A Kernel Estimator of a Conditional Quantile

Xiaojing Xiang

Biostatistics Department, Ciba Pharmaceuticals, Summit, New Jersey 07901

Let \((X_1, Y_1), (X_2, Y_2), \ldots\) be two-dimensional random vectors which are independent and distributed as \((X, Y)\). For \(0 < p < 1\), let \(\xi(p \mid x)\) be the conditional \(p\)th quantile of \(Y\) given \(X = x\); that is, \(\xi(p \mid x) = \inf\{ y : P(Y \leq y \mid X = x) \geq p \}\). We consider the problem of estimating \(\xi(p \mid x)\) from the data \((X_1, Y_1), (X_2, Y_2), \ldots (X_n, Y_n)\). In this paper, a new kernel estimator of \(\xi(p \mid x)\) is proposed. The asymptotic normality and a law of the iterated logarithm are obtained.

1. INTRODUCTION

Let \((X_1, Y_1), (X_2, Y_2), \ldots\) be two-dimensional random vectors which are independent and distributed as \((X, Y)\). For \(0 < p < 1\), let \(\xi(p \mid x)\) be the conditional \(p\)th quantile of \(Y\) given \(X = x\); that is,

\[
\xi(p \mid x) = \inf\{ y : P(Y \leq y \mid X = x) \geq p \}. \tag{1.1}
\]

We consider the problem of estimating \(\xi(p \mid x)\) from the data \((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\). Usefulness of conditional quantile functions as good descriptive statistics has been discussed by Hogg [7] who calls them percentile regression lines. The estimation of conditional quantile has received the attention of several authors, Bhattacharya [2], Stone [12], and more recently Stute [13], Bhattacharya and Gangopadhyay [3], and Mehra, Rao, and Upadrashta [9]. The estimator proposed in [3] is basically a conditional version of the sample quantile estimator in the unconditional case. Using an approach similar to that in Bahadur [1], Bhattacharya and Gangopadhyay [3] showed the asymptotic normality and established a Bahadur-type representation of the estimator. The estimator proposed in [9] is a smooth conditional quantile estimator. Mehra, Rao, and Upadrashta [9] proved the asymptotic normality and found an almost sure convergence rate for their estimator.
Let $F(y, x) = P(Y \leq y, X \leq x)$ be the joint distribution function of $Y$ and $X$. Let $G(x) = P(X \leq x)$ be the distribution function of $X$. The empirical distribution functions corresponding to $F(y, x)$ and $G(x)$ are

$$F_n(y, x) = \frac{1}{n} \sum_{i=1}^{n} I(Y_i \leq y, X_i \leq x)$$

(1.2)

and

$$G_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) = F_n(\infty, x),$$

(1.3)

respectively, where $I(A)$ denotes the indicator function of set $A$. Let

$$z_n(y, x) = \sqrt{n} F_n(y, x) - F(y, x)$$

(1.4)

and

$$t_n(x) = \sqrt{n} G_n(x) - G(x)$$

(1.5)

be the corresponding empirical processes. Write $h(y, x) = (\partial \partial x) F(y, x)$, $f(y, x) = (\partial^2 \partial x \partial y) F(y, x)$, $g(x) = G'(x)$, and $m(y \mid x) = P(Y \leq y \mid X = x)$. The kernel estimators of $h(y, x)$ and $g(x)$ are

$$\hat{h}_n(y, x) = \frac{1}{h_n} \int K \left( \frac{x-u}{h_n} \right) dF_n(y, u)$$

(1.6)

and

$$\hat{g}_n(x) = \frac{1}{h_n} \int K \left( \frac{x-u}{h_n} \right) dG_n(u),$$

(1.7)

respectively, for an appropriate kernel function $K(x)$ and a bandwidth $h_n$. Using the notations in Horváth and Yandell [8], the kernel estimator of $m(y \mid x)$ is

$$\hat{m}_n(y \mid x) = \frac{\hat{h}_n(y, x)}{\hat{g}_n(x)} = \sum_{i=1}^{n} I(Y_i \leq y) K \left( \frac{x-X_i}{h_n} \right) / \sum_{i=1}^{n} K \left( \frac{x-X_i}{h_n} \right),$$

(1.8)

being a kernel type conditional empirical function. $\hat{m}_n(y \mid x)$ is nondecreasing and right continuous in $y$.

In this article, we propose a new kernel estimator of a conditional quantile which is a conditional version of Parzen’s estimator in the univariate
case (see Parzen [10]). Let \( \hat{\mu}_n^{-1}(y \mid x) = \inf\{t : \hat{\mu}_n(t \mid x) \geq y\} \). Our estimator of \( \zeta(p \mid x) \) is of the form

\[
\hat{\zeta}_n(p \mid x) = \frac{1}{a_n} \int \hat{\mu}_n^{-1}(y \mid x) w \left( \frac{y - p}{a_n} \right) dy
\]

(1.9)

for an appropriate kernel function \( w(x) \) and a bandwidth \( a_n \). Using an approach different from that in [9], we shall prove the asymptotic normality for this estimator under weaker assumptions (see Remark 2.2 of Theorem 2.1). A law of the iterated logarithm for \( \hat{\zeta}_n(p \mid x) \) is also obtained. This LIL improves the almost sure convergence result in [9].

The present paper is organized in the following manner. Section 2 gives the main results and the remarks. The proofs of the theorems are provided in Section 3.

2. MAIN RESULTS

In this paper, we assume that for \( x = x_0 \),

\[
g(x_0) > 0
\]

(2.1)

and

\[
f(\zeta(p \mid x_0), x_0) > 0.
\]

(2.2)

We further assume that there exist neighborhoods of \( x_0 \) and \( \zeta(p \mid x_0) \), say \( I_{x_0} \) and \( I_\zeta \), for \( m \geq 2 \),

\[
\sup_{(y, u) \in I_{x_0} \times I_\zeta} \left| \frac{\partial^m}{\partial u^m} h(y, u) \right| < \infty, \quad \sup_{u \in I_\zeta} |g^{(m)}(u)| < \infty,
\]

(2.3)

\[
\sup_{(y, u) \in I_{x_0} \times I_\zeta} \left| \frac{\partial^2}{\partial u^2} h(y, u) \right| < \infty, \quad \frac{\partial^2}{\partial y^2} h(y, u) \text{ is continuous on } I_\zeta \times I_{x_0},
\]

(2.4)

\[
\frac{\partial^m}{\partial u^m} h(y, u) \text{ is continuous on } I_\zeta \times I_{x_0}, \quad g^{(m)}(u) \text{ is continuous on } I_\zeta.
\]

(2.5)

For the kernel functions \( K(x) \) and \( w(x) \), we assume that both \( K(x) \) and \( w(x) \) are symmetric and continuous with compact support in \([-1, 1]\), that \( K(x) \) has bounded variation on \( \mathbb{R} \) and that

\[
|w(x_1) - w(x_2)| \leq C |x_1 - x_2|
\]

(2.6)
for a universal constant $C$ and $x_1, x_2 \in (-\infty, \infty)$. We also assume that
\[ \int_{-1}^{1} w(x) \, dx = 1 \quad (2.7) \]
\[ \int_{-1}^{1} K(x) \, dx = 1, \quad \int_{-1}^{1} x^j K(x) \, dx = 0, \quad j = 1, \ldots, m - 1, \]
\[ \int_{-1}^{1} x^m K(x) \, dx = \sigma_m \neq 0. \quad (2.8) \]

For the sequence of bandwidth $\{h_n\}$, we require that as $n \to \infty$,
\[ h_n \to 0 \quad \text{for some} \quad \gamma_1 > 0, \quad nh_n = O(n^{\gamma_2}) \quad \text{for some} \quad \gamma_2 > 0. \quad (2.9) \]

Finally, we assume that for some $C > 0, 0 < \tau \leq 1$, and $y, z \in I_{ij}$,
\[ \sup_{x \in I_{i0}} |m(y \mid x) - m(z \mid x)| \leq C \left( |m(y \mid x_0) - m(z \mid x_0)|^\tau. \quad (2.10) \]

Let
\[ \beta_1 = \frac{g^{(m)}(x_0)}{f(\xi(p \mid x_0), x_0)} m^2, \quad \beta_2 = -\frac{\sigma_m}{f(\xi(p \mid x_0), x_0) m^2} \frac{\partial^2}{\partial x^m} h(\xi(p \mid x_0), x_0). \]

**Theorem 2.1.** Under the assumptions (2.1)-(2.10), if $h_n = n^{-1/(2m + 1)}$,
\[ a_n = o(n^{-m/(4m + 2)}), \quad a_n = n^{m/(2m + 1)} \log \log n \to \infty \quad \text{as} \quad n \to \infty, \]
\[ n^{m/(2m + 1)} \left( \frac{\xi_n(p \mid x_0)}{\xi(p \mid x_0)} - \frac{\xi_n(p \mid x_0)}{\xi(p \mid x_0)} \right) \rightarrow N(\Delta, \sigma^2), \]
where $\Delta = \beta_1 + \beta_2$ and $\sigma^2 = p(1 - p) g(x_0) \int_{-1}^{1} K^2(u) du / f^2(\xi(p \mid x_0), x_0)$.

**Remark 2.1.** First of all, for each $m \geq 2$, (2.9) holds for $h_n = n^{-1/(2m + 1)}$.

**Remark 2.2.** For $m = 2$, we obtain the same convergence rate as that in [3, 9]. To obtain this result, Bhattacharya and Gangopadhyay [3] imposed extra Hölder conditions in their 2(b); Mehra, Rao, and Upadrashta [9] required that the kernel functions be twice continuously differentiable which rules out some popular kernel functions such as the triangular kernel. Here these extra assumptions are relaxed.

The following result is a law of the iterated logarithm for $\xi_n(p \mid x_0)$.

**Theorem 2.2.** Under the assumptions (2.1)-(2.10), if $h_n = (\log_2 n/n)^{1/(2m + 1)}$,
\[ a_n = o(n^{-m/(4m + 2)}), \quad a_n = n^{m/(2m + 1)} \log_2 n \to \infty \quad \text{as} \quad n \to \infty, \]
\[
\lim_{n \to \infty} \left( \frac{\log_2 n}{n} \right)^{-m/(2m+1)} (\xi_n(p | x_0) - \zeta(p | x_0)) = A + \frac{(2p(1-p) g(x_0) \int K'(u) du)^{1/2}}{f(z(p | x_0), x_0)}.
\]

\[
\lim_{n \to \infty} \left( \frac{\log_2 n}{n} \right)^{-m/(2m+1)} (\xi_n(p | x_0) - \zeta(p | x_0)) = A + \frac{(2p(1-p) g(x_0) \int K'(u) du)^{1/2}}{f(z(p | x_0), x_0)},
\]

where \( \log_2 n = \log \log n \) and \( A \) is given in Theorem 2.1.

Remark 2.3. For \( m = 2 \), the strong consistency rate obtained by Mehra, Rao, and Upadrashta [9] is \( O((\log n/n)^{2/5}) \) (see their Theorem 3.1). They also claim that if one chooses the bandwidth of the order \( O(n^{-1/5}) \) (it should be \( O((\log_2 n/n)^{1/5}) \) actually), the almost sure convergence rate is of the order \( O((\log_2 n/n)^{2/5}) \). In contrast, the result in above Theorem 2.2 is more accurate than that in [9].

3. PROOFS OF THE THEOREMS

Our approach is based on the strong embedding results in [5, 8]. Define

\[
\hat{h}_n(y, x_0) = \frac{1}{h_n} \int K \left( \frac{x_0 - u}{\hat{h}_n} \right) dF(y, u), \quad (3.1)
\]

\[
\hat{g}_n(x_0) = \frac{1}{h_n} \int K \left( \frac{x_0 - u}{\hat{h}_n} \right) dG(u), \quad (3.2)
\]

\[
\hat{m}_n(y | x_0) = \frac{\hat{h}_n(y, x_0)}{\hat{g}_n(x_0)}, \quad (3.3)
\]

and

\[
\beta_n(y | x_0) = \sqrt{\hat{m}_n(y | x_0) - \hat{m}_n(y | x_0)}. \quad (3.4)
\]

We first state two useful lemmas. The following lemma is the consequence of the theorem of Csörgö and Harváth [5] and the proof of Theorem 3.1 of Horváth and Yandell [8].
Lemma 3.1. Under the conditions (2.1), (2.3), (2.8), and (2.9) there exists a sequence of Brownian bridges \( \{ B_n(t), 0 \leq t \leq 1 \} \) such that for a constant \( \beta > 0 \),

\[
\sup_{y \in K_x} \left| \beta_n(y \mid x_0) - \left( \int K^2(u) \, du / g(x_0) \right)^{1/2} B_n(m(y \mid x_0)) \right| = O(n^{-\beta}) \quad \text{a.s.}
\]

and

\[
\sqrt{mn} E B_n(x) B_m(y) = \min(m, n)(\min(x, y) - xy). \tag{3.5}
\]

Lemma 3.2. Assume (2.1), (2.3), (2.5), and (2.8) hold. Then

\[
\sup_{y \in K_x} \left| m_n(y \mid x_0) - m(y \mid x_0) + \frac{\partial^m h(y, x_0) x_m h^m}{m! g'(x_0)} \right| = o(h^m). \tag{36}
\]

Proof. Write

\[
m_n(y \mid x_0) - m(y \mid x_0) = \frac{1}{g(x_0)} (\xi_n(y, x_0) - h(y, x_0))
\]

\[
\quad - \frac{h(y, x_0)}{g'(x_0)} (\tilde{g}_n(x_0) - g(x_0)) + \frac{h(y, x_0)}{g'(x_0)} \frac{\partial h}{\partial x^m} (\tilde{g}_n(x_0) - g(x_0))^2
\]

\[
\quad - \frac{1}{g(x_0)} (\tilde{g}_n(x_0) - g(x_0))(\xi_n(y, x_0) - h(y, x_0)).
\]

From (2.3), (2.5), and (2.8),

\[
\left| \tilde{g}_n(x_0) - g(x_0) - \frac{h_n x_n h^m}{m!} g'(x_0) \right| = o(h^m)
\]

\[
\sup_{y \in K_x} \left| \xi_n(y, x_0) - h(y, x_0) - \frac{h_n x_n h^m}{m!} \frac{\partial h}{\partial x^m} \right| = o(h^m).
\]

The lemma follows easily.

In the following proofs, we use the notation \( b_n \sim d_n \) if and only if \( b_n/d_n \to 1 \) as \( n \to \infty \). Define

\[
\tilde{h}_n(p \mid x_0) = \frac{1}{a_n} \int \tilde{z}(y \mid x_0) w \left( \frac{y - p}{a_n} \right) dy. \tag{3.7}
\]
Proof of Theorem 2.1. Write
\[ h_n(p \mid x_0) - \zeta(p \mid x_0) = I_{1n} + I_{2n}, \]  
with \( I_{1n} = \tilde{h}_n(p \mid x_0) - \tilde{h}_n(p \mid x_0) \) and \( I_{2n} = \tilde{h}_n(p \mid x_0) - \zeta(p \mid x_0) \). Write \( H_n(y, p) = \int_0^1 w(s - p)/a_n \, ds \). From the change of variable theorem (Billingsley, 4, p. 219),

\[ I_{1n} = \frac{1}{a_n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( H_n(m(y \mid x_0), p) - H_n(m(y \mid x_0), p) \right) \, dy \]

\[ = \frac{1}{a_n} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \left( \frac{w(y \mid x_0) - p}{a_n} \right) \left( m(y \mid x_0) - m(y \mid x_0) \right) \right) \, dy \]

\[ = \frac{1}{a_n} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \left( \frac{w(y \mid x_0) - p}{a_n} \right) \left( m(y \mid x_0) - m(y \mid x_0) \right) \right) \, dy \]

\[ = r_{1n} + r_{2n} + r_{3n}. \]

Let \( A(K, g) = \left( \int K^2(u) \, du / g(x_0) \right)^{1/2} \) and

\[ R_n(y \mid x_0) = \tilde{h}_n(y \mid x_0) - \left( \int K^2(u) \, du / g(x_0) \right)^{1/2} B_n(m(y \mid x_0)). \]

From Lemma 3.1, with probability 1,

\[ \sqrt{n} \tilde{m}_n r_{1n} = - \frac{A(K, g)}{a_n} \int \left( \frac{m(y \mid x_0) - p}{a_n} \right) B_n(m(y \mid x_0)) \, dy \]

\[ = \frac{1}{a_n} \int \left( \frac{m(y \mid x_0) - p}{a_n} \right) R_n(y \mid x_0) \, dy \]

\[ = \frac{A(K, g)}{a_n} \int \left( \frac{t - p}{a_n} \right) B_n(t) \frac{g(x_0)}{f(t \mid x_0, x_0)} \, dt \]

\[ - \frac{1}{a_n} \int \left( \frac{t - p}{a_n} \right) R_n(t \mid x_0) \frac{g(x_0)}{f(t \mid x_0, x_0)} \, dt \]
\[
\sim - \frac{A(K, g) g(x_0)}{f(\xi(p \mid x_0), x_0)} \int_{-1}^{1} w(u) B_n(p + a_n u) \, du
\]
and
\[
= - \frac{g(x_0)}{f(\xi(p \mid x_0), x_0)} \int_{-1}^{1} w(u) R_n(\xi(p + a_n u \mid x_0)) \, du.
\]

If \(a_n\) is small enough, \(\xi(p + a_n u \mid x_0) \in I_p\), for \(u \in [-1, 1]\). Hence from Lemma 3.1, with probability 1,
\[
\left| \frac{g(x_0)}{f(\xi(p \mid x_0), x_0)} \int_{-1}^{1} w(u) R_n(\xi(p + a_n u \mid x_0)) \, du \right| = O(n^{-\delta}). \tag{3.9}
\]

On the other hand, from Lévy’s theorem (see Shorack and Wellner, 11, p. 534), with probability 1,
\[
\frac{A(K, g) g(x_0)}{f(\xi(p \mid x_0), x_0)} \int_{-1}^{1} w(u) B_n(p + a_n u) \, du = \frac{A(K, g) g(x_0)}{f(\xi(p \mid x_0), x_0)} B_n(p) + O(a_n^{1/2}(\log a_n)^{1/2}). \tag{3.10}
\]

(3.9) and (3.10) imply, with probability 1,
\[
\sqrt{n} r_n \sim - \frac{A(K, g) g(x_0)}{f(\xi(p \mid x_0))} B_n(p) \tag{3.11}
\]

having a normal distribution. To estimate \(r_{2n}\), we choose an \(\epsilon > 0\) such that \(a = \xi(p - (1 + \epsilon)a_n \mid x_0)\) and \(b = \xi(p + (1 + \epsilon)a_n \mid x_0)\) both are finite. Write
\[
r_{2n} = - \int_{-b}^{b} \left( \int_{-\infty}^{b} \left( \int_{-\infty}^{b} \left( \int_{-\infty}^{b} \left( \int_{-\infty}^{b} \left( w(u) - w\left( \frac{m(y \mid x_0) - p}{a_n} \right) \right) \, dy \right) \, dy \right) \, dy \right) \, dy = S_{1n} + S_{2n} + S_{3n}.
\]
Since \( w(x) \) has compact support in \([-1, 1]\),

\[
|S_{2n}| \leq \int_{-1}^{1} I \left( \frac{\hat{m}_n(y \mid x_0) - p}{a_n} < 1 \right) \int_{-1}^{1} \left| \frac{w(u)}{r(\hat{m}_n(y \mid x_0) - p)/a_n} \right| w(u) \, du \, dy
\]

\[
\leq \int_{-1}^{1} I \left( \frac{\hat{m}_n(b \mid x_0) - p}{a_n} < 1 \right) \int_{-1}^{1} \left| \frac{w(u)}{r(\hat{m}_n(y \mid x_0) - p)/a_n} \right| w(u) \, du \, dy
\]

\[
\leq I \left( \frac{\hat{m}_n(y \mid x_0) - m(b \mid x_0)}{a_n} < -\varepsilon \right) \int_{-1}^{1} \left| \frac{w(u)}{r(\hat{m}_n(y \mid x_0) - p)/a_n} \right| w(u) \, du \, dy.
\]

From Corollary 5.1 of Horváth and Yandell (1988) and \( a_n n^{(2m+1)/\log_2 n} \to \infty \), with probability 1 the indicator function in above last inequality is zero for large \( n \). Hence \( |S_{2n}| = 0 \) a.s. for large \( n \). With the same reason, \( |S_{3n}| = 0 \) a.s. for large \( n \). It follows that

\[
r_{2n} = O(S_{1n})
\]

\[
= O(a_n^{-1}) \sup_{z \in [-1, 1]} |\hat{m}_n(\zeta(p + a_n z \mid x_0) - m(\zeta(p + a_n z \mid x_0))^2)
\]

\[
= O(a_n^{-1}) \sup_{y \in l_n} |\hat{m}_n(y \mid x_0) - \hat{m}_n(y \mid x_0)|^2
\]

\[
+ O(a_n^{-1}) \sup_{y \in l_n} |\hat{m}_n(y \mid x_0) - m(y \mid x_0)|^2.
\]

Again by Corollary 5.1 of Horváth and Yandell [8], together with Lemma 3.2, with probability 1,

\[
\sqrt{n} \hat{m}_n r_{2n} = O(\log_n n + \sqrt{n} h_{n} \hat{m}_n^{-1/2}) = O\left( \frac{\log n}{n^{1/2} a_n^{1/2}} n^{-m-1/(2m+1)} \right).
\]

Hence from \( a_n n^{m/(2m+1)} / \log_2 n \to \infty \) as \( n \to \infty \), we have

\[
\sqrt{n} \hat{m}_n r_{2n} = o(1) \quad \text{a.s.}
\]

For \( r_{3n} \), by Lemma 3.2,

\[
\sqrt{n} \hat{m}_n r_{3n} \sim -\sqrt{n} \int_{-1}^{1} w(u) \left( \frac{g(x_0)}{f(\hat{\zeta}(p \mid x_0), x_0)} \right) \frac{\hat{m}_n((\zeta(p + a_n u \mid x_0) - (p + a_n u) \, du}{\hat{m}_n((\zeta(p \mid x_0), x_0))}
\]

\[
\sim \frac{g(m) \delta_n^{m-1/2} n^{1/2}}{f(\zeta(p \mid x_0), x_0) m!} - \frac{\alpha_m h_{n}^{m+1/2} n^{1/2}}{f(\zeta(p \mid x_0), x_0) m!} \frac{\gamma^{m}}{\hat{m}_n((\zeta(p \mid x_0), x_0))}.
\]
On the other hand, from (2.4), \( I_{2n} = O(a_n^2) \), implying by taking \( a_n = o(n^{-m/(4m+2)}) \)

\[
\sqrt{n h_n} I_{2n} = o(1). \tag{3.15}
\]

Combining (3.11), (3.13)–(3.15), we have

\[
\sqrt{n h_n} \left( \hat{\xi}_n(p | x_0) - \xi(p | x_0) \right) \\
\sim \frac{A(K, g) g(x_0)}{f(\xi(p | x_0), x_0)} B_n(p) \\
+ \frac{g''(x_0) p \sigma_n h_n^m + n^{1/2} n^{1/2}}{f(\xi(p | x_0), x_0) m!} \\
- \frac{\sigma_n h_n^{m + n^{1/2}} n^{1/2}}{f(\xi(p | x_0), x_0) m!} \frac{\partial^m}{\partial x^m} h(\xi(p | x_0), x_0).
\]

By the definitions of \( \beta_i, i = 1, 2 \), and \( h_n = n^{-1/(2m+1)} \), we obtain

\[
n^{m/(2m+1)}(\hat{\xi}_n(p | x_0) - \xi(p | x_0)) \sim \frac{A(K, g) g(x_0)}{f(\xi(p | x_0), x_0)} B_n(p) + \beta_1 + \beta_2 \text{ a.s.}
\]

The proof is complete.

**Proof of Theorem 2.2.** Along with the line of the proof of Theorem 2.1, with probability 1,

\[
\xi_n(p | x_0) - \xi(p | x_0) \sim \frac{A(K, g) g(x_0)}{\sqrt{m h_n f(\xi(p | x_0), x_0)}} B_n(p) + h_n^m (\beta_1 + \beta_2). \tag{3.16}
\]

From (3.5) and a LIL in Córsugó and Hall [6], we have

\[
\lim_{n \to \infty} {\pm \frac{B_n(p)}{(\log n)^{1/2}}} = (2p(1-p))^{1/2} \tag{3.17}
\]

for either choice of sign. Hence, by taking \( h_n = (\log n/n)^{1/(2m+1)} \), the conclusion follows from (3.16) and (3.17).

**ACKNOWLEDGMENTS**

This work was done when I was a student at the Department of Statistics at University of Chicago. I am very grateful to my advisor, Professor S. M. Stigler for his encouragement and advice and to Professor R. R. Bahadur for helpful discussions. Many thanks to the referee and the editor for their help.
REFERENCES