The Annals of Statistics
1990, Vol. 18, No. 3, 1400–1415

KERNEL AND NEAREST-NEIGHBOR ESTIMATION OF A CONDITIONAL QUANTILE

BY P. K. BHATTACHARYA[1] AND ASHIS K. GANGOPADHYAY

University of California at Davis and University of North Carolina at Chapel Hill

Let $(X_1, Z_1), (X_2, Z_2), \ldots, (X_n, Z_n)$ be iid as (X, Z), Z taking values in R^1, and for $0 < p < 1$, let $\xi_p(x)$ denote the conditional p-quantile of Z given $X = x$, i.e., $P(Z \le \xi_p(x) | X = x) = p$. In this paper, kernel and nearest-neighbor estimators of $\xi_p(x)$ are proposed. In order to study the asymptotics of these estimates, Bahadur-type representations of the sample conditional quantiles are obtained. These representations are used to examine the important issue of choosing the smoothing parameter by a local approach (for a fixed x) based on weak convergence of these estimators with varying k in the k-nearest-neighbor method and with varying h in the kernel method with bandwidth h. These weak convergence results lead to asymptotic linear models which motivate certain estimators.

1. Introduction. Let $(X_1, Z_1), (X_2, Z_2), \ldots$ be two-dimensional random vectors which are iid as (X, Z), and for $0 < p < 1$, let $\xi_p(x)$ denote the conditional p-quantile of Z given $X = x$. We consider the problem of estimating $\xi_p(x)$ from the data $(X_1, Z_1), \ldots, (X_n, Z_n)$, and study the asymptotic properties of the kernel and nearest-neighbor (NN) estimators as $n \to \infty$.

Usefulness of conditional quantile functions as good descriptive statistics has been discussed by Hogg (1975) who calls them percentile regression lines. The problem of conditional quantile estimation has been investigated by Bhattacharya (1963) following the fractile approach, and is also included in the general scheme of nonparametric regression considered by Stone (1977). More recently asymptotic normality of estimators of conditional quantiles has been proved by Cheng (1983), who considered kernel estimators in the fixed design case, and by Stute (1986), who considered NN-type estimators in the random design case. However, both of these authors took the bandwidth h_n to be $O(n^{-1/3})$ which kept the bias smaller than the random error by an order of magnitude, and thereby made the rate of convergence slower than optimal.

In this paper, we obtain Bahadur-type representations [Bahadur (1966)] of NN and kernel estimators of conditional quantiles in order to study their asymptotics (Theorems N1 and K1). Bias terms show up in these representations because of our choice of $k = k_n = O(n^{4/5})$ in the k-NN method and $h = h_n = O(n^{-1/5})$ in the kernel method (with uniform kernel), for the purpose of achieving optimum balance between bias and random error. Using

Received February 1988; revised June 1989.
[1]Research supported by NSF grants DMS-86-01641 and DMS-87-03401.
AMS 1980 *subject classifications*. Primary 62G05, 62J02; secondary 62G20, 62G30, 60F17.
Key words and phrases. Conditional quantile, kernel estimator, nearest-neighbor estimator, Bahadur representation, weak convergence, Brownian motion, asymptotic linear model, order statistics, induced order statistics.

these representations, we prove weak convergence results (Theorems N2 and K2) for the NN estimators with $n^{-4/5}k$ varying over $[a, b]$, $0 < a < b$, and for kernel estimators with $n^{1/5}h$ varying similarly. The results are analogous to the one obtained by Bhattacharya and Mack (1987) for NN regression estimators with varying k and lead to asymptotic models similar to theirs. The conditional quantile appears as a parameter in these models along with another parameter which governs the bias. Consideration of best linear unbiased estimators in these models provides a different approach to the choice of smoothing parameters.

2. The main results. For the random vector (X, Z), let f denote the pdf of X and $g(\cdot|x)$ the conditional pdf of Z given $X = x$, with corresponding conditional cdf $G(\cdot|x)$. We want to estimate $\xi_p(x_0)$, the conditional p-quantile of Z given $X = x_0$. Since $p \in (0, 1)$ and x_0 will remain fixed throughout our discussion, we shall write $\xi_p(x_0) = \xi$.

The following regularity conditions are assumed.

1. (a) $f(x_0) > 0$.
   (b) $f''(x)$ exists in a neighborhood of x_0, and there exist $\varepsilon > 0$ and $A < \infty$ such that $|x - x_0| \le \varepsilon$ implies $|f''(x) - f''(x_0) \le A|x - x_0|$.
2. (a) $g(\xi|x_0) > 0$, where $G(\xi|x_0) = p$.
   (b) The partial derivatives $g_z(z|x)$ and $g_{xx}(z|x)$ of $g(z|x)$ and $G_{xx}(z|x)$ of $G(z|x)$ exist in a neighborhood of (x_0, ξ), and there exist $\varepsilon > 0$ and $A < \infty$ such that $|x - x_0| \le \varepsilon$ and $|z - \xi| \le \varepsilon$ together imply

$$|g_z(z|x)| \le A, \qquad |g_x(z|x_0)| \le A, \qquad |g_{xx}(z|x_0)| \le A,$$
$$|g_{xx}(z|x) - g_{xx}(z|x_0)| \le A|x - x_0|, \qquad |G_{xx}(z|x) - G_{xx}(z|x_0)| \le A|x - x_0|.$$

By condition 2, ξ is uniquely defined by $G(\xi|x_0) = p$.

Now let $\{(X_i, Z_i),\ i = 1, 2, \dots\}$ be iid as (X, Z), and let $Y_i = |X_i - X_0|$, so that $\{(Y_i, Z_i),\ i = 1, 2, \dots\}$ are iid as $(Y = |X - x_0|, Z)$ with the pdf f_Y of Y, the conditional pdf $g^*(\cdot|y)$ of Z given $Y = y$ and the corresponding conditional cdf $G^*(\cdot|y)$ given by

$$f_Y(y) = f(x_0 + y) + f(x_0 - y),$$
$$(1)\quad g^*(z|y) = [f(x_0 + y)g(z|x_0 + y) + f(x_0 - y)g(z|x_0 - y)]/f_Y(y),$$
$$G^*(z|y) = [f(x_0 + y)G(z|x_0 + y) + f(x_0 - y)G(z|x_0 - y)]/f_Y(y).$$

Note that

$$g^*(z|0) = g(z|x_0) = g(z), \qquad G^*(z|0) = G(z|x_0) = G(z).$$

Here and in what follows, we write $g(z|x_0) = g(z)$ and $G(z|x_0) = G(z)$ for simplicity.

Let $Y_{n1} < \cdots < Y_{nn}$ denote the order statistics and $Z_{n1}, \dots, Z_{nn}$ the induced order statistics of $(Y_1, Z_1), \dots, (Y_n, Z_n)$, i.e., $Z_{ni} = Z_j$ if $Y_{ni} = Y_j$. For any positive integer $k \le n$, the k-NN empirical cdf of Z (with respect to x_0) is now defined as

$$\hat{G}_{nk}(z) = k^{-1} \sum_{i=1}^{k} 1(Z_{ni} \le z), \tag{2}$$

where $1(S)$ denotes the indicator of the event S. The k-NN estimator of ξ can now be expressed as the p-quantile of $\hat{G}_{nk}$, i.e.,

$$\begin{aligned} \hat{\xi}_{n,k} &= \text{the } [kp]\text{th order statistic of } Z_{n1}, \dots, Z_{nk} \\ &= \inf\{z: \hat{G}_{nk}(z) \ge [kp]/k\}. \end{aligned} \tag{3}$$

The kernel estimator of ξ with uniform kernel and bandwidth h can also be expressed in the same manner, viz.,

$$\begin{aligned} \tilde{\xi}_{nh} &= \inf\{z: \hat{G}_{nK_n(h)}(z) \ge [K_n(h)p]/K_n(h)\}, \\ K_n(h) &= \sum_{i=1}^{n} 1(Y_i \le h/2) = n\hat{F}_{Y,n}(h/2), \end{aligned} \tag{4}$$

$\hat{F}_{Y,n}$ being the empirical cdf of $Y_1, \dots, Y_n$. The kernel estimators are thus related to the NN estimators by

$$\tilde{\xi}_{nh} = \hat{\xi}_{nK_n(h)}, \tag{5}$$

where $K_n(h)$ is the random integer given by (4).

We now state our main results in the following two theorems of which Theorem N1 gives a Bahadur-type representation for the k-NN estimator $\hat{\xi}_{nk}$ of ξ with k lying in

$$I_n(a,b) = \{k: k_0 = [n^{4/5}a] \le k \le [n^{4/5}b] = k_1\}, \qquad 0 < a < b,$$

and Theorem K1 gives a corresponding representation for the kernel estimator $\tilde{\xi}_{nh}$ with h lying in

$$J_n(c,d) = [n^{-1/5}c, n^{-1/5}d], \qquad 0 < c < d.$$

THEOREM N1.

$$\hat{\xi}_{nk} - \xi = \beta(\xi)(k/n)^2 + \{kg(\xi)\}^{-1} \sum_{i=1}^{k} [1(Z_{ni}^* > \xi) - (1-p)] + R_{nk},$$

where

$$\beta(\xi) = -[f(x_0)G_{xx}(\xi|x_0) + 2f'(x_0)G_x(\xi|x_0)]/\{24 f^3(x_0)g(\xi)\},$$

$$Z_{ni}^* = G^{-1} \circ G^*(Z_{ni}|Y_{ni}), \qquad 1 \le i \le n, n \ge 1,$$

and

$$\max_{k \in I_n(a,b)} |R_{nk}| = O(n^{-3/5} \log n), \quad a.s.$$

THEOREM K1.

$$\tilde{\xi}_{nh} - \xi = \beta(\xi) f^2(x_0) h^2 + \{[nhf(x_0)] g(\xi)\}^{-1} \times \sum_{i=1}^{[nhf(x_0)]} [1(Z^*_{ni} > \xi) - (1-p)] + R^*_{nh},$$

where $\beta(\xi)$ *and* Z^*_{ni} *are as in Theorem* N1, *and*

$$\sup_{h \in J_n(c,d)} |R^*_{nh}| = O(n^{-3/5} \log n), \quad a.s.$$

REMARKS.

1. Let $\mathscr{A} = \sigma\{Y_1, Y_2, \dots\}$ denote the σ-field of $Y_1, Y_2, \dots$. Then $Z_{n1}, \dots, Z_{nn}$ are conditionally independent given $\mathscr{A}$, with Z_{ni} having conditional cdf $G^*(\cdot|Y_{ni})$, as shown by Bhattacharya (1974). Hence $G^*(Z_{ni}|Y_{ni})$, $1 \le i \le n$, are conditionally independent and uniform (0, 1) given $\mathscr{A}$, and therefore,

$$P[Z^*_{ni} \le z_i, 1 \le i \le n] = EP[G^*(Z_{ni}|Y_{ni}) \le G(z_i), 1 \le i \le n | \mathscr{A}] = \prod_{i=1}^{n} G(z_i).$$

Thus for each n, $Z^*_{n1}, \dots, Z^*_{nn}$ are iid with cdf G. Since $G(\xi) = p$, it follows that for each n, the summands $1(Z^*_{ni} > \xi) - (1-p)$ in the above representations are independent random variables with mean 0 and variance $p(1-p)$.

2. The remainder terms in both theorems are $O(n^{-3/5} \log n)$, a.s. In Theorem N1, this corresponds to $O(k^{-3/4} \log k)$ with $k = O(n^{4/5})$, as one would expect. The same explanation applies to Theorem K1, because $[nhf(x_0)] = O(n^{4/5})$ for $h = O(n^{-1/5})$.

3. Weaker versions of the above theorems were proved by Gangopadhyay (1987). His remainder terms were $o(n^{-2/5})$, a.s. in Theorem N1 and $o_p(n^{-2/5})$ in Theorem K1.

3. Weak convergence properties of NN estimators with varying k and kernel estimators with varying bandwidth. Consider the stochastic processes $\{\hat{\xi}_{nk}, k \in I_n(a,b)\}$ and $\{\tilde{\xi}_{nh}, h \in J_n(c,d)\}$. The two theorems in this section describe the weak convergence properties of suitably normalized versions of these processes, as $n \to \infty$. The symbol $\to_{\mathscr{D}}$ indicates convergence in distribution, i.e., weak convergence of the distributions of the stochastic processes (or random vectors) under consideration and $\{B(t), t \ge 0\}$ denotes a standard Brownian motion.

THEOREM N2. *Let* $T_n(t) = \hat{\xi}_{n,[n^{4/5}t]}$. *Then for any* $0 < a < b$,

$$\{n^{2/5}[T_n(t) - \xi] - \beta t^2, a \le t \le b\} \to_{\mathscr{D}} \{\sigma t^{-1} B(t), a \le t \le b\},$$

where $\beta = \beta(\xi)$ *given in Theorem* N1 *and* $\sigma^2 = p(1-p)/g^2(\xi)$.

PROOF. In the representation for $\hat{\xi}_{nk}$ given in Theorem N1, take $k = [n^{4/5}t] = n^{4/5}t + \varepsilon_n(t)$ with $0 \le \varepsilon_n(t) < 1$. After a little rearrangement of terms, this leads to

$$n^{2/5}[T_n(t) - \xi] - \beta(\xi)t^2 = \frac{\sqrt{p(1-p)}}{g(\xi)} t^{-1} n^{-2/5} \sum_{i=1}^{[n^{4/5}t]} W_{ni} + \sum_{j=1}^{3} R_{nj}(t),$$

where

$$W_{ni} = \frac{1(Z_{ni}^* > \xi) - (1-p)}{\sqrt{p(1-p)}}, \qquad 1 \le i \le n, \tag{6}$$

are iid with mean 0 and variance 1 for each n in view of Remark 1. However, it can be shown easily that $\sum_{j=1}^{3} R_{nj}(t) = o_p(1)$. We thus have, with $\sigma = \sqrt{p(1-p)}/g(\xi)$,

$$n^{2/5}[T_n(t) - \xi] - \beta(\xi)t^2 = \sigma t^{-1} n^{-2/5} \sum_{1}^{[n^{4/5}t]} W_{ni} + o_p(1)$$

uniformly in $a \le t \le b$. Now use Theorem 1, page 452 of Gikhman and Skorokhod (1969) to see that $\{n^{-2/5}\sum_1^{[n^{4/5}t]} W_{ni},\ a \le t \le b\} \to_{\mathscr{D}} \{B(t),\ a \le t \le b\}$. This proves the theorem. □

In the kernel case, we can use a similar approach to establish the following theorem.

THEOREM K2. *Let* $S_n(t) = \tilde{\xi}_{n, n^{-1/5}t}$. *Then for any* $0 < c < d$,

$$\{n^{2/5}[S_n(t) - \xi] - \gamma t^2, c \le t \le d\} \to_{\mathscr{D}} \{\tau t^{-1} B(t), c \le t \le d\},$$

where $\gamma = \beta f^2(x_0)$ *and* $\tau = \sigma/\sqrt{f}(x_0)$, *with* β *and* σ *as in Theorem* N2.

REMARK. The uniformity of the order of magnitude of the remainder terms in Theorems N1 and K1 was crucial in proving Theorems N2 and K2.

From Theorems N2 and K2, it follows that $n^{2/5}[T_n(t) - \xi] \to_{\mathscr{D}} N(\beta t^2, \sigma^2 t^{-1})$ and $n^{2/5}[S_n(t) - \xi] \to_{\mathscr{D}} N(\gamma t^2, \tau^2 t^{-1})$ for each t, where $N(\mu, \sigma^2)$ denotes a Gaussian r.v. with mean μ and variance σ^2. Hence the asymptotic mean-squared errors (AMSE) of $T_n(t)$ is $n^{-4/5}(\beta^2 t^4 + \sigma^2 t^{-1})$, which is minimized at $t_1^* = \{\sigma^2/(4\beta^2)\}^{1/5}$ and the AMSE of $S_n(t)$ is $n^{-4/5}(\gamma^2 t^4 + \tau^2 t^{-1})$, which is minimized at $t_2^* = \{\tau^2/(4\gamma^2)\}^{1/5}$. However, these optimum t_1^* and t_2^* involve unknown quantities involving the marginal distribution of X and the conditional distribution of Z given $X = x_0$. Although one could attempt to use consistent (but possibly nonoptimal) estimates of β, σ^2, γ and τ^2 to approximate t_1^* and t_2^* in the spirit of Woodroofe (1970) and Krieger and Pickands (1981), we shall take another approach in the next section by considering linear combinations of k-NN estimators with varying k and kernel estimators with varying bandwidth.

4. Asymptotic linear models and linear combinations of $\hat{\xi}_{nk}$ and $\tilde{\xi}_{nk}$. Neglect the remainder term in Theorem N1 to obtain the following asymptotic linear model for $\{\hat{\xi}_{nk},\ k_0 \le k \le k_1\}$:

$$\hat{\xi}_{nk} \simeq \xi + (k/n)^2\beta + \sigma\Delta_{nk}, \qquad k_0 \le k \le k_1, \tag{7}$$

where β and σ are as in Theorem N2 and

$$\Delta_{nk} = k^{-1}\sum_{1}^{k} W_{ni},$$

in which W_{ni}, $1 \le i \le n$, are the iid r.v.'s with mean 0 and variance 1 given by (6). Hence

$$E(\Delta_{nk}) = 0, \qquad \operatorname{Cov}(\Delta_{nj}, \Delta_{nk}) = \min(j^{-1}, k^{-1}).$$

Due to the covariance structure of $\{\Delta_{nk}\}$, we have that

$$\begin{aligned} \varepsilon_{nk} &= \{k(k+1)\}^{1/2}(\Delta_{nk+1} - \Delta_{nk}\}, \qquad k_0 \le k \le k_1 - 1,\\ \varepsilon_{nk_1} &= (k_1)^{1/2}\Delta_{nk_1} \end{aligned} \tag{7a}$$

are mutually uncorrelated with mean 0 and variance 1. This is exactly like the asymptotic linear model for k-NN regression obtained by Bhattacharya and Mack (1987). Following their approach, we take normalized differences in (7a) to get

$$\begin{aligned} V_{nk} &= \{k(k+1)\}^{1/2}(\hat{\xi}_{nk+1} - \hat{\xi}_{nk})\\ &= u_{nk}\beta + (g(\xi))^{-1}[p(1-p)]^{1/2}\varepsilon_{nk}, \qquad k_0 \le k \le k_1 - 1,\\ V_{nk_1} &= (k_1)^{1/2}\hat{\xi}_{nk_1}\\ &= (k_1)^{1/2}\xi + u_{uk_1}\beta + (g(\xi))^{-1}[p(1-p)]^{1/2}\varepsilon_{nk_1}, \end{aligned} \tag{7b}$$

where

$$\begin{aligned} u_{nk} &= [k(k+1)]^{1/2}(2k+1)n^{-2}, \qquad k_0 \le k \le k_1 - 1,\\ u_{nk_1} &= k_1^{5/2}n^{-2}. \end{aligned} \tag{7c}$$

So, the BLUE (best linear unbiased estimator) of ξ in the asymptotic linear model given by (7a), (7b) and (7c) is

$$\hat{\xi} = \hat{\xi}_{nk_1} - \hat{\beta}(k_1/n)^2,$$

where

$$\hat{\beta} = \left\{\sum_{k=k_0}^{k_1-1} u_{nk}^2\right\}^{-1}\left[\sum_{k=k_0}^{k_1-1} u_{nk}V_{nk}\right].$$

Using Theorem N2 and arguing in exactly the same way as in Bhattacharya and Mack (1987), the asymptotic distribution of the BLUE of ξ is obtained as

$$n^{2/5}(\hat{\xi} - \xi) \to_{\mathscr{D}} N(0, \sigma^2(A+1)b^{-1}), \qquad A = (5/4)\left[1 - (a/b)^5\right]^{-1}.$$

On the other hand, let $[n^{4/5}t^*]$ denote the optimum number of NN's. Then the ARE of $\hat{\xi}$ with respect to ξ (based on a comparison of asymptotic mean-squared errors) is

$$(5/4)b(A+1)^{-1}(4\beta^2/\sigma^2)^{1/5}.$$

It should be noted that we can choose b sufficiently large for any $(a/b) < 1$ to make ARE's arbitrarily large. However, due to practical limitation imposed by $k_1 = [n^{4/5}b] \le n$, the choice of b is restricted by a finite quantity for any given sample size.

The kernel estimators $\tilde{\xi}_{nh}$ with $n^{-1/5}c \le h \le n^{-1/5}d$ also satisfy an asymptotic linear model similar to (7). For this, first neglect the remainder term in Theorem K1 to obtain

$$\tilde{\xi}_{nh} \simeq \xi + \beta f^2(x_0)h^2 + \sigma\Delta_{n,[nhf(x_0)]}, \qquad h \in J_n(c,d), \tag{8}$$

where β, σ and Δ_{nk} are as in (7). In this model, the indexing parameter h is continuous, but can be discretized by letting $m_0 = [n^{4/5}cf(x_0)]$, $m_1 = [n^{4/5}df(x_0)]$ and $h(m) = m/\{nf(x_0)\}$ for $m = m_0, m_0+1, \dots, m_1$. Then

$$\tilde{\xi}_{nh} \simeq \tilde{\xi}_{nh(m)} \simeq \xi + (m/n)^2\beta + \sigma\Delta_{nm}, \tag{9}$$

for $h(m) \le h < h(m+1)$ and $m_0 \le m \le m_1$. The asymptotics of the linear combinations of $\tilde{\xi}_{nh}$ given by (9) are, therefore, essentially the same as in the model given by (7). The details are as in Gangopadhyay (1987).

5. Proof of Theorem N1: Preliminary lemmas. The k-NN estimator $\hat{\xi}_{nk}$ of ξ is the p-quantile of the empirical cdf $\hat{G}_{nk}$ of $Z_{n1}, \dots, Z_{nk}$, which are conditionally independent with Z_{ni} having conditional cdf $G^*(\cdot|Y_{ni})$. It is, therefore, natural to think of $\hat{\xi}_{nk}$ as an estimator of the p-quantile ξ_{nk} of the random cdf $k^{-1}\Sigma_1^k G^*(\cdot|Y_{ni})$. The discrepancy between this random cdf and the cdf $G^*(\cdot|0) = G(\cdot)$ of which ξ is the p-quantile, is going to give rise to a bias in addition to the random error in estimating ξ_{nk} by $\hat{\xi}_{nk}$. To facilitate the examination of this bias and the random error, we introduce some notation. Let

$$\begin{aligned} g^*(\cdot|Y_{ni}) &= g_{ni}(\cdot), & G^*(\cdot|Y_{ni}) &= G_{ni}(\cdot), \\ \bar{g}_{nk}(\cdot) &= k^{-1}\sum_1^k g_{ni}(\cdot), & \bar{G}_{nk}(\cdot) &= k^{-1}\sum_1^k G_{ni}(\cdot). \end{aligned} \tag{10}$$

Then ξ_{nk}, which is the target of $\hat{\xi}_{nk}$ is given by

$$\bar{G}_{nk}(\xi_{nk}) = p = G(\xi). \tag{11}$$

To examine the asymptotic properties of $\hat{\xi}_{nk} - \xi_{nk}$ and $\xi_{nk} - \xi$ for $k \in I_n(a,b)$, we now analyze the corresponding properties of $\hat{G}_{nk}(\cdot) - \bar{G}_{nk}(\cdot)$ and $\bar{G}_{nk}(\cdot) - G(\cdot)$.

Note that the Lemma 1 (stated below) implies that the order statistics $0 < Y_{n1} < \cdots < Y_{n,[n^{4/5}b]}$ are of the order of $n^{-1/5}$. Consequently, for $k \in I_n(a,b)$, it should be possible to approximate the pdf's $\bar{g}_{nk}$ and the cdf's $\bar{G}_{nk}$ defined in (10), by the first few terms of their expansions in powers of Y_{ni}, $i \le [n^{4/5}b]$. To this end, we have the following lemmas.

LEMMA 1. *For $B > b/f(x_0)$ and for sufficiently large n,*

$$P\left[Y_{n,[n^{4/5}b]} > n^{-1/5}B\right] \le \exp\left[-2n^{3/5}\{Bf(x_0) - b\}^2\right].$$

PROOF. This is proved in Bhattacharya and Mack (1987). □

LEMMA 2. $k^{-1}\sum_1^k Y_{ni}^2 = \{12 f^2(x_0)\}^{-1}(k/n)^2 + R_{nk}$, *where*

$$\max_{k \le [n^{4/5}b]} |R_{nk}| = O(n^{-3/5}), \quad a.s.$$

PROOF. Let $0 < U_{n1} < \cdots < U_{nn} < 1$ denote the order statistics of a random sample of size n from uniform (0, 1). Then it follows easily that

$$k^{-1}\sum_1^k Y_{ni}^2 = \{2f(x_0)\}^{-2}k^{-1}\sum_1^k U_{ni}^2 + R_{nk}(1) \tag{12}$$

and

$$\begin{aligned} k^{-1}\sum_1^k U_{ni}^2 &= k^{-1}\sum_1^k (i/n)^2 + R_{nk}(2) \\ &= (1/3)(k/n)^2 + R_{nk}(3) + R_{nk}(2). \end{aligned} \tag{13}$$

By the law of iterated logarithm [see Csörgő and Révész (1981), page 157], it can be shown that

$$\max_{k \le [n^{4/5}b]} |R_{nk}(2)| = O\left(n^{-7/10}\sqrt{\log\log n}\right), \quad \text{a.s.},$$

and it is easy to show that $\max_{k \le [n^{4/5}b]}|R_{nk}(1)| = O(n^{-4/5})$, a.s. and $\max_{k \le [n^{4/5}b]}|R_{nk}(3)| = O(n^{-6/5})$, a.s. Thus combining (12) and (13) the lemma is proved. □

LEMMA 3. *The following expansions hold for the conditional pdf $g^*(z|y)$ and the conditional cdf $G^*(z|y)$:*

$$g^*(z|y) = g(z) + \tfrac{1}{2}y^2 q(z) + y^3 r(y,z),$$

$$G^*(z|y) = G(z) + \tfrac{1}{2}y^2 Q(z) + y^3 R(y,z),$$

where

$$g(z) = g(z|x_0), \qquad G(z) = G(z|x_0),$$

$$q(z) = g_{xx}(z|x_0) + 2f'(x_0)g_x(z|x_0)/f(x_0),$$

$$Q(z) = G_{xx}(z|x_0) + 2f'(x_0)G_x(z|x_0)/f(x_0),$$

and there exist $\varepsilon > 0$ *and* $M < \infty$ *such that* $|q(z)|$, $|Q(z)|$, $|r(y,z)|$ *and* $|R(y,z)|$ *are all bounded by* M *for* $0 \le y \le \varepsilon$ *and* $|z - \xi| \le \varepsilon$.

PROOF. The expansions for $g^*(z|y)$ and $G^*(z|y)$ follow easily by expanding $f(x_0 \pm y)$, $g(z|x_0 \pm y)$ and $G(z|x_0 \pm y)$ about $y = 0$, and substituting them in (1). The boundedness of $|q(z)|$, $|Q(z)|$, $|r(y,z)|$ and $|R(y,z)|$ follows from conditions 1 and 2. □

6. Proof of Theorem N1: Bias in $\hat{\xi}_{nk}$. Recall the target of $\hat{\xi}_{nk}$ is ξ_{nk}, the p-quantile of the random cdf $\overline{G}_{nk}(\cdot) = k^{-1}\Sigma_1^k G^*(\cdot|Y_{ni})$, while ξ is the p-quantile of $G(\cdot)$. The leading term of $\xi_{nk} - \xi$ is nonstochastic with probability 1, which is determined in this section.

LEMMA 4. *For every* B, *there exist* N *and* C *such that in the sample space of infinite sequences* $\{((y_1, z_1), (y_2, z_2), \dots): y_i \ge 0,\ z_i \text{ real}\}$, $Y_{n,[n^{4/5}b]} \le Bn^{-1/5}$ *implies* $\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| \le Cn^{-2/5}$ *for all* $n \ge N$.

PROOF. Fix $B < \infty$ and $0 < a < b$, and assume that $Y_{n,[n^{4/5}b]} \le Bn^{-1/5}$. By (11), it is enough to show the existence of N and C such that for $n \ge N$ and $k \in I_n(a,b)$,

$$\overline{G}_{nk}(\xi - Cn^{-2/5}) \le G(\xi) \le \overline{G}_{nk}(\xi + Cn^{-2/5}).$$

For this, choose N and C so that

$$\max(BN^{-1/5}, CN^{-2/5}) \le \min(\varepsilon, \tfrac{1}{2}), \tag{14}$$

and use Lemma 3 to obtain

$$\left|\overline{G}_{nk}(\xi \pm Cn^{-2/5}) - G(\xi \pm Cn^{-2/5})\right| \le MB^2 n^{-2/5}, \tag{15}$$

for $n \ge N$. Moreover, since by condition 2, $G(\xi + Cn^{-2/5}) \ge G(\xi) + \frac{1}{2}Cn^{-2/5}g(\xi)$ and $G(\xi - Cn^{-2/5}) \le G(\xi) - \frac{1}{2}Cn^{-2/5}g(\xi)$ for $Cn^{-2/5} \le \varepsilon$, (15) implies

$$\overline{G}_{nk}(\xi + Cn^{-2/5}) \ge G(\xi) + \tfrac{1}{2}Cn^{-2/5}g(\xi) - MB^2n^{-2/5},$$
$$\overline{G}_{nk}(\xi - Cn^{-2/5}) \le G(\xi) - \tfrac{1}{2}Cn^{-2/5}g(\xi) + MB^2n^{-2/5},$$

for $n \ge N$. The lemma is proved by choosing $C > 2MB^2/g(\xi)$ and then choosing N so as to satisfy (14). □

COROLLARY. *For* $0 < a < b$, $\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| = O(n^{-2/5})$, *a.s.*

PROOF. Take $B > b/f(x_0)$ and apply Lemma 4 using C and N appropriately determined by B. Then

$$\sum_{n=N}^{\infty} P\Big[\max_{k \in I_n(a,b)} |\xi_{nk} - \xi| > Cn^{-2/5}\Big] \le \sum_{n=N}^{\infty} P\big[Y_{n,[n^{4/5}b]} > Bn^{-1/5}\big]$$
$$\le \sum_{n=N}^{\infty} \exp\big[-2n^{3/5}\{Bf(x_0) - b\}^2\big] < \infty.$$

□

We now determine the leading term of $\xi_{nk} - \xi$.

LEMMA 5. *For* $0 < a < b$,

$$\max_{k \in I_n(a,b)} \left|\xi_{nk} - \xi - \beta(\xi)(k/n)^2\right| = O(n^{-3/5}), \quad a.s.,$$

where $\beta(\xi) = -Q(\xi)\{24 f^2(x_0)g(\xi)\}^{-1}$.

PROOF. By (11) and Lemma 3,

$$\begin{aligned} G(\xi) = \overline{G}_{nk}(\xi_{nk}) &= k^{-1}\sum_1^k \left[G_{ni}(\xi) + (\xi_{nk} - \xi)g_{ni}(z_{ni})\right] \\ &= G(\xi) + k^{-1}\sum_1^k \left\{\tfrac{1}{2}Y_{ni}^2 Q(\xi) + Y_{ni}^3 R(Y_{ni}, \xi)\right\} \\ &\quad + (\xi_{nk} - \xi)\left[g(\xi) + k^{-1}\sum_1^k \left\{(g(z_{ni}) - g(\xi)) + \tfrac{1}{2}Y_{ni}^2 q(z_{ni})\right.\right. \\ &\qquad\qquad \left.\left. + Y_{ni}^3 r(Y_{ni}, z_{ni})\right\}\right], \end{aligned}$$

where for each i, z_{ni} lies between ξ_{nk} and ξ, so that

$$(16) \qquad \max_{k \in I_n(a,b)} \max_{1 \le i \le k} |z_{ni} - \xi| \le \max_{k \in I_n(a,b)} |\xi_{nk} - \xi| = O(n^{-2/5}), \quad \text{a.s.}$$

Hence

$$\begin{aligned} \xi_{nk} - \xi &= -\frac{1}{2}\,\frac{Q(\xi)k^{-1}\Sigma_1^k Y_{ni}^2 + 2k^{-1}\Sigma_1^k Y_{ni}^3 R(Y_{ni}, \xi)}{g(\xi) + k^{-1}\Sigma_1^k\left[\{g(z_{ni}) - g(\xi)\} + \frac{1}{2}Y_{ni}^2 q(z_{ni}) + Y_{ni}^3 r(Y_{ni}, z_{ni})\right)} \\ &= -\frac{1}{2}\,\frac{Q(\xi)\left[\{12 f^2(x_0)\}^{-1}(k/n)^2 + R_{nk}(1)\right] + R_{nk}(2)}{g(\xi) + R_{nk}(3)}, \end{aligned}$$

where $\max_{k \in I_n(a,b)}|R_{nk}(1)| = O(n)^{-3/5}$, a.s. by Lemma 2, $\max_{k \in I_n(a,b)}|R_{nk}(2)| = O(n^{-3/5})$, a.s., by Lemmas 1 and 3, and $\max_{k \in I_n(a,b)}|R_{nk}(3)| = O(n^{-2/5})$, a.s., by (16), Lemmas 1 and 3, and condition 2. Since $(k/n)^2 \le n^{-2/5}b^2$ for $k \in I_n(a, b)$, the lemma is proved. □

7. Proof of Theorem N1: Conclusion. The first representation of $\hat{\xi}_{nk} - \xi_{nk}$ rests on Lemmas 8 and 9, which run parallel to Bahadur's proof [Bahadur (1966)]. However, we start with a lemma which provides an exponential bound for deviation of sums of independent Bernoulli variables from their mean and then prove another lemma dealing with fluctuations of $\hat{G}_{nk}(\cdot) - \overline{G}_{nk}(\cdot)$.

LEMMA 6. *Let $U_{n1}, \dots, U_{nn}$ be independent Bernoulli variables with $P(U_{ni} = 1) = \pi_{ni}$. Then*

$$\text{(a)}\quad P\left[\left|n^{-1}\sum_{1}^{n}(U_{ni} - \pi_{ni})\right| > t_n\right] \le 2\exp\left[-\tfrac{1}{2}nt_n^2\Big/\left\{\max_{1\le i\le n}\pi_{ni} + t_n\right\}\right].$$

In particular, if $t_n/\max_{1\le i\le n}\pi_{ni} \to 0$, then for large n,

$$\text{(b)}\qquad P\left[\left|n^{-1}\sum_{1}^{n}(U_{ni} - \pi_{ni})\right| > t_n\right] \le 2\exp\left[-\tfrac{1}{4}nt_n^2\Big/\max_{1\le i\le n}\pi_{ni}\right]$$

and if $\max_{1\le i\le n}\pi_{ni}/t_n \to 0$, then for large n,

$$\text{(c)}\qquad P\left[\left|n^{-1}\sum_{1}^{n}(U_{ni} - \pi_{ni})\right| > t_n\right] \le 2\exp\left[-\tfrac{1}{4}nt_n\right].$$

PROOF. The first inequality is a simplified version of Bernstein's inequality [see Uspensky (1937), page 205], from which the other two follow as special cases. □

LEMMA 7. *Suppose ζ_{nk} are $\mathscr{A}$-measurable random variables with $|\zeta_{nk} - \xi_{nk}| \le Cn^{-2/5}\log n = \varepsilon_n(C)$. Then for any γ, there exists M such that*

$$\sum_{n=1}^{\infty} n^{\gamma}\max_{k\in I_n(a,b)} P\left[\left|\{\hat{G}_{nk}(\zeta_{nk}) - \hat{G}_{nk}(\xi_{nk})\} - \{\overline{G}_{nk}(\zeta_{nk}) - \overline{G}_{nk}(\xi_{nk})\}\right| > Mn^{-3/5}\log n\right] < \infty.$$

PROOF. Write $U_{nki} = 1(Z_{ni} \le \zeta_{nk}) - 1(Z_{ni} \le \xi_{nk})$ and $\mu_{nki} = G_{ni}(\zeta_{nk}) - G_{ni}(\xi_{nk}) = E(U_{nki}|\mathscr{A})$. Then

$$\{\hat{G}_{nk}(\zeta_{nk}) - \hat{G}_{nk}(\xi_{nk})\} - \{\overline{G}_{nk}(\zeta_{nk}) - \overline{G}_{nk}(\xi_{nk})\} = k^{-1}\sum_{i=1}^{k}(U_{nki} - \mu_{nki}).$$

Choose $B > b/f(x_0)$, and for each n, let $S_n = \{Y_{n,[n^{4/5}b]} \le B^{-1/5}\}$. Since $|\zeta_{nk} - \xi_{nk}| \le \varepsilon_n(C) = Cn^{-2/5}\log n$, Lemma 4 implies that there exist C' and N such that for $n \ge N$ and for z lying between ζ_{nk} and ξ_{nk}, $|z - \xi| \le Cn^{-2/5}\log n + C'n^{-2/5} \le 2\varepsilon_n(C)$ holds on the set S_n. Using Lemma 3, we now conclude that when n is large, then on S_n,

$$\begin{aligned}\max_{1\le i\le[n^{4/5}b]}|\mu_{nki}| &\le \varepsilon_n(C)\sup_{|z-\xi|\le 2\varepsilon_n(C)} g(z) + B^2n^{-2/5}\sup_{|z-\xi|\le 2\varepsilon_n(C)}|Q(z)| \\ &\quad + 2B^3n^{-3/5}\sup_{0\le y\le Bn^{-1/5},\ |z-\xi|\le 2\varepsilon_n(C)}|R(y,z)| \\ &\le \tfrac{3}{2}\varepsilon_n(c)g(\xi) + MB^2n^{-2/5} + 2MB^3n^{-3/5} \le 2g(\xi)\varepsilon_n(c),\end{aligned}$$

since $\varepsilon_n(C) = C_n^{-2/5}\log n$ dominates $MB^2n^{-2/5} + 2MB^3n^{-3/5}$. Now use Lemma 6(b), replacing $\max_{1\le i\le[n^{4/5}b]}|\mu_{nki}|$ by its upper bound obtained above, to conclude that for large n,

$$\max_{k\in I_n(a,b)} P\left[\left|k^{-1}\sum_1^k (U_{nki}-\mu_{nki})\right| > Mn^{-3/5}\log n\right]$$
$$= \max_{k\in I_n(a,b)} EP\left[\left|k^{-1}\sum_1^k (U_{nki}-\mu_{nki})\right| > Mn^{-3/5}\log n \,\middle|\, \mathscr{A}\right]$$
$$\le 2\exp\left[-M^2 a\{5Cg(\xi)\}^{-1}\log n\right] + P(S_n^c).$$

To complete the proof, observe that $\sum_{n=1}^{\infty} n^{\gamma - M^2 a\{5Cg(\xi)\}^{-1}} < \infty$ for sufficiently large M, and $\sum_{n=1}^{\infty} P(S_n^c) < \infty$ by Lemma 1. □

We now define $a_n = n^{-2/5}\log n$, $b_n = n^{1/5}$ and divide the interval $[\xi_{nk} - a_n, \xi_{nk} + a_n]$ into $2b_n$ equal intervals:

$$J_{nk,r} = [\xi_{nk} + ra_n/b_n, \xi_{nk} + (r+1)a_n/b_n] = [\eta_{nk,r}, \eta_{nk,r+1}],$$
$$r = -b_n, \dots, -1, 0, 1, \dots, b_n - 1,$$

each of length $a_n/b_n = n^{-3/5}\log n$. Let

$$H_{nk}(z) = \{\hat{G}_{nk}(z) - \hat{G}_{nk}(\xi_{nk})\} - \{\overline{G}_{nk}(z) - \overline{G}_{nk}(\xi_{nk})\},$$

$$(17) \qquad H_{nk}^* = \sup_{|z-\xi_{nk}|\le a_n} |H_{nk}(z)| = \max_{-b_n\le r\le b_n-1}\ \sup_{z\in J_{nk,r}} |H_{nk}(z)|,$$

$$H_n^* = \max_{k\in I_n(a,b)} |H_{nk}^*|.$$

Lemma 8. $P[\max_{k\in I_n(a,b)}|\hat{\xi}_{nk} - \xi_{nk}| > a_n \text{ i.o.}] = 0$.

Proof. $\hat{\xi}_{nk} \le \xi_{nk} - a_n$ implies

$$k^{-1}\sum_1^k \{1(Z_{ni}\le \xi_{nk} - a_n) - G_{ni}(\xi_{nk} - a_n)\} \ge [kp]/k - \overline{G}_{nk}(\xi_{nk} - a_n).$$

Fix $B > b/f(x_0)$ and let $S_n = \{Y_{n,[n^{4/5}b]} \le Bn^{-1/5}\}$. Then by Lemma 3,

$$\min_{k\in I_n(a,b)} \{[kp]/k - \overline{G}_{nk}(\xi_{nk} - a_n)\} \ge \tfrac{1}{2}g(\xi)a_n$$

on the set S_n when n is large. Hence for large n, by Theorem 1 of Hoeffding (1963), we have

$$P\left[\min_{k\in I_n(a,b)} (\hat{\xi}_{nk} - \xi_{nk}) \le -a_n \text{ i.o.}\right] = 0.$$

In the same way,

$$P\left[\max_{k\in I_n(a,b)} (\hat{\xi}_{nk} - \xi_{nk}) \ge a_n \text{ i.o.}\right] = 0,$$

and the lemma is proved. □

LEMMA 9. $P[H_n^* > Cn^{-3/5} \log n \ i.o.] = 0$ *for large* C.

PROOF. It follows from the monotonicity of $\hat{G}_{nk}(\cdot)$ and $\overline{G}_{nk}(\cdot)$ that for $z \in J_{nk,r} = [\eta_{nk,r}, \eta_{nk,r+1}]$,

$$H_{nk}(\eta_{nk,r}) - \alpha_{nk,r} \le H_{nk}(z) \le H_{nk}(\eta_{nk,r+1}) + \alpha_{nk,r},$$

where $H_{nk}(\cdot)$ is given by (17), and

$$\alpha_{nk,r} = \overline{G}_{nk}(\eta_{nk,r+1}) - \overline{G}_{nk}(\eta_{nk,r}).$$

Hence

$$H_{nk}^* = \sup_{|z-\xi_{nk}| \le a_n} |H_{nk}(z)| \le \max_{-b_n \le r \le b_n} |H_{nk}(\eta_{nk,r})| + \max_{-b_n \le r \le b_n - 1} \alpha_{nk,r}.$$

Let $S_n = \{Y_{n,[n^{4/5}b]} \le Bn^{-1/5}\}$ as in the previous proofs. Then by Lemmas 3 and 4,

$$\max_{-b_n \le r \le b_n - 1} \alpha_{nk,r} \le 2g(\xi) n^{-3/5} \log n$$

on the set S_n when n is large. Hence

$$\begin{aligned} &P\big[H_n^* > \{M + 2g(\xi)\} n^{-3/5} \log n\big] \\ &\quad \le P\Big[\max_{k \in I_n(a,b)} \max_{-b_n \le r \le b_n} |H_{nk}(\eta_{nk,r})| > Mn^{-3/5} \log n\Big] + P(S_n^c) \\ &\quad \le 2n(b-a) \max_{k \in I_n(a,b)} \max_{-b_n \le r \le b_n} P\big[|H_{nk}(\eta_{nk,r})| > Mn^{-3/5} \log n\big] + P(S_n^c). \end{aligned}$$

But $\max_{-b_n \le r \le b_n} |\eta_{nk,r} - \xi_{nk}| \le n^{-2/5} \log n$, and by Lemma 7,

$$\sum_{n=1}^{\infty} n \max_{k \in I_n(a,b)} \max_{-b_n \le r \le b_n} P\big[|H_{nk}(\eta_{nk,r})| > Mn^{-3/5} \log n\big] < \infty,$$

while $\sum_{n=1}^{\infty} P(S_n^c) < \infty$ by Lemma 1. This completes the proof. □

By Lemmas 8 and 9, we now have

$$\begin{aligned} p - \hat{G}_{nk}(\xi_{nk}) &= \overline{G}_{nk}(\hat{\xi}_{nk}) - \overline{G}_{nk}(\xi_{nk}) + R_{nk}(1) \\ &= (\hat{\xi}_{nk} - \xi_{nk}) \bar{g}_{nk}(\xi_{nk}^*) + R_{nk}(1), \end{aligned} \tag{18}$$

where ξ_{nk}^* lies between $\hat{\xi}_{nk}$ and ξ_{nk}, and

$$\max_{k \in I_n(a,b)} |R_{nk}(1)| = O(n^{-3/5} \log n), \quad \text{a.s.} \tag{19}$$

By the corollary to Lemma 4 and Lemma 8, $\max_{k \in I_n(a,b)} |\xi_{nk}^* - \xi| = O(n^{-2/5} \log n)$, a.s. Consequently, Lemmas 1 and 3 now imply

$$\max_{k \in I_n(a,b)} |\bar{g}_{nk}(\xi_{nk}^*) - g(\xi)| = O(n^{-2/5} \log n), \text{ a.s.} \tag{20}$$

Furthermore, it follows from Theorem 1 of Hoeffding (1963) that

$$p - \hat{G}_{nk}(\xi_{nk}) = O(n^{-2/5} \log n), \quad \text{a.s.} \tag{21}$$

From (18), (19), (20) and (21), we have

$$\max_{k \in I_n(a,b)} \left| (\hat{\xi}_{nk} - \xi_{nk}) - \{g(\xi)\}^{-1}[p - \hat{G}_{nk}(\xi_{nk})] \right| = O(n^{-3/5} \log n), \quad \text{a.s.}$$

Since

$$p - \hat{G}_{nk}(\xi_{nk}) = k^{-1} \sum_1^k [1(Z_{ni} > \xi_{nk}) - \{1 - G_{ni}(\xi_{nk})\}],$$

we now have the following representation:

$$\text{(22a)} \quad \begin{aligned} \hat{\xi}_{nk} &= \xi_{nk} + \{kg(\xi)\}^{-1} \sum_1^k [1(Z_{ni} > \xi_{nk}) - \{1 - G_{ni}(\xi_{nk})\}] + R_{nk}, \\ &\max_{k \in I_n(a,b)} |R_{nk}| = O(n^{-3/5} \log n), \quad \text{a.s.} \end{aligned}$$

This representation can be easily modified to two other slightly different forms, viz.,

$$\text{(22b)} \quad \hat{\xi}_{nk} = \xi_{nk} + \{kg(\xi)\}^{-1} \sum_1^k [1(Z_{ni} > \xi) - \{1 - G_{ni}(\xi)\}] + R_{nk}$$

and

$$\text{(22c)} \quad \hat{\xi}_{nk} = \xi_{nk} + \{kg(\xi)\}^{-1} \sum_1^k [1(Z_{ni}^* > \xi) - \{1 - G(\xi)\}] + R_{nk},$$

where $\max_{k \in I_n(a,b)} |R_{nk}| = O(n^{-3/5} \log n)$, a.s., in both (22b) and (22c), and $Z_{ni}^* = G^{-1} \circ G_{ni}(Z_{ni})$ and $G(\cdot) = G(\cdot|x_0)$ is the conditional cdf of Z given $X = x_0$.

Combine Lemma 5 with (22c) and note that $G(\xi) = p$ to complete the proof of Theorem N1. □

8. Proof of Theorem K1. The kernel estimator $\tilde{\xi}_{nh}$ can be regarded as the NN estimator $\hat{\xi}_{nK_n(h)}$ in which $K_n(h)$ is a random integer given by (4). A formal substitution for k by $K_n(h)$ in the representation given in Theorem N1 leads to

$$\text{(23)} \quad \begin{aligned} \tilde{\xi}_{nh} - \xi &= \beta(\xi)\{K_n(h)/n\}^2 \\ &+ \{K_n(h) g(\xi)\}^{-1} \sum_1^{K_n(h)} [1(Z_{ni}^* > \xi) - (1 - p)] + R_{nK_n(h)}. \end{aligned}$$

However, this is of no use unless we can show that

(a) $\sup_{h \in J_n(c,d)} |R_{nK_n(h)}|$ converges at a fast rate, where $J_n(c,d) = [n^{-1/5}c, n^{-1/5}d]$, $0 < c < d$, and

(b) in the first two terms of the RHS of (23), $K_n(h)$ can be replaced by the leading term of its deterministic component without slowing down the rate of convergence of the remainder term.

To establish (a) and (b), we first examine the magnitude of $\{K_n(h) - nhf(x_0)\}$ in the following lemma which is proved by routine calculations.

LEMMA 10. *Let* $\Delta_n(h) = K_n(h) - nhf(x_0)$. *Then*

$$\sup_{h \in J_n(c,d)} |\Delta_n(h)| = O(n^{2/5} \log n), \quad a.s.$$

From Lemma 10 and the fact that $\max_{k \in I_n(a,b)} |R_{nk}| = O(n^{-3/5} \log n)$, a.s., it now follows that

$$\sup_{h \in J_n(c,d)} |R_{nK_n(h)}| = O(n^{-3/5} \log n), \quad \text{a.s.}$$

We now consider the first two terms on the RHS of (23). Of these,

$$\beta(\xi)\{K_n(h)/n\}^2 = \beta(\xi) f^2(x_0) h^2 + R'_{nh},$$

where

$$R'_{nh} = \beta(\xi) f^2(x_0) h^2 [\Delta_n(h)/\{nhf(x_0)\}][2 + \Delta_n(h)/\{nhf(x_0)\}],$$

and by Lemma 10,

$$\sup_{h \in J_n(c,d)} |R'_{nh}| = O(n^{-4/5} \log n), \quad \text{a.s.}$$

To examine the other term, let

$$U_{ni} = 1(Z^*_{ni} > \xi) - (1-p), \qquad m_n(h) = [nhf(x_0)].$$

Then

$$\{K_n(h) g(\xi)\}^{-1} \sum_1^{K_n(h)} U_{ni} = \{m_n(h) g(\xi)\}^{-1} \sum_1^{m_n(h)} U_{ni} + R''_{nh} + R'''_{nh},$$

$$(24) \quad \begin{aligned} R''_{nh} &= -\{\Delta_n(h)/K_n(h)\}\{m_n(h) g(\xi)\}^{-1} \sum_1^{m_n(h)} U_{ni}, \\ R'''_{nh} &= \{1 - \Delta_n(h)/K_n(h)\}\{m_n(h) g(\xi)\}^{-1} \left[\sum_1^{K_n(h)} U_{ni} - \sum_1^{m_n(h)} U_{ni} \right], \end{aligned}$$

where $U_{n1}, \dots, U_{nn}$ are conditionally independent given $\mathscr{A}$, with $E(U_{ni} | \mathscr{A}) = 0$. Of the two remainder terms $\sup_{h \in J_n(c,d)} |R''_{nh}| = O(n^{-3/5} \log n)$, a.s., by Lemma 10 and Theorem 1 of Hoeffding (1963). Now let $h_{n0} < h_{n1} < \cdots < h_{n\nu_n}$ denote the jump points of $m_n(h) = [nhf(x_0)]$ in $J_n(c,d)$. Since for each j and for all $h_{nj} \le h < h_{n,j+1}$,

$$\left| \sum_1^{K_n(h)} U_{ni} - \sum_1^{m_h(h)} U_{ni} \right| \le \left| \sum_1^{K_n(h_{nj})} U_{ni} - \sum_1^{m_n(h_{nj})} U_{ni} \right| + \{K_n(h_{n,j+1}) - K_n(h_{nj})\}$$

we only need to verify

$$(25) \qquad \max_{0 \le j \le \nu_n} \left| \sum_{1}^{K_n(h_{nj})} U_{ni} - \sum_{1}^{m_n(h_{nj})} U_{ni} \right| = O(n^{1/5} \log n), \quad \text{a.s.},$$

$$(26) \qquad \max_{0 \le j \le \nu_n} \{K_n(h_{n,j+1}) - K_n(h_{nj})\} = O(n^{1/5} \log n), \quad \text{a.s.},$$

in order to conclude that in (24), $\sup_{h \in J_n(c,d)} |R'''_{nh}| = O(n^{-3/5} \log n)$, a.s. To prove (25) and (26), note that $h_{n,j+1} - h_{nj} \le \{nf(x_0)\}^{-1}$ and $\nu_n \le n^{4/5}(d-c) f(x_0)$. Now $K_n(h_{n,j+1}) - K_n(h_{nj})$ is binomial (n, π_{nj}) with $\pi_{nj} = O(n^{-1})$, and (26) follows from Lemma 6(b). Finally, apply Hoeffding's inequality [Hoeffding (1963)], use Lemma 10 and the fact that $\nu_n = O(n^{4/5})$ to prove (25). This completes the proof of Theorem K1. The details are as in Bhattacharya and Gangopadhyay (1987). □

REFERENCES

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

BHATTACHARYA, P. K. (1963). On an analog of regression analysis. *Ann. Math. Statist.* **34** 1459–1473.

BHATTACHARYA, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics. *Ann. Statist.* **2** 1034–1039.

BHATTACHARYA, P. K. and GANGOPADHYAY, A. K. (1987). Kernel and nearest neighbor estimation of a conditional quantile. Technical Report 104, Intercollege Div. of Statist., Univ. California, Davis.

BHATTACHARYA, P. K. and MACK, Y. P. (1987). Weak convergence of k-NN density and regression estimators with varying k and applications. *Ann. Statist.* **15** 976–994.

CHENG, K. F. (1983). Nonparametric estimators for percentile regression functions. *Comm. Statist. Theory Methods* **12** 681–692.

CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics.* Academic, New York.

GANGOPADHYAY, A. K. (1987). Nonparametric estimation of conditional quantile function. Ph.D. dissertation, Univ. California, Davis.

GIKHMAN I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes.* W. B. Saunders, Philadelphia.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

HOGG, R. V. (1975). Estimates of percentile regression lines using salary data. *J. Amer. Statist. Assoc.* **70** 56–59.

KRIEGER, A. M. and PICKANDS, J., III (1981). Weak convergence and efficient density estimation at a point. *Ann. Statist.* **9** 1066–1078.

STONE, C. J. (1977). Consistent nonparametric regrpssion. *Ann. Statist.* **5** 595–645.

STUTE, W. (1986). Conditional empirical processes. *Ann. Stistist.* **14** 638–647.

USPENSKY, J. V. (1937). *Introduction to Mathematical Probability.* McGraw-Hill, New York.

WOODROOFE, M. (1970). On choosing a delta-sequence. *Ann. Math. Statist.* **41** 1665–1671.

DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616

DEPARTMENT OF MATHEMATICS
BOSTON UNIVERSITY
111 CUMMINGTON STREET
BOSTON, MASSACHUSETTS 02215