JOURNAL OF MULTIVARIATE ANALYSIS 26, 184–206 (1988)

Asymptotics of Conditional Empirical Processes

LAJOS HORVÁTH*

Bolyai Institute, Szeged University, Szeged, Hungary

AND

BRIAN S. YANDELL†

University of Wisconsin–Madison

Communicated by M. Csörgő

Using two definitions of the conditional empirical processes we obtain some approximations for these processes. We also prove the functional law of the iterated logarithm for the conditional processes. Our results say that the asymptotic behavior of the conditional and unconditional empirical processes are very similar.

1. INTRODUCTION

We investigate asymptotic approximations of two estimators of the conditional empirical process of Y given $X = x$, leading to a functional law of the iterated logarithm. Our main results show that the weak behavior of the conditional empirical processes at a given x is essentially the same as that of the empirical process, although the rates now depend on the bandwidth.

We examine the kernel estimator of Nadaraya [20] and Watson [29] and the nearest neighbor (NN) estimator of Yang [31] and others. The subject of nonparametric regression was reviewed by Stone [24, 25] and Collomb [4]. Recently Révész [22], Johnston [12], Liero [17], and Konakov and Piterbarg [15] developed strong approximations of residual mean regression functions. Mack and Silverman [18] and Cheng [2]

Received February 10, 1987; revised October 5, 1987.

AMS 1980 subject classifications: Primary 60F17; Secondary 60F15.

Key words and phrases: conditional distribution, multivariate Wiener process, stochastic integrals, U-statistics, functional laws of iterated logarithm.

* Research done at Carleton University, Ottawa, Canada, supported by NSERC Canada and EMR Canada grants of M. Csörgő.

† Research supported by USDA-CSRS Grant 511-100 and NSF Grant DMS 8404970.

0047-259X/88 $3.00

derived uniform convergence rates for kernel estimates. Stute [26, 27] proved almost sure and weak convergence results for kernel and nearest neighbor estimates of the conditional empirical function. Several authors have recently explored robust nonparametric regression (cf. Cheng and Cheng [3] and Härdle [11]).

Let (Y, X) be a random vector in $\mathbf{R}^2$ with continuous distribution function $F(y, x) = P(Y \leqslant y, X \leqslant x\}$ and marginals $H(y) = P\{Y \leqslant y\}$, $G(x) = P\{X \leqslant x\}$. In statistical applications one may need an estimator of the conditional distribution function

$$m(y \mid x) = P\{Y \leqslant y \mid X = x\}, \qquad y \in \mathbf{R},$$

at a fixed point $x \in \mathbf{R}$. Based on the random i.i.d. sample $\{(Y_i, X_i)\}_{i=1}^n$, the usual estimators for F and G are

$$F_n(y, x) = n^{-1} \# \{1 \leqslant i \leqslant n \colon Y_i \leqslant y, X_i \leqslant x\}$$

and

$$G_n(x) = F_n(\infty, x) = n^{-1} \# \{1 \leqslant i \leqslant n \colon X_i \leqslant x\}.$$

(Throughout this paper, $\mu(\infty) = \lim \mu(x)$, $x \to \infty$.)

Our first estimator for $m(y \mid x)$ is essentially due to Nadaraya [20] and Watson [29]. Let K be an appropriate kernel function, a_n be a sequence of bandwidths and define

$$h_n(y, x) = a_n^{-1} \int_{-\infty}^{\infty} K\left(\frac{x-u}{a_n}\right) d_u F_n(y, u).$$

(Integrals are over $\mathbf{R}$ unless otherwise noted.) Standard arguments involving multivariate densities show that $h_n(y, x)$ is a reasonable estimate for

$$h(y, x) = \frac{\partial}{\partial x} F(y, x).$$

We assume that $g = G'$ exists in a neighborhood of x and estimate it by

$$g_n(x) = h_n(\infty, x) = a_n^{-1} \int K\left(\frac{x-u}{a_n}\right) dG_n(u).$$

The estimator

$$\begin{aligned} m_n(y \mid x) &= h_n(y, x)/g_n(x) \\ &= \sum_{i=1}^n I\{Y_i \leqslant y\} K\left(\frac{x - X_i}{a_n}\right) \Big/ \sum_{i=1}^n K\left(\frac{x - X_i}{a_n}\right), \end{aligned}$$

where $I(A)$ is the indicator of the set A, will be referred to as a kernel-type estimator of $m(y \mid x)$.

A different estimator is proposed by Yang [31] and Stute [28, 26]. Let

$$r_n(y \mid x) = a_n^{-1} \int K\left(\frac{G_n(x) - G_n(u)}{a_n}\right) d_u F_n(y, u)$$

and

$$l_n(x) = r_n(\infty \mid x) = a_n^{-1} \int K\left(\frac{G_n(x) - G_n(u)}{a_n}\right) dG_n(u).$$

The estimator

$$\begin{aligned} k_n(y \mid x) &= r_n(y \mid x)/l_n(x) \\ &= \sum_{i=1}^{n} I\{Y_i \leq y\}\, K\left(\frac{G_n(x) - G_n(X_i)}{a_n}\right) \Big/ \sum_{i=1}^{n} K\left(\frac{G_n(x) - G_n(X_i)}{a_n}\right) \end{aligned}$$

is called a nearest-neighbor (NN) type estimator.

The central part of this paper is the almost sure behavior of the processes

$$\beta_n(y \mid x) = (na_n)^{1/2}\,(m_n(y \mid x) - m_{(n)}(y \mid x)) \tag{1.1}$$

and

$$\gamma_n(y \mid x) = (na_n)^{1/2}\,(k_n(y \mid x) - k_{(n)}(y \mid x)), \tag{1.2}$$

where

$$m_{(n)}(y \mid x) = h_{(n)}(y, x)/g_{(n)}(x),$$

$$k_{(n)}(y \mid x) = a_n^{-1} \int K\left(\frac{G(x) - G(u)}{a_n}\right) d_u F(y, u),$$

and $g_{(n)}$ and $h_{(n)}$ are defined analogously to g_n and h_n, respectively, with F_n and G_n replaced by F and G. Standard methods can give the usual rates for the "numerical errors" $m_{(n)} - m$ and $k_{(n)} - m$ by assuming more regularity conditions on F. Thus $m_{(n)}$ and $k_{(n)}$ can be replaced by m in (1.1) and (1.2) and our results will remain true (cf. Lemmas 3.1 and 4.1).

We assume that $x \in \mathbf{R}$ is a fixed point throughout this paper. It is well known that $\beta_n(y \mid x)$ and $\gamma_n(y \mid x)$ have normal limits for any fixed $y \in \mathbf{R}$ as $n \to \infty$. Stute [26] proves that $\gamma_n(y \mid x)$ converges weakly to a time-transformed Brownian bridge. The main aim of this paper is a further investigation of the asymptotic properties of the conditional empirical processes $\beta_n(y \mid x)$ and $\gamma_n(y \mid x)$. We obtain rates for the Gaussian approximations and derive functional laws of the iterated logarithm for these processes. The proofs are based on the observation that β_n and γ_n can be expressed as integrals with respect to

$$\alpha_n(y, x) = n^{1/2}(F_n(y, x) - F(y, x)),$$
$$t_n(x) = \alpha_n(\infty, x) = n^{1/2}(G_n(x) - G(x)),$$

with a small remainder term.

In Section 2 we list some results which will be very useful later on. Section 3 contains the approximation of β_n and in Section 4 we prove a similar result for γ_n. The functional laws of the iterated logarithm will be proven in Section 5.

Throughout this paper, C in proofs stands for a generic constant whose value may differ from line to line. Suprema and integrals are over $\mathbf{R}$ unless otherwise stated. We can assume without loss of generality that our probability space $(\Omega, A, \mathbf{P})$ is so rich that all the random variables and processes introduced so far or later on can be defined on it (cf. de Acosta [8]).

2. Preliminaries

This section contains some approximations for α_n, t_n and a symmetric kernel based on results in the literature. The following multivariate approximation is due to Borisov [1].

THEOREM A. *We can define a sequence of Gaussian processes* $\{\Gamma_n^{(1)}(y, x),\ -\infty < y, x < \infty\}$ *such that*

$$P\{\sup_{y,x} |\alpha_n(y, x) - \Gamma_n^{(1)}(y, x)| > C_1 n^{-1/6} \log n\} \leqslant C_2 n^{-2}$$

and $E\Gamma_n^{(1)}(y, x) = 0$,

$$E\Gamma_n^{(1)}(y, x)\, \Gamma_n^{(1)}(y', x') = F(y \wedge y', x \wedge x') - F(y, x)\, F(y', x').$$

(Here $a \wedge b = \min(a, b)$.) Using Komlós, Major, and Tusnády [14] and Lemma 1.1.1 of Csörgő and Révész [5] one can establish the next inequality.

THEOREM B. *If G is continuous in a neighborhood of x, then*

$$P\{\sup_{|G(x) - G(u)| \leqslant d} |t_n(x) - t_n(u)| > C_3(n^{-1/2} \log n + (d \log n)^{1/2})\}$$
$$\leqslant n^{-2} + d^{-1} n^{-4},$$

for all $0 < d < 1$.

Let $\mu(\mathbf{z}_1;\mathbf{z}_2)$, $\mathbf{z}_1,\mathbf{z}_2\in\mathbf{R}^2$, be a symmetric function, i.e., $\mu(\mathbf{z}_1;\mathbf{z}_2)=\mu(\mathbf{z}_2;\mathbf{z}_1)$ for all $\mathbf{z}_1,\mathbf{z}_2\in\mathbf{R}^2$. Assume that

$$E\mu(Y_1,X_1;Y_2,X_2)=0 \tag{2.1}$$

and

$$E\{\mu(Y_1,X_1;Y_2,X_2)\mid Y_1,X_1\}=0 \quad \text{a.s.} \tag{2.2}$$

Using the symmetric kernel μ we can define a multivariate V-statistic by

$$U_n=\sum_{1\leqslant i,j\leqslant n}\mu(Y_i,X_i;Y_j,X_j).$$

Condition (2.2) implies that U_n is essentially a degenerate U-statistic. For a survey on U- and V-statistics, see Serfling [23, Chap. 5]. The following result is essentially due to Dehling, Denker, and Philipp [9]. They proved the result for a bivariate kernel μ. It is easy to see that their method also works for a multivariate kernel.

THEOREM C. *Assume that* (2.1) *and* (2.2) *hold, and*

$$|\mu(\mathbf{z}_1;\mathbf{z}_2)|\leqslant C_4. \tag{2.3}$$

Then for all positive integers v,

$$E(U_n)^{2v}\leqslant n^{2v}\{4v^{2v}(4C_4/e)^{2v}+n^{-1}v^{2v+1}(80/e)^{2v+1}C_4^{2v}\}.$$

Theorem A is tailored for getting a rate in weak invariance principles but it is not enough to obtain exact laws of the iterated logarithm. In Section 5 we need the following result.

THEOREM D. *We can define a Gaussian process* $\{\Gamma^{(2)}(y,x,t),\ -\infty<y,\ x<\infty,\ t>0)$ *such that*

$$\sup_{y,x}|\alpha_n(y,x)-n^{-1/2}\Gamma^{(2)}(y,x,n)| \overset{\text{a.s.}}{=} o(n^{-\lambda})$$

with some $\lambda>0$ *and* $E\Gamma^{(2)}(y,x,t)=0$,

$$E\Gamma^{(2)}(y,x,t)\,\Gamma^{(2)}(y',x',t')=(t\wedge t')(F(y\wedge y',x\wedge x')-F(y,x)\,F(y',x')).$$

Csörgő and Révész [5] obtained a similar result with $\lambda<\frac{5}{16}$ in the case of a smooth distribution function. Theorem D was proven by Philipp and Pinzur [21] with $\lambda<\frac{1}{20000}$ without assuming any regularity conditions on F. Recently Csörgő and Horváth [6] improved the Philipp–Pinzur result to $\lambda<\frac{1}{8}$ without assuming any regularity conditions on F.

3. Kernel Type Estimator

We now show that when the conditions below hold we can approximate β_n using a sequence of Brownian bridges. From this we have immediately a strong law. The following conditions which we will use are not very restrictive, requiring some smoothness and an appropriate rate for the bandwidth.

$$g(x) > 0 \tag{3.1}$$

$$\sup_{y,u} h(y, u) < \infty \tag{3.2}$$

$$\sup_{y,u} \left| \frac{\partial}{\partial u} h(y, u) \right| < \infty, \qquad \sup_u |g'(u)| < \infty \tag{3.3}$$

$$K \geqslant 0, \quad \lim_{u \to \pm\infty} K(u) = 0, \qquad \int K(u)\, du = 1 \tag{3.4}$$

$$\int |u|\, K^2(u)\, du < \infty \tag{3.5}$$

$$\int K^2(u)\, du < \infty \tag{3.6}$$

$$K \text{ has bounded variation on } \mathbf{R} \tag{3.7}$$

and

$$a_n \log n \to 0 \qquad \text{as} \quad n \to \infty \tag{3.8}$$

$$n a_n^3 (\log n)^{-6} \to \infty \qquad \text{as} \quad n \to \infty. \tag{3.9}$$

If we assume that K vanishes outside of a finite interval, then (3.3) can be replaced with

$$|g'(u)| < \infty, \qquad \sup_y \left| \frac{\partial}{\partial u} h(y, u) \right| < \infty$$

uniformly in u in a neighborhood of x.

THEOREM 3.1. *We assume that* (3.1)–(3.9) *hold. We can define a sequence of Brownian bridges* $\{B_n^{(1)}(t),\ 0 \leqslant t \leqslant 1\}$ *such that*

$$P\left\{ \sup_y \left| \beta_n(y \mid x) - \left(\int K^2(u)\, du / g(x) \right)^{1/2} B_n^{(1)}(m(y \mid x)) \right| \right.$$
$$\left. > C_5((a_n \log n)^{1/2} + n^{-1/6} a_n^{-1/2} \log n) \right\} \leqslant C_6 n^{-2}.$$

Theorem 3.1 and the Borel–Cantelli lemma immediately imply the following strong law.

COROLLARY 3.1. *If* (3.1)–(3.9) *hold, then*

$$\limsup_{n\to\infty}(\log n)^{-1/2}\sup_y|\beta_n(y\mid x)|\leqslant 2^{-1/2}\left(\int K^2(u)\,du/g(x)\right)^{1/2}\qquad a.s.$$

Assuming more regularity conditions we improve Corollary 3.1 in Section 5 and prove a law of the iterated logarithm.

The results of this section will hold for $(na_n)^{1/2}\,(m_n(y\mid x)-m(y\mid x))$ if we can prove that $m_{(n)}(y\mid x)-m(y\mid x)$ is asymptotically negligible. This is done in the following lemma, whose proof is immediate using a two-term Taylor expansion.

LEMMA 3.1. *We assume that* (3.1)–(3.4) *hold*, $\int u^2K(u)\,du<\infty$ *and*

$$\sup_{y,u}\left|\frac{\partial^2}{\partial u^2}h(y,u)\right|<\infty.$$

Then

$$\sup_y|m_{(n)}(y\mid x)-m(y\mid x)|=O(a_n^3).$$

If we assume that K vanishes outside a finite interval, then the extra condition on h in Lemma 3.1 can be replaced by

$$\sup_y\left|\frac{\partial^2}{\partial u^2}h(y,u)\right|<\infty$$

uniformly in u in a neighborhood of x.

In order to prove Theorem 3.1, we first approximate β_n by stochastic integrals involving α_n and t_n in the following lemma.

LEMMA 3.2. *If* (3.1)–(3.9) *hold, then*

$$\beta_n(y\mid x)=\frac{a_n^{-1/2}}{g(x)}\int K\left(\frac{x-u}{a_n}\right)d_u\alpha_n(y,u)$$
$$-\frac{a_n^{-1/2}m(y\mid x)}{g(x)}\int K\left(\frac{x-u}{a_n}\right)dt_n(u)+R_n^{(1)}(y,x)$$

and

$$P\{\sup|R_n^{(1)}(y,x)|>C_7((a_n\log n)^{1/2}+n^{-1/2}a_n^{-3/2}\log n)\}\leqslant C_8n^{-2}.$$

Proof. First we note that

$$\begin{aligned}\beta_n(y \mid x) = & \frac{a_n^{-1/2}}{g_{(n)}(x)} \int K\left(\frac{x-u}{a_n}\right) d_u \alpha_n(y, u) \\ & - \frac{a_n^{-1/2} h_{(n)}(y, x)}{g_{(n)}^2(x)} \int K\left(\frac{x-u}{a_n}\right) dt_n(u) \\ & - \frac{n^{-1/2} a_n^{-3/2}}{g^{(n)}(x)\, g_n(x)} \int K\left(\frac{x-u}{a_n}\right) d_u \alpha_n(y, u) \int K\left(\frac{x-u}{a_n}\right) dt_n(u) \\ & + \frac{n^{-1/2} a_n^{-3/2} h_{(n)}(y, x)}{g_{(n)}^2(x)\, g_n(x)} \left(\int K\left(\frac{x-u}{a_n}\right) dt_n(u)\right)^2. \end{aligned} \tag{3.10}$$

By (3.3)–(3.5) we have

$$\sup_y |h_{(n)}(y, x) - h(y, x)| = O(a_n) \tag{3.11}$$

and

$$|g_{(n)}(x) - g(x)| = O(a_n). \tag{3.12}$$

Integration by parts and Kiefer's [13] inequality imply that

$$\begin{aligned} & P\left\{\sup_y \left|\int K\left(\frac{x-u}{a_n}\right) d_u \alpha_n(y, u)\right| > C(\log n)^{1/2}\right\} \\ & \quad = P\left\{\sup_y \left|\int \alpha_n(y, x - t a_n)\, dK(t)\right| > C(\log n)^{1/2}\right\} \\ & \quad \leqslant P\left\{\sup_{y, u} |\alpha_n(y, u)| \int d\,|K(t)| > C(\log n)^{1/2}\right\} \\ & \quad \leqslant C n^{-2}, \end{aligned} \tag{3.13}$$

and similarly

$$P\left\{\left|\int K\left(\frac{x-u}{a_n}\right) dt_n(u)\right| > C(\log n)^{1/2}\right\} \leqslant C n^{-2}. \tag{3.14}$$

Using now (3.11)–(3.14), we obtain that

$$\begin{aligned} & P\left\{a_n^{-1/2} \left|\frac{1}{g_{(n)}(x)} - \frac{1}{g(x)}\right| \sup_y \left|\int K\left(\frac{x-u}{a_n}\right) d_u \alpha_n(y, u)\right|\right. \\ & \quad \left. > C(a_n \log n)^{1/2}\right\} \leqslant C n^{-2} \end{aligned} \tag{3.15}$$

and, for the second term in (3.10),

$$P\left\{\sup_y \left|\frac{h_{(n)}(y,x)}{g_{(n)}^2(x)} - \frac{m(y\mid x)}{g(x)}\right| a_n^{-1/2} \left|\int K\left(\frac{x-u}{a_n}\right) dt_n(u)\right| \right.$$
$$\left. > C(a_n \log n)^{1/2}\right\} \leqslant Cn^{-2}. \tag{3.16}$$

Observing that

$$g_n(x) - g_{(n)}(x) = a_n^{-1} \int K\left(\frac{x-u}{a_n}\right) d(G_n(u) - G(u))$$

we get from (3.1), (3.9), (3.13), and (3.14) that

$$P\{g_n(x) > C\} \leqslant Cn^{-2}. \tag{3.17}$$

Hence (3.1), (3.14), and (3.17) yield

$$P\left\{\frac{n^{1/2}a_n^{-3/2}}{g_{(n)}(x)\, g_n(x)} \left|\int K\left(\frac{x-u}{a_n}\right) dt_n(u)\right| \sup_y \left|\int K\left(\frac{x-u}{a_n}\right) d_u\alpha_n(y,u)\right| \right.$$
$$\left. > Cn^{-1/2}a_n^{-3/2} \log n\right\} \leqslant Cn^{-2} \tag{3.18}$$

and

$$P\left\{\sup_y \frac{n^{1/2}a_n^{-3/2}h_{(n)}(y,x)}{g_{(n)}^2(x)\, g_n(x)} \left(\int K\left(\frac{x-u}{a_n}\right) dt_n(u)\right)^2 \right.$$
$$\left. > Cn^{-1/2}a_n^{-3/2} \log n\right\} \leqslant Cn^{-2}. \tag{3.19}$$

Lemma 3.2 now follows from (3.15), (3.16), (3.18), and (3.19).

If we assume that

$$\inf_{t_1 \leqslant x \leqslant t_2} g(x) > 0$$

instead of (3.1), then Lemma 3.2 holds uniformly in $x \in [t_1^*, t_2^*]$ for $t_1 < t_1^* \leqslant t_2^* < t_2$.

Proof of Theorem 3.1. By Lemma 3.2 an approximation of

$a_n^{-1/2} \int K((x-u)/a_n)\, d_u\alpha_n(y, u)$ results in an approximation of β_n. Now applying Theorem A we have

$$P\left\{a_n^{-1/2} \sup_y \left|\int K\left(\frac{x-u}{a_n}\right) d_u(\Gamma_n^{(1)}(y, u) - \alpha_n(y, u))\right| > Ca_n^{-1/2}n^{-1/6} \log n\right\} \leqslant Cn^{-2}. \tag{3.20}$$

It is well known (cf. Wichura [30]) that there is a two-dimensional distribution function J with uniform marginals such that

$$F(y, x) = J(H(y), G(x)). \tag{3.21}$$

Let $\{W_J(s, t),\ 0 \leqslant s,\ t \leqslant 1\}$ be a two-dimensional Wiener process with $EW_J(s, t) = 0$ and $EW_J(s, t)\, W_J(s', t') = J(s \wedge s', t \wedge t')$. Then for each n we have

$$\{\Gamma_n^{(1)}(y, x), y, x \in \mathbf{R}\} \stackrel{D}{=} \{W_J(H(y), G(x)) - F(y, x)\, W_J(1, 1), y, x \in \mathbf{R}\}, \tag{3.22}$$

and

$$\left\{\int K\left(\frac{x-u}{a_n}\right) d_u\Gamma_n^{(1)}(y, u), y, x \in \mathbf{R}\right\}$$
$$\stackrel{D}{=} \left\{\int K\left(\frac{x-u}{a_n}\right) d_u W_J(H(y), G(u)) - W_J(1, 1) \int K\left(\frac{x-u}{a_n}\right) h(y, u)\, du, y, x \in \mathbf{R}\right\}.$$

Now applying (3.2) and (3.4) we get that

$$P\left\{|W_J(1, 1)| \sup_y \left|\int K\left(\frac{x-u}{a_n}\right) h(y, u)\, du\right| > Ca_n(\log n)^{1/2}\right\} \leqslant Cn^{-2}. \tag{3.24}$$

Elementary calculations show that

$$E \int K\left(\frac{x-u}{a_n}\right) d_u W_J(H(y), G(u)) = 0$$

and

$$Ea_n^{-1} \int K\left(\frac{x-u}{a_n}\right) d_u W_J(H(y), G(u)) \int K\left(\frac{x-u}{a_n}\right) d_u W_J(H(y'), G(u))$$

$$= a_n^{-1} \int K^2\left(\frac{x-u}{a_n}\right) h(y \wedge y', u)\, du = l_n^*(y \wedge y').$$

Thus we get

$$\left\{a_n^{-1/2} \int K\left(\frac{x-u}{a_n}\right) d_u W_J(H(y), G(u)), y \in \mathbf{R}\right\} \stackrel{D}{=} \{W(l_n^*(y)), y \in \mathbf{R}\}, \quad (3.25)$$

where $\{W(t), 0 \leqslant t < \infty\}$ is a standard Wiener process. Let $l(y) = h(y, x) \int K^2(u)\, du$. A one-term Taylor expansion and (3.3), (3.5) imply that

$$\sup_y |l_n^*(y) - l(y)| = \sup_y \left|\int K^2(u)\{h(y, x - ua_n) - h(y, x)\}\, du\right| = O(a_n). \quad (3.26)$$

Combining (3.26) and Lemma 1.1.1 of Csörgő and Révész [5] we get that

$$P\{\sup_y |W(l_n^*(y)) - W(l(y))| > C(a_n \log n)^{1/2}\} \leqslant Cn^{-2}. \quad (3.27)$$

By Lemma 3.2, (3.20), (3.23)–(3.25), and (3.27) we have that

$$\{\beta_n(y \mid x), y \in \mathbf{R}\}$$
$$\stackrel{D}{=} \left\{\frac{1}{g(x)} W(l(y)) - \frac{m(y \mid x)}{g(x)} W(l(\infty)) + R_n^{(2)}(y), y \in \mathbf{R}\right\} \quad (3.28)$$

and

$$P\{\sup_y |R_n^{(2)}(y)| > C((a_n \log n)^{1/2} + n^{-1/6} a_n^{-1/2} \log n)\} \leqslant Cn^{-2}. \quad (3.29)$$

It is easy to check that

$$\left\{\frac{1}{g(x)} W(l(y)) - \frac{m(y \mid x)}{g(x)} W(l(\infty)), y \in \mathbf{R}\right\}$$
$$\stackrel{D}{=} \left\{\left(\int K^2(u)\, du/g(x)\right)^{1/2} B(m(y \mid x)), y \in \mathbf{R}\right\}, \quad (3.30)$$

where $\{B(t), 0 \leqslant t \leqslant 1\}$ is a Brownian bridge. Theorem 3.1 now follows from (3.28)–(2.30) and Lemma 4.4.4 of Csörgő and Révész [5].

4. Nearest-Neighbor Type Estimator

We start again with the conditions. We assume that there is a neighborhood of x, denoted by $V=V(x)$, such that the following conditions hold true:

$$\inf_{u\in V} g(u)>0, \tag{4.1}$$

$$\sup_{u\in V}\sup_{y} h(y,u)<\infty, \tag{4.2}$$

$$\sup_{u\in V}\sup_{y}\left|\frac{\partial}{\partial u}h(y,u)\right|<\infty \tag{4.3}$$

and

$$\sup_{u\in V}|g'(u)|<\infty. \tag{4.4}$$

We need more restrictive conditions on K than those assumed in Section 3, namely,

$$K \text{ is bounded and vanishes outside of a finite interval,} \tag{4.5}$$

$$\sup_{u}|K'(u)|<\infty, \tag{4.6}$$

and

$$\sup_{u}|K''(u)|<\infty. \tag{4.7}$$

Without loss of generality, we can assume that $K(u)=0$ unless $u\in[-1,1]$. It will follow from the proofs that instead of (4.6) and (4.7) it is enough to assume that K' and K'' exist and are uniformly bounded almost everywhere with respect to Lebesgue measure.

THEOREM 4.1. *Assume that* (4.1)–(4.7) *and* (3.8)–(3.9) *hold. We can define a sequence of Brownian bridges* $\{B_n^{(2)}(t), 0\leqslant t\leqslant 1\}$ *such that*

$$P\left\{\sup_{y}\left|\gamma_n(y\mid x)-\left(\int K^2(u)\,du\right)^{1/2} B_n^{(2)}(m(y\mid x))\right|\right.$$
$$\left.>C_9((a_n\log n)^{1/2}+n^{-1/6}a_n^{-1/2}\log n)\right\}\leqslant C_{10}n^{-2}.$$

Theorem 4.1 and the Borel–Cantelli lemma immediately imply the following.

COROLLARY 4.1. *If* (4.1)–(4.7) *and* (3.8)–(3.9) *hold, then*

$$\limsup_{n\to\infty} (\log n)^{-1/2} \sup_y |\gamma_n(y \mid x)| \leqslant \left(\int K^2(u)\, du/2 \right)^{1/2} \qquad a.s.$$

The next lemma gives a rate for $k_{(n)} - m$.

LEMMA 4.1. *Assume that* (3.4) *and* (4.1)–(4.4) *hold and*

$$\sup_{u \in V(x)} \sup_y \left| \frac{\partial^2}{\partial u^2} h(y, u) \right| < \infty.$$

Then

$$\sup_y |k_{(n)}(y \mid x) - m(y \mid x)| = O(a_n^3).$$

The proof of the Theorem 4.1 uses Theorem A and calculations similar to those in the proof of Theorem 3.1. The proof is omitted, except for the following

LEMMA 4.2. *If* (4.1)–(4.7) *and* (3.8), (3.9) *hold, then*

$$\gamma_n(y \mid x) = a_n^{-1/2} \int K\left(\frac{G(x) - G(u)}{a_n} \right) d_u \alpha_n(y, u)$$
$$- a_n^{-1/2} m(y \mid x) \int K\left(\frac{G(x) - G(u)}{a_n} \right) dt_n(u) + R_n^{(3)}(y, x)$$

and

$$P\{\sup_y |R_n^{(3)}(y, x)| > C_{11} n^{-1/2} a_n^{-3/2} \log n\} \leqslant C_{12} n^{-2}. \tag{4.8}$$

Proof. A two-term Taylor expansion yields

$$r_n(y \mid x) = a_n^{-1} \int K\left(\frac{G(x) - G(u)}{a_n} \right) d_u F_n(y, u)$$
$$+ n^{-1/2} a_n^{-2} \int (t_n(x) - t_n(u))\, K'\left(\frac{G(x) - G(u)}{a_n} \right) d_u F_n(y, u)$$
$$+ \tfrac{1}{2} n^{-1} a_n^{-3} \int (t_n(x) - t_n(u))^2\, K''(\delta)\, d_u F_n(y, u)$$
$$= R_n^{(4)} + R_n^{(5)} + R_n^{(6)},$$

where δ is between $(G_n(x) - G_n(u))/a_n$ and $(G(x) - G(u))/a_n$. We approximate the first two terms of the expansion of $\gamma_n(y \mid x)$ by

$(na_n)^{1/2}\,(R_n^{(4)} - k_{(n)}(y \mid x))$ and $(na_n)^{1/2}\,R_n^{(5)}$, respectively, and show that $R_n^{(6)}$ is negligible.

First we consider $R_n^{(6)}$. Using the Dvoretzky, Kiefer, and Wolfowitz [10] inequality we get that

$$P\{\sup_u |G_n(u) - G(u)| > C(\log n/n)^{1/2}\} \leqslant Cn^{-2}. \tag{4.10}$$

By condition (3.8) this means that δ is essentially in the interval $(G(x) - G(u))/a_n \pm Ca_n^{1/2}$. Using (4.5), (4.7), (4.10), and Theorem B with (3.8) we obtain that

$$\begin{aligned} &P\{\sup_y |R_n^{(6)}(y)| > Cn^{-1}a_n^{-2}\log n\} \\ &\quad \leqslant P\{\sup_{|G(x)-G(u)| \leqslant Ca_n} (t_n(u) - t_n(x))^2 > Ca_n \log n\} \\ &\quad \leqslant Cn^{-2}. \end{aligned} \tag{4.11}$$

To estimate $R_n^{(5)}$ we observe that

$$\begin{aligned} n^{1/2}a_n^2 R_n^{(5)} = &\int (t_n(x) - t_n(u))\, m(y \mid u)\, K'\left(\frac{G(x) - G(u)}{a_n}\right) dG_n(u) \\ &+ R_n^{(7)} - R_n^{(8)}, \end{aligned} \tag{4.12}$$

where

$$R_n^{(7)} = t_n(x) \int K'\left(\frac{G(x) - G(u)}{a_n}\right) \{d_u F_n(y, u) - m(y \mid u)\, dG_n(u)\},$$

$$R_n^{(8)} = \int t_n(u)\, K'\left(\frac{G(x) - G(u)}{a_n}\right) \{d_u F_n(y, u) - m(y \mid u)\, dG_n(u)\}.$$

We show that $R_n^{(7)}$ and $R_n^{(8)}$ are asymptotically small. For convenience, let

$$\mu_y^1(s, t) = I\{t \leqslant y\}\, K'\left(\frac{G(x) - G(s)}{a_n}\right),$$

$$\mu_y^2(s, t) = m(y \mid s)\, K'\left(\frac{G(x) - G(s)}{a_n}\right).$$

By (4.10) it is enough for $R_n^{(7)}$ to consider

$$\begin{aligned} R_n^{(9)} &= \sum_{i=1}^{n} (I\{Y_i \leqslant y\} - m(y \mid X_i))\, K'\left(\frac{G(x) - G(X_i)}{a_n}\right) \\ &= \sum_{i=1}^{n} \{\mu_y^1(X_i, Y_i) - \mu_y^2(X_i, Y_i)\}. \end{aligned} \tag{4.13}$$

It is easy to see that there exist $N = Cn$ points $-\infty = y_1 < y_2 < \cdots < y_N = \infty$ such that

$$\begin{aligned} |E\mu^1_{y_j}(X_1, Y_1) - E\mu^1_{\mu_{j+1}}(X_1, Y_1)| &\leqslant 1/n, \\ EI\{y_j < Y_1 \leqslant y_{j+1}\} &\leqslant 1/n, \end{aligned} \tag{4.14}$$

$1 \leqslant j \leqslant N-1$. Using the monotonicity of μ^1_y and μ^2_y in y, we get that

$$\begin{aligned} \sup_y |R_n^{(9)}(y)| &\leqslant 3 \max_{1 \leqslant j \leqslant N} \left| \sum_{i=1}^n \{\mu^1_{y_j}(X_i, Y_i) - \mu^2_{y_j}(X_i, Y_i)\} \right| \\ &\quad + 3 \max_{1 \leqslant j \leqslant N-1} \left| \sum_{i=1}^n \{\mu^1_{y_j}(X_i, Y_i) - \mu^1_{y_{j+1}}(X_i, Y_i)\} \right| \\ &= R_n^{(10)} + R_n^{(11)}. \end{aligned}$$

Observing that $E\mu^1_y(X_1, Y_1) = E\mu^2_y(X_1, Y_1)$, we have by Hoeffding's inequality (cf. Serfling [23, p. 75]) that

$$P\left\{ \max_{1 \leqslant j \leqslant N} \left| \sum_{i=1}^n (\mu^1_{y_j}(X_i, Y_i) - \mu^2_{y_j}(X_i, Y_i)) \right| > C(n \log n)^{1/2} \right\} \leqslant Cn^{-2}. \tag{4.15}$$

By (4.14) we have that

$$R_n^{(11)} \leqslant 3 + 6 \max_{1 \leqslant j \leqslant N} \left| \sum_{i=1}^n (\mu^1_{y_j}(X_i, Y_i) - E\mu^1_{y_j}(X_i, Y_i)) \right|. \tag{4.16}$$

Arguing similarly to (4.15) we obtain

$$P\{R_n^{(11)} > C(n \log n)^{1/2}\} \leqslant Cn^{-2}. \tag{4.17}$$

Combining (4.15), (4.17), and (4.10) we have

$$P\{R_n^{(7)} > Cn^{-1/2} \log n\} \leqslant Cn^{-2}. \tag{4.18}$$

Estimation of $R_n^{(8)}$ is similar to $R_n^{(7)}$, but is somewhat lengthier. We can write

$$\begin{aligned} R_n^{(8)} = n^{-3/2} \sum_{1 \leqslant i, j \leqslant n} & (I\{X_j \leqslant X_i\} - G(X_i))(I\{Y_i \leqslant y\} \\ & - m(y \mid X_i)) K'\left(\frac{G(x) - G(X_i)}{a_n}\right). \end{aligned}$$

Let

$$\mu_y^3(z_1, z_2; z_3, z_4) = I\{z_2 \leq y\}(I\{z_3 \leq z_1\} - G(z_1))\, K'\left(\frac{G(x) - G(z_1)}{a_n}\right)$$
$$+ I\{z_4 \leq y\}(I\{z_1 \leq z_3\} - G(z_3))\, K'\left(\frac{G(x) - G(z_3)}{a_n}\right),$$
$$\mu_y^4(z_1, z_2; z_3, z_4) = m(y \mid z_1)(I\{z_3 \leq z_1\} - G(z_1))\, K'\left(\frac{G(x) - G(z_1)}{a_n}\right)$$
$$+ m(y \mid z_3)(I\{z_1 \leq z_3\} - G(z_3))\, K'\left(\frac{G(x) - G(z_3)}{a_n}\right)$$

and

$$R_n^{(12)} = \sum_{1 \leq i, j \leq n} \{\mu_y^3(X_i, Y_i; X_j, Y_j) - \mu_y^4(X_i, Y_i; X_j, Y_j)\}.$$

It is easy to check that

$$R_n^{(8)} = \tfrac{1}{2} n^{-3/2} R_n^{(12)}. \tag{4.19}$$

Using (4.14) we can show by elementary arguments that

$$\sup_y |R_n^{(12)}(y)| \leq 3 \max_{1 \leq N} \left| \sum_{1 \leq i, j \leq n} (\mu_{y_k}^3(X_i, Y_i; X_j, Y_j) - \mu_{y_k}^4(X_i, Y_i; X_j, Y_j)) \right|$$
$$+ 3 \max_{1 \leq k \leq N-1} \left| \sum_{1 \leq i, j \leq n} (\mu_{y_k}^3(X_i, Y_i; X_j, Y_j) \right.$$
$$\left. - \mu_{y_{k+1}}^3(X_i, Y_i; X_j, Y_j)) \right|$$
$$= 3 \max_{1 \leq k \leq N} |R_{n,k}^{(13)}| + 3 \max_{1 \leq k \leq N} |R_{n,k}^{(14)}|. \tag{4.20}$$

For each k, $R_{n,k}^{(13)}$ is a degenerate U-statistic, and an application of Theorem C implies that

$$E(R_{n,k}^{(13)})^{2v} \leq C^{2v} n^{2v} \{v^{2v} + n^{-1} v^{2v+1}\}, \tag{4.21}$$

and C does not depend on k. Therefore, with $v = \log n$,

$$P\{|R_{n,k}^{(13)}| > bn \log n\} = P\{(R_{n,k}^{(13)})^{2v} > (bn \log n)^{2v}\}$$
$$\leq \frac{n^{2v}}{(bn \log n)^{2v}} C^{2v} [(\log n)^{2v} + n^{-1} (\log n)^{2v+1}]$$
$$= (C/b)^{\log n} [1 + n^{-1} \log n] \leq n^{-3}, \tag{4.22}$$

if b is large enough, which implies that

$$P\{\max_{1 \leqslant k \leqslant N} |R_{n,k}^{(13)}| > Cn \log n\} \leqslant Cn^{-2}. \tag{4.23}$$

Using the definition of μ_y^3 we get that

$$|R_{n,k}^{(14)}| \leqslant Cn \sum_{i=1}^{n} I\{y_k < Y_i \leqslant y_{k+1}\}.$$

Now (4.14) and Bernstein's inequality (cf. Serfling [23, p. 95]) imply that

$$P\{\max_{1 \leqslant k \leqslant N-1} |R_{n,k}^{(14)}| \geqslant Cn \log n\} \leqslant Cn^{-2}. \tag{4.24}$$

Collecting together (4.19), (4.20), (4.23), and (4.24) we get that

$$P\{\sup_y |R_n^{(8)}(y)| > Cn^{-1/2} \log n\} \leqslant Cn^{-2}. \tag{4.25}$$

Therefore, for $R_n^{(5)}$ it is enough to consider

$$\int (t_n(x) - t_n(u))\, m(y \mid u)\, K'\left(\frac{G(x) - G(u)}{a_n}\right) dG_n(u).$$

Using (4.5) we get that

$$\begin{aligned} R_n^{(15)} &= \sup_y \left| \int (t_n(x) - t_n(u))(m(y \mid u) - m(y \mid x))\, K'\left(\frac{G(x) - G(u)}{a_n}\right) dG_n(u) \right| \\ &\leqslant C \sup_y \sup_{|G(x) - G(u)| \leqslant a_n} |m(y \mid u) - m(y \mid x)| \\ &\quad \times \sup_{|G(x) - G(u)| \leqslant a_n} \{ |t_n(x) - t_n(u)|\, |G_n(u) - G_n(x)| \}. \end{aligned} \tag{4.26}$$

We obtain from (4.1), (4.3), and (4.4) that

$$\sup_y \sup_{|G(x) - G(u)| \leqslant a_n} |m(y \mid u) - m(y \mid)| = O(a_n). \tag{4.27}$$

We get from (4.4) that

$$\begin{aligned} &\sup_{|G(x) - G(u)| \leqslant a_n} |G_n(u) - G_n(x)| \\ &\quad \leqslant \sup_{|G(x) - G(u)| \leqslant a_n} n^{-1/2} |t_n(x) - t_n(u)| + O(n^{-1/2} a_n). \end{aligned} \tag{4.28}$$

Using now (4.27), (4.28), and Theorem B, we obtain that

$$P\{R_n^{(15)} > C(\log n/n)^{1/2}\} \leqslant Cn^{-2}. \tag{4.29}$$

In a similar fashion to the estimation of $R_n^{(8)}$, an application of Theorem C yields

$$P\left\{\sup_y \left|\int (t_n(x)-t_n(u))\, K'\left(\frac{G(x)-G(u)}{a_n}\right) d(G_n(u)-G(u))\right| > Cn^{-1/2}\log n\right\} \leqslant Cn^{-2}. \tag{4.30}$$

Therefore, since $R_n^{(7)}$ and $R_n^{(8)}$ are negligible, it is enough to consider the first term on the righthand side of (4.12). Integration by parts gives

$$\int (t_n(x)-t_n(u))\, K'\left(\frac{G(x)-G(u)}{a_n}\right) dG(u) = -\,a_n \int K\left(\frac{G(x)-G(u)}{a_n}\right) dt_n(u). \tag{4.31}$$

Combining (4.12), (4.18), (4.25), (4.26), and (4.29)–(4.31) we obtain

$$P\left\{\sup_y \left|R_n^{(5)}(y)+\frac{m(y\mid x)}{n^{1/2}a_n}\int K\left(\frac{G(x)-G(u)}{a_n}\right) dt_n(u)\right| > Cn^{-1}a_n^{-2}\log n\right\} \leqslant Cn^{-2}. \tag{4.32}$$

All that remains is the approximation of $R_n^{(4)}-k_{(n)}(y\mid x)$. Condition (4.5) implies that $l_{(n)}(x)=k_{(n)}(\infty\mid x)=1$ if n is large enough. Using (4.9), (4.11), and (4.32) we obtain that

$$P\{(na_n)^{1/2}\,|l_n(x)-1|>Cn^{-1/2}a_n^{-3/2}\log n\} \leqslant Cn^{-2}. \tag{4.33}$$

By Kiefer's [13] inequality we have

$$P\left\{\sup_y \left|\int K\left(\frac{G(x)-G(u)}{a_n}\right) d_u\alpha_n(y,u)\right| > C(\log n)^{1/2}\right\} \leqslant Cn^{-2}. \tag{4.34}$$

Observing that

$$(na_n)^{-1/2}\,\gamma_n(y\mid x)=r_n(y\mid x)-k_{(n)}(y\mid x)-(l_n(x)-1)(r_n(y\mid x) - k_{(n)}(y\mid x))/l_n(x)+k_{(n)}(y\mid x)(1-l_n(x))/l_n(x),$$

Lemma 4.2 follows from (4.9), (4.11), and (4.32)–(4.34).

5. Limit Points of Conditional Empirical Processes

Theorems 3.1 and 4.1 say that the weak behavior of the conditional empirical processes are essentially the same as for the usual empirical process. Now we consider the exact almost sure properties of the conditional processes. Let

$$r(s) = r(s \mid x) = \inf\{y: m(y \mid x) \geqslant s\}, \tag{5.1}$$

the inverse (quantile) of $m(y \mid x)$. We will assume that there is a neighborhood of x, say $V = V(x)$, such that

$$\sup_{u \in V} |m(r(s \mid x) \mid u) - m(r(s' \mid x) \mid u)| \leqslant C\,|s - s'|^{\tau} \tag{5.2}$$

with some $C > 0$ and $0 < \tau \leqslant 1$. This condition is implied by the condition that for $y < y'$,

$$\sup_{u \in V} |m(y \mid u) - m(y' \mid u)| \leqslant C\,|m(y \mid x) - m(y' \mid x)|^{\tau}.$$

Let $\boldsymbol{H}$ be the set of absolutely continuous functions (with respect to Lebesgue measure) on $[0, 1]$ for which $f(0) = f(1) = 0$ and $\int_0^1 (f'(t))^2\, dt \leqslant 1$.

Theorem 5.1. *Assume that* (3.1)–(3.4), (3.8), (3.9), (4.5), *and* (5.2) *hold and that* $a_n = n^{-\mu}$, $0 < \mu < 2\lambda$, *where* λ *is given in Theorem* D.

(i) *If* (3.1)–(3.4) *hold, then*

$$\left\{\left(2 \log\log n \int K^2(u)\, du/g(x)\right)^{-1/2} \beta_n(r(s \mid x) \mid x), 0 \leqslant s \leqslant 1\right\}$$

is almost surely relatively compact in $D[0, 1]$ *with* $\boldsymbol{H}$ *as its set of limit points.*

(ii) *If* (4.1)–(4.4) *hold instead, then*

$$\left\{\left(2 \log\log n \int K^2(u)\, du\right)^{-1/2} \gamma_n(r(s \mid x) \mid x),\ 0 \leqslant s \leqslant 1\right\}$$

is almost surely relatively compact in $D[0, 1]$ *with* $\boldsymbol{H}$ *as its set of limit points.*

The next result follows immediately from Theorem 5.1.

COROLLARY 5.1. *Assume the conditions of Theorem* 5.1. *For the kernel estimator, if* (3.1)–(3.4) *hold, then*

$$\limsup_{n \to \infty} (\log \log n)^{-1/2} \sup_y |\beta_n(y \mid x)| = \left(\frac{\int K^2(u)\, du}{2g(x)} \right)^{1/2} \qquad a.s.$$

For the NN estimator, if (4.1)–(4.4) *hold, then*

$$\limsup_{n \to \infty} (\log \log n)^{-1/2} \sup_y |\gamma_n(y \mid x)| = \left(\int K^2(u)\, du/2 \right)^{1/2} \qquad a.s.$$

Proof of Theorem 5.1. We show the proof only for β_n, as the result for the NN estimator follows along similar lines. By Lemma 3.2, it is enough to consider

$$a_n^{-1/2} \left\{ \int K\left(\frac{x-u}{a_n} \right) d_u \alpha_n(y, u) - m(y \mid x) \int K\left(\frac{x-u}{a_n} \right) dt_n(u) \right\}.$$

Applying Theorem D, we get

$$\sup_y \left| a_n^{-1/2} \int K\left(\frac{x-u}{a_n} \right) d_u(\alpha_n(y, u) - n^{-1/2} \Gamma^{(2)}(y, u, n)) \right| \overset{\text{a.s.}}{=} o(n^{-\lambda} a_n^{-1/2}).$$

Let $\{ W_J(s, t, u),\ 0 \leqslant s,\ t \leqslant 1,\ u > 0 \}$ be a three-dimensional Wiener process with $EW_J(s, t, u) = 0$ and $EW_J(s, t, u)\, W_J(s', t', u') = (u \wedge u')\, J(s \wedge s', t \wedge t')$. Then from (3.21) we have

$$\{ \Gamma^{(2)}(y, u, n),\ y, u \in \mathbf{R},\ n \geqslant 1 \}$$

$$\overset{D}{=} \{ W_J(H(y), G(u), n) - F(y, u)\, W_J(1, 1, n),\ y, u \in \mathbf{R},\ n \geqslant 1 \}. \qquad (5.3)$$

Integration by parts and the law of the iterated logarithm for Wiener processes gives

$$(na_n)^{-1/2}\, |W_J(1, 1, n)| \sup_y \left| \int K\left(\frac{x-u}{a_n} \right) h(y, u)\, du \right| \overset{\text{a.s.}}{=} O((a_n \log \log n)^{1/2}).$$

Hence it suffices to investigate

$$\Gamma_n^{(3)}(s) = (na_n)^{-1/2} \left\{ \int K\left(\frac{x-u}{a_n} \right) d_u W_J(H(r(s), G(u), n) - s \int K\left(\frac{x-u}{a_n} \right) dW_J(1, G(u), n) \right\}.$$

We obtain from (3.26) that

$$\sup_{0 \leqslant s, t \leqslant 1} \left| E\Gamma_n^{(3)}(s)\, \Gamma_n^{(3)}(t) - \int K^2(u)\, du\, g(x)(t \wedge s - ts) \right| = o(1). \quad (5.4)$$

Elementary calculations show that

$$E(\Gamma_n^{(3)}(t) - \Gamma_n^{(3)}(s))^2 \leqslant C\, |t-s|^{\tau},$$

where τ is given by (5.2).

Let $c > 1$ and define $n_\nu = [c^\nu]$, with ν an integer. We next show that for every $\varepsilon > 0$ there is a $c = c(\varepsilon) > 1$ such that

$$\limsup_{\nu \to \infty} \max_{n_\nu \leqslant n \leqslant n_{\nu+1}} (\log \nu)^{-1/2} \sup_{0 \leqslant s \leqslant 1} |\Gamma_n^{(3)}(s) - \Gamma_{n_\nu}^{(3)}(s)| \leqslant \varepsilon \qquad \text{a.s.} \quad (5.5)$$

Let $n \leqslant m$ such that $\frac{1}{2} \leqslant n/m$ and define

$$\Gamma^{(4)}(t, s) = \Gamma_{n+t(m-n)}^{(3)}(s) - \Gamma_n^{(3)}(s), \qquad 0 \leqslant s, t \leqslant 1. \quad (5.6)$$

It is not difficult to calculate the covariance function of $\Gamma^{(4)}$ because the process is given in terms of integrals with respect to a Wiener measure. Long but elementary calculations give the following results:

$$E(\Gamma^{(4)}(t, s) - \Gamma^{(4)}(t', s))^2 \leqslant C\, |1 - m/n|\, |t - t'|, \quad (5.7)$$

$$E(\Gamma^{(4)}(t, s) - \Gamma^{(4)}(t, s'))^2 \leqslant C\, |1 - m/n|\, |s - s'|^{\tau} \quad (5.8)$$

and

$$E(\Gamma^{(4)}(t, s))^2 \leqslant C\, |1 - m/n|, \quad (5.9)$$

where C is an absolute constant. By (5.7)–(5.9) we can apply Lemma 2 of Lai [16] and get that for all $x \geqslant 10$

$$P\left\{ \sup_{0 \leqslant s, t \leqslant 1} |\Gamma^{(4)}(t, s)| > xC\, |1 - m/n| \right\} \leqslant 100 \int_x^{\infty} \exp(-u^2/2)\, du. \quad (5.10)$$

Hence (5.5) follows immediately from (5.6) and (5.10) with $x = (4 \log \nu)^{1/2}$. Now we observe that ε can be arbitrarily small in (5.5), and $\log\log n_\nu / \log \nu \to 1$ as $\nu \to \infty$ for all $c > 1$. Thus Theorem 1.1a of Mangano [19] implies that $\{(2 \log\log n \int K^2(u)\, du / g(x))^{-1/2}\, \Gamma_n^{(3)}(s),\ 0 \leqslant s \leqslant 1\}$, is almost surely relatively compact in C[0, 1] with limit points in $\boldsymbol{H}$.

Next we show that H must be in the cluster set. Let $\eta > \nu$. Then straightforward calculations yield

$$\begin{aligned}
&E\{E(\Gamma^{(3)}_{n_\eta}(s) \mid \Gamma^{(3)}_u(t), 0 \leqslant t \leqslant 1, 0 \leqslant u \leqslant n_\nu)\}^2 \\
&\quad \leqslant E\{E(\Gamma^{(3)}_{n_\eta}(s) \mid W_J(t, v, u), 0 \leqslant t, v \leqslant 1, 0 \leqslant u \leqslant n_\nu)\}^2 \\
&\quad \leqslant Cc^{\nu - \eta},
\end{aligned}$$

which goes to zero as $\eta - \nu \to \infty$. Now we can apply Theorem 1.1b of Mangano [19] (cf. his remark on p. 912) to show that the cluster set is H. The proof of Theorem 5.1 is now complete.

References

[1] Borisov, I. S. (1982). Approximation of empirical fields, constructed with respect to vector observations with dependent components. *Sibirsk. Mat. Zh.* **23** 31–41.

[2] Cheng, P. E. (1984). Strong consistency of nearest neighbor regression function estimates. *J. Multivariate Anal.* **15** 63–72.

[3] Cheng, K. F., and Cheng, P. E. (1987). Robust nonparametric estimation of a regression function. *Sankhyā Ser. A* **49**.

[4] Collomb, G. (1981). Estimation non-parametrique de la regression: revue bibliographique. *Internat. Statist. Rev.* **49** 75–93.

[5] Csörgő, M., and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Akad. Kiadó, Budapest and Academic Press, New York.

[6] Csörgő, M. and Horváth, L. (1988). A note on strong approximation of multivariate empirical processes. *Stochastic Process Appl.* (to appear).

[7] Csörgő, M. and Révész, P. (1975). A strong approximation of the multivariate empirical process. *Studia Sci. Math. Hungar.* **10** 427–434.

[8] de Acosta, A. (1982). Invariance principles in probability for triangular arrays of B-valued random vectors and some applications. *Ann. Probab.* **10** 346–373.

[9] Dehling, H., Denker, M., and Philipp, W. (1987). The almost sure invariance principle for the empirical process of U-statistic structure. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 121–134.

[10] Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1956). Asymptotic minimal character of the sample distribution function and of the multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

[11] Härdle, W. (1984). Robust regression function estimation. *J. Multivariate Anal.* **14** 169–180.

[12] Johnston, G. J. (1982). Probabilities of maximal deviations for nonparametric regression function estimates. *J. Multivariate Anal.* **12** 402–414.

[13] Kiefer, J. (1961). On large deviations of the empiric distribution function of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649–660.

[14] Komlós, J., Major, P., and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF, I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131.

[15] Konakov, V. D., and Piterbarg, V. I. (1984). On the convergence rate of maximal deviation distribution for kernel regression estimates. *J. Multivariate Anal.* **15** 279–294.

[16] Lai, T. L. (1974). Reproducing kernel Hilbert spaces and the law of the iterated logarithm for Gaussian processes. *Z. Wahrsch. Verw. Gebiete* **29** 7–19.

[17] LIERO, H. (1982). On the maximal deviation of the kernel regression function estimate. *Math. Operationsforsch. Statist. Ser. Statist.* **13** 171–182.

[18] MACK, Y. P., AND SILVERMAN, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. Verw. Gebiete* **61** 405–415.

[19] MANGANO, G.-C. (1976). Sequential compactness of certain sequences of Guassian random variables with values in C[0, 1]. *Ann. Probab.* **4** 901–913.

[20] NADARAYA, E. A. (1964). On estimating regression. *Theory Probab. Appl.* **9** 141–142.

[21] PHILIPP, W., AND PINZUR, L. (1980). Almost sure approximation theorems for the multivariate empirical process. *Z. Wahrsch. Verw. Gebiete* **54** 1–13.

[22] RÉVÉSZ, P. (1979). On the nonparametric estimation of the regression function. *Problems Control Inform. Theory* **8** 297–302.

[23] SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

[24] STONE, C. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.* **8** 1348–1360.

[25] STONE, C. (1977). Consistent nonparametric regression. *Ann. Statist.* **5** 595–620.

[26] STUTE, W. (1986). Conditional empirical processes. *Ann. Statist.* **14** 638–647.

[27] STUTE, W. (1986). *On Almost Sure Convergence of Conditional Empirical Distribution Functions.* Mathematisches Institut der Justus-Liebig-Universität. Preprint.

[28] STUTE, W. (1984). Asymptotic normality of nearest neighbor regression function estimates. *Ann. Statist.* **12** 917–926.

[29] WATSON, G. S. (1964). Smooth regression analysis. *Sankhyā Ser. A* **26** 359–372.

[30] WICHURA, M. (1973). Some Strassen-type laws of the iterated logarithm for multi-parameter stochastic processes with independent increments. *Ann. Probab.* **1** 272–296.

[31] YANG, S.-S. (1981). Linear functions of concomitants of order statistics with application to non-parametric estimation of a regression function. *J. Amer. Statist. Assoc.* **76** 658–662.