\ U''_n=N^{-1}\sum_{j=1}^{n} I(A_j)g''_{nj},\\ &g'''_{nj}=g_j-g'_j-g''_{nj},\ U'''_n=N^{-1}\sum_{j=1}^{n} I(A_j)g'''_{nj}. \end{aligned}\right\} \tag{13}$$

Since $|y|^p$ is a convex function of y for $p>1$, and for fixed $t>0$ and $\lambda>0$, $\exp(ty^\lambda)I(y>a)$ is a convex function of $y\in(a,+\infty)$ for snfficiently large a, from the Jensen's inequality it is easy to see that $E|g(X_1)|^p<\infty$ or $E\{\exp(t|g(X_1)|^\lambda)\}<\infty$ according to the conditions of Theorem 1 or Thesrem 2

$$\sum_j P(g_j>c_j)<\infty,$$

respectively. Hence

and by Borel-Cantelli's lemma, we have

$$P(g_j>c_j,\ \text{i. o.})=0.$$

Therefore

$$\sum_{j=1}^{\infty} g_j^2 I(g_j>c_j)<\infty \qquad \text{a. s.} \tag{14}$$

By lemma 2, we have

$$np_n=n\mu(S_{rh})\geqslant c(x)\cdot n^{\alpha},\ \forall n, \tag{15}$$

in the case of Theorem 1, and

$$np_n\geqslant c(x)(\log n)^{1+\alpha},\ \forall n, \tag{15'}$$

in the case of Theorem 2. Therefore for any $\varepsilon>0$, we have $\frac{1}{n\varepsilon}\leqslant\frac{1}{2}p_n$ for sufficiently large n. By Hoeffding's inequality, we have

$$\begin{aligned}P(N^{-1}>\varepsilon)&=P(N<1/\varepsilon)=P\Big(\frac{N}{n}-p_n<\frac{1}{n\varepsilon}-p_n\Big)\\&\leqslant P\Big(\frac{N}{n}-p_n<-\frac{1}{2}p_n\Big)\leqslant\exp\Big\{-n\Big(\frac{1}{2}p_n\Big)^2\Big/\Big(2p_n+\frac{1}{2}p_n\Big)\Big\}\\&\leqslant\exp\{-np_n/10\}\end{aligned}$$

when n is large enough. Hence, for any $\varepsilon>0$,

$$\sum_{n=1}^{\infty}P(N^{-1}>\varepsilon)<\infty.$$

By Borel-Cantelli's lemma, we have

$$\lim_{n\to\infty} N^{-1}=0 \qquad \text{a. s.} \tag{16}$$

From Schwarz's inequality, noticing (14), we have

$$\begin{aligned}(U_n')^2&\leqslant N^{-2}\sum_{i=1}^{n}I(A_i)\sum_{j=1}^{n}g_j^2I(g_j>c_j)\\&\leqslant\frac{1}{N}I(N>0)\sum_{j=1}^{n}g_j^2I(g_j>c_j)\to 0 \qquad \text{a. s.}\end{aligned} \tag{17}$$

Now, we proceed to prove

$$U_n''\to 0 \qquad \text{a. s.} \tag{18}$$

Let $\widetilde{E}(\cdot)=E(\cdot\,|\,I(A_1),\ \cdots,\ I(A_n))$, $Z_j=d_nI(A_j)N^{-1}[g_{nj}''-E(g_{nj}''|A_j)]$, then $\widetilde{E}Z_j=0$, $Z_j\leqslant 1$, $j=1,\ \cdots,\ n$.

Suppose that $E|Y|^p<\infty$, or $E\{\exp(t|Y|^{\lambda})\}<\infty$, one can take $b<p$, $1<b\leqslant 2$, since $e^z\leqslant 1+z+|z|^b$ when $z\leqslant 1$, we have

$$\widetilde{E}\{\exp(Z_j)\}\leqslant 1+\widetilde{E}|Z_j|^b\leqslant\exp(\widetilde{E}|Z_j|^b).$$

When $I(A_1)$, $\cdots$, $I(A_n)$ are given, Z_1, $\cdots$, Z_n are iid, hence

$$\begin{aligned}&\widetilde{E}\{\exp[d_n(U_n''-\widetilde{E}U_n'')]\}\leqslant\exp\Big\{d_n^bN^{-1}\sum_{j=1}^{n}I(A_j)\widetilde{E}|g_{nj}''-\widetilde{E}g_{nj}''|^b\Big\}\\&\leqslant\exp\Big\{d_n^bN^{-b}\sum_{j=1}^{n}I(A_j)\widetilde{E}g^b(X_j)\Big\}=\exp\Big\{d_n^bN^{-(b-1)}I(N>0)\int_{S_{rh}}g^b(v)\mu(dv)/\mu(S_{rh})\Big\}\\&\leqslant\exp\{d_n^bN^{-(b-1)}I(N>0)g^{b*}(x)\},\end{aligned} \tag{19}$$

where we have written

$$\varphi^*(x)=\sup_{\rho>0}\int_{S_\rho}\varphi(u)\mu(du)/\mu(S_\rho)$$

for any μ-integrable function $\varphi(x)$. Note that by Lemma 1 and the choice of x, we have $g^{b*}(x)<\infty$. By (15) or (15′) and $d_n=\log n\log_2 n$, we know that for any $\varepsilon>0$, there exists $n_0=n_0(x)$ such that

$$d_n^b\left(\frac{2}{np_n}\right)^{b-1} g^{b*}(x)<\frac{1}{2}d_n\varepsilon$$

for $n\geqslant n_0$. So, by Hoeffding's inequality and Markov's inequality, we have

$$\begin{aligned}P(U_n''-\tilde{E}U_n''\geqslant\varepsilon)&\leqslant P\left(N<\frac{1}{2}np_n\right)+P\left(U_n''-\tilde{E}U_n''\geqslant\varepsilon,\ N\geqslant\frac{1}{2}np_n\right)\\&\leqslant P\left(\frac{N}{n}-p_n<-\frac{1}{2}p_n\right)+e^{-d_n\varepsilon}\int_{N\geqslant\frac{1}{2}np_n}\tilde{E}\{\exp[d_n(U_n''-\tilde{E}U_n'')]\}dP\\&\leqslant\exp(-np_n/10)+\exp(-d_n\varepsilon)\cdot\exp\{d_n^b(2/np_n)^{b-1}g^{b*}(x)\}\\&\leqslant\exp(-np_n/10)+\exp\left(-\frac{1}{2}d_n\varepsilon\right).\end{aligned}\tag{20}$$

By (15′) and by the choice of d_n, we get

$$\sum_n P(U_n''-\tilde{E}U_n''\geqslant\varepsilon)<\infty,\ \text{for any } \varepsilon>0.\tag{21}$$

By Borel-Cantelli's lemma

$$\limsup_{n\to\infty}(U_n''-\tilde{E}U_n'')\leqslant 0\qquad\text{a. s.}\tag{22}$$

Recalling (11), we see that $\tilde{E}U_n''\to 0$ a. s., thus,

$$\limsup_{u\to\infty} U_n''\leqslant 0\qquad\text{a. s.,}$$

therefore, (18) is true.

Now we come to prove: $\lim\limits_{n\to\infty} U_n'''=0$ a. s. Putting

$$D_n=\{j:1\leqslant j\leqslant n,\ I(A_j)\cdot N^{-1}>d_n^{-1}\},\tag{23}$$

we have

$$N^{-1}\sum_{j=1}^n I(A_j)g_{nj}'''\leqslant\sum_{j\in D_n}N^{-1}I(A_j)c_j\leqslant N^{-1}c_n\#(D_n).$$

Hence, for any $\varepsilon>0$, in order to get $U_n'''\geqslant\varepsilon$, we must have

$$\#(D_n)\geqslant\varepsilon Nc_n^{-1}.\tag{24}$$

Take a positive integer $k>\dfrac{1}{p\alpha-1}$ in the case of Theorem 1, and $k=1$ in another case, by $c_n=n^{1/p}$ (in Theorem 1) or $c_n=2\cdot(t^{-1}\log n)^{1/\lambda}$ (in Theorem 2), recalling (15) or (15′), we know that there exists $n_0=n_0$ (x) such that

$$N\geqslant\frac{1}{2}np_n,\ U_n'''\geqslant\varepsilon\Rightarrow\#(D_n)\geqslant\varepsilon\frac{1}{2}np_nc_n^{-1}\geqslant k\tag{25}$$

for $n\geqslant n_0$, and that

$$N\geqslant\frac{1}{2}np_n,\ j\in D_n\Rightarrow g_j>\frac{np_n}{2d_n}.\tag{26}$$

Hence, when $n\geqslant n_0$, we get

$$P(U_n'''\geqslant\varepsilon)\leqslant P\left(N<\frac{1}{2}np_n\right)+P\left(N\geqslant\frac{1}{2}np_n,\ U_n'''\geqslant\varepsilon\right)$$
$$\leqslant P\left(N<\frac{1}{2}np_n\right)+P\left(\#\left\{j:j\leqslant n,\ g_j>\frac{np_n}{2d_n}\right\}>k\right). \tag{27}$$

In the case of Theorem 1, since $g_1, \cdots, g_n$ are iid and $np_n\geqslant c(x)n^\alpha$, we have

$$P(U_n'''\geqslant\varepsilon)\leqslant P\left(N<\frac{1}{2}np_n\right)+\left\{\sum_{j=1}^n P\left(g_j>\frac{np_n}{2d_n}\right)\right\}^k$$
$$\leqslant P\left(N<\frac{1}{2}np_n\right)+\left\{n\left(\frac{2d_n}{np_n}\right)^p E|g_1|^p\right\}^k$$
$$\leqslant P\left(N<\frac{1}{2}np_n\right)+\{n\,d_n^pn^{-p\alpha}c\,(x)\}^k$$
$$\leqslant\exp(-np_n/10)+c(x)d_n^{pk}n^{-(p\alpha-1)k},\ n\geqslant n_0. \tag{28}$$

In the case of Theorem 2, from (15′), there exsits $n_0'=n_0'(x)$ such that

$$t\left|\frac{np_n}{2d_n}\right|^\lambda\geqslant 3\log n \text{ for } n\geqslant n_0'.$$

Choosing $k=1$ in (27), we get

$$P(U_n'''\geqslant\varepsilon)\leqslant P\left(N<\frac{1}{2}np_n\right)+nP\left(g_1>\frac{np_n}{2d_n}\right)$$
$$\leqslant P\left(N<\frac{1}{2}np_n\right)+n\cdot\exp\left(-t\left|\frac{np_n}{2d_n}\right|^\lambda\right)E[\exp(t|g_1|^\lambda)]$$
$$\leqslant\exp(-np_n/10)+c(x)n\exp(-3\log n)$$
$$\leqslant\exp(-np_n/10)+c(x)n^{-2},\ n\geqslant n_0'. \tag{28′}$$

From (28) or (28′), we get

$$\sum_n P(U_n'''\geqslant\varepsilon)<\infty, \text{ for any } \varepsilon>0. \tag{29}$$

By Borel-Cantelli's lemma and $U_n'''\geqslant 0$, we have

$$\lim_{n\to\infty}U_n'''=0 \quad \text{a. s.} \tag{30}$$

From (17), (18), (30) and $U_n(x)=U_n'+U_n''+U_n'''$, we know that (12) is true.

In the following we prove that

$$T_n(x)\triangleq\sum_{j=1}^n W_{nj}(x)(Y_j-m(X_j))\to 0 \quad \text{a. s.} \tag{31}$$

as $n\to\infty$, where $W_{nj}(x)=K((X_j-x)/h)\Big/\sum_{i=1}^n K((X_i-x)/h)$ has been defined in (2).

Write $e_j=Y_j-m(X_j)$, take c_j, d_n as above, and let

$$\left.\begin{aligned}&e_j'=e_jI(e_j>c_j);\ T_n'=\sum_{j=1}^n W_{nj}e_j',\\&e_{nj}''=e_jI(W_{nj}e_j\leqslant d_n^{-1}),\ T_n''=\sum_{j=1}^n W_{nj}e_{nj}'',\\&e_{nj}'''=e_j-e_j'-e_j'',\ T_n'''=\sum_{j=1}^n W_{nj}e_{nj}'''.\end{aligned}\right\} \tag{32}$$

Similarily, we get

$$\sum_j P(e_j>c_j)\leqslant\sum_j P\left(|Y_j|<\frac{1}{2}c_j\right)+\sum_j P\left(|m(X_j)|>\frac{1}{2}c_j\right)<\infty,$$

$$P(e_j>c_j,\ \text{i. o.})=0,$$

$$\sum_{j=1}^{\infty} e_j^2 I(e_j>c_j)<\infty \quad \text{a. s.}$$

By Schwarz's inequality and (16), we have

$$(T'_n)^2\leqslant\sum_{j=1}^{n} W_{nj}^2(x)\sum_{j=1}^{\infty} e_j^2 I(e_j>c_j)$$
$$\leqslant(\tilde{c}_2/\tilde{c}_1)^2 N^{-1} I(N>0)\sum_{j=1}^{\infty} e_j^2 I(e_j>c_j)\xrightarrow{\text{a.s.}} 0. \tag{33}$$

Writing $Z_j=d_n W_{nj} e''_{nj}$, we have $Z_j\leqslant 1$. When $\Delta=(X_1, X_2, \cdots)$ is given, $Z_1, \cdots, Z_n$ are conditionally independent, and $E(Z_j|X_j)\leqslant 0$.

Assume that $1<b\leqslant 2$, write $g_b(X_j)=E(|Y_j|^b|X_j)$, then we have as before

$$E(e^{Z_j}|X_j)\leqslant 1+E(|Z_j|^b|X_j)\leqslant\exp\{E(|Z_j|^b|X_j)\},$$

$$E\{\exp(d_n T''_u)\,|\,\Delta\}\leqslant\exp\left\{cd_n^b\sum_{j=1}^{n} W_{nj}^b g_b(X_j)\right\}$$
$$\leqslant\exp\left\{cd_n^b N^{-b}\sum_{j=1}^{n} I(A_j) g_b(X_j)\right\}. \tag{34}$$

Suppose $p>1$, by $\alpha>p^{-1}$, one can take $b>1$ sufficiently close to 1 such that $q=p/b>1$ and $\alpha>1/q$. Suppose $\lambda>0$, by $\alpha>\frac{1}{\lambda}$, one can take $b>1$ sufficiently close to 1 such that the inequality $\alpha>1/\nu$ is valid for $\nu=\lambda/b$. In this case, $E|g_b(X_j)|^q\leqslant E|Y_j|^p<\infty$, and $1/q<\alpha\leqslant 1$ for the occasion of Theorem 1. For the case of Theorem 2, we have $E\{\exp(t|g_b(X_j)|^\nu)\}<\infty$, and $\alpha>1/\nu$. Therefore, from the proof of (12), we obtain

$$N^{-1}\sum_{j=1}^{n} I(A_j) g_b(X_j)\to g_b(x) \quad \text{a. s}, \tag{35}$$

as $n\to\infty$. Writing $\mu^\infty=\mu\times\mu\times\cdots$, we have

$$cN^{-1}\sum_{j=1}^{n} I(A_j) g_b(X_j)\leqslant c(x, \Delta) \qquad \text{a. e. } \Delta(\mu^\infty). \tag{36}$$

Just as before, from Hoeffding's inequality

$$\sum_n P\left(N\leqslant\frac{1}{2}np_n\right)<\infty.$$

By Borel-Cantelli's lemma

$$P\left(N\leqslant\frac{1}{2}np_n,\ \text{i. o.}\right)=0. \tag{37}$$

Hence, when n is sufficiently large

$$N>\frac{1}{2}np_n, \quad \text{a. e. } \Delta(\mu^\infty). \tag{38}$$

From (34), recalling $np_n\geqslant c(x)(\log n)^{1+\alpha}$, $d_n=\log n\log_2 n$, for a. e. $\Delta(\mu^\infty)$ and given $\varepsilon>0$, there exists $n_0=n_0(x, \Delta)$ such that

$$P(T''_n \geqslant \varepsilon \mid \Delta) \leqslant e^{-d_n\varepsilon} E\{\exp(d_n T''_n) \mid \Delta\} \leqslant e^{-d_n\varepsilon} \exp\{d_n^b N^{-(b-1)} I(N>0) c(x, \Delta)\}$$
$$\leqslant e^{-d_n\varepsilon} \exp\left\{d_n^b \left(\frac{2}{np_n}\right)^{b-1} c(x, \Delta)\right\} \leqslant e^{-d_n\varepsilon} e^{d_n\varepsilon/2} = \exp\left(-\frac{1}{2} d_n \varepsilon\right) \quad (39)$$

for $n \geqslant n_0$. Here, we get the second inequality by (36). By (38) we get the third one. From (39), we have

$$\sum_n P(T''_n \geqslant \varepsilon \mid \Delta) < \infty, \text{ for any } \varepsilon > 0. \quad (40)$$

Hence, by Borel-Cantelli's lemma, we have

$$P(\limsup_{n\to\infty} T''_n > 0 \mid \Delta) = 0, \text{ for a.e. } \Delta(\mu^\infty). \quad (41)$$

By Fubini's theorem, we have

$$P(\limsup_{n\to\infty} T''_n > 0) = 0, \quad (42)$$

i. e,

$$\limsup_{n\to\infty} T''_n \leqslant 0 \quad \text{a. s.,} \quad (43)$$

For the proof of $\limsup_{n\to\infty} T'''_n \leqslant 0$ a. s., write

$$D_n = \{j: 1 \leqslant j \leqslant n,\ W_{nj} e_j > d_n^{-1}\}. \quad (44)$$

Then

$$\sum_{j=1}^n W_{nj} e'''_{nj} \leqslant \sum_{j\in D_n} (\tilde{c}_2/\tilde{c}_1) N^{-1} I(A_j) c_j \leqslant (\tilde{c}_2/\tilde{c}_1) N^{-1} c_n \#(D_n).$$

Therefore, in order to get $T'''_n \geqslant \varepsilon$ for any $\varepsilon > 0$, we must have

$$\#(D_n) \geqslant \tilde{c}_1 \tilde{c}_2^{-1} \varepsilon N c_n^{-1}. \quad (45)$$

Taking a positive integer $k > \frac{1}{p\alpha - 1}$, just as before, there exists $n_0 = n_0(x)$ such that

$$N \geqslant \frac{1}{2} np_n, \quad T'''_n \geqslant \varepsilon \Rightarrow \#(D_n) \geqslant k \quad (46)$$

for $n \geqslant n_0$, and

$$N \geqslant \frac{1}{2} np_n,\ j \in D_n \Rightarrow e_j > \tilde{c}_1 np_n / (2 d_n \tilde{c}_2) \triangleq c np_n / d_n. \quad (47)$$

Hence, when $n \geqslant n_0$, we have

$$P(T'''_n \geqslant \varepsilon) \leqslant P\left(N < \frac{1}{2} np_n\right) + P\left(N > \frac{1}{2} np_n,\ T'''_n \geqslant \varepsilon\right)$$
$$\leqslant P\left(N < \frac{1}{2} np_n\right) + P(\#\{j: j \leqslant n,\ e_j > c np_n/d_n\} > k). \quad (48)$$

In the case of Theorem 1, since $e_1, \cdots, e_n$ are iid, and $np_n \geqslant c(x) n^\alpha$, we have

$$P(T'''_n \geqslant \varepsilon) \leqslant P\left(N < \frac{1}{2} np_n\right) + \left\{\sum_{j=1}^n P(e_j > c np_n/d_n)\right\}^k$$
$$\leqslant P\left(N < \frac{1}{2} np_n\right) + \{n (d_n/(c np_n))^p E|e_1|^p\}^k$$
$$\leqslant P\left(N < \frac{1}{2} np_n\right) + \left\{n d_n^p n^{-p\alpha} c(x)\right\}^k$$
$$\leqslant \exp(-np_n/10) + c(x) d_n^{pk} n^{-(p\alpha-1)k}, \quad n \geqslant n_0. \quad (49)$$

In the case of Theorem 2, take $k=1$ in (48). Similar to the proof of (28), there exists $n' = n'_0(x)$ such that

$$P(T_n'''' \geqslant \varepsilon) \leqslant \exp(-np_n/10) + c(x)n^{-2} \quad (49')$$

for $n > n_0'$. From (49) and (49'), we get

$$\sum_n P(T_n'''' \geqslant \varepsilon) < \infty \text{ for any } \varepsilon > 0. \quad (50)$$

Hence, by Borel-Cantelli's lemma, we have

$$\limsup_{n\to\infty} T_n'''' \leqslant 0 \quad \text{a. s.} \quad (51)$$

From (33), (43), (51) and $T_n(x) = T_n' + T_n'' + T_n'''$, we have

$$\limsup_{n\to\infty} T_n(x) \leqslant 0 \quad \text{a. s.}, \quad (52)$$

Replacing e_j by $-e_j$, (52) implies

$$\liminf_{n\to\infty} T_n(x) \geqslant 0 \quad \text{a. s.}. \quad (53)$$

From (52) and (53), we get (31).

Finally by (8), (9), (12) and (31), noticing

$$|m_n(x) - m(x)| \leqslant \left|\sum_{i=1}^n W_{ni}(x)(m(X_i) - m(x))\right| + \left|\sum_{i=1}^n W_{ni}(x)(Y_i - m(X_i))\right|$$
$$= J_n(x) + |T_n(x)|, \quad (54)$$

for $x \in F$, we have

$$\lim_{n\to\infty} |m_n(x) - m(x)| = 0 \quad \text{a. s.} \quad (55)$$

Since F is the support of μ, we have $\mu(F) = 1$. By Fubini's theorem, we get (4). This completes the proof of Theorem 1 and 2.

References

[1] Devroye, L., On the Almost Everywhere Convergence of Nonparametric Regression Function Estimates, *Ann. Statist.*, **9** (1981), 1310—1319.

[2] Wheeden, R. L. & Zygmund, A., Measure and Integral, Marcel Dekker, New York, 1977.

[3] Hoeffding, W., Probability Inequalities for sums of Bounded Random Variables, *J. Amer. Statist. Assoc.*, **58** (1963), 13—30.