The Annals of Probability
1994, Vol. 22, No. 2, 1078–1091

A LAW OF THE LOGARITHM FOR KERNEL QUANTILE DENSITY ESTIMATORS[1]

BY XIAOJING XIANG

University of Oregon and University of Chicago

In this article we derive a law of the logarithm for the maximal deviation between two kernel-type quantile density estimators and the true underlying quantile density function in the randomly right-censored case. Extensions to higher derivatives are included. The results are applied to get optimal bandwidths with respect to almost sure uniform convergence.

1. Introduction. Let $X_1, X_2, \dots, X_n$ be independent and identically distributed random variables with common distribution function $F(x)$. Let $f(x) = F'(x)$ be the density function of X_1. A very popular estimator of $f(x)$ is the kernel estimator defined by

$$f_n(t) = \frac{1}{h_n} \int K\left(\frac{t-x}{h_n}\right) dF_n(x), \tag{1.1}$$

where F_n is the empirical distribution function of the sample $X_1, \dots, X_n$, $\{h_n\}$ is a sequence of bandwidths with $h_n \downarrow 0$ and $K(x)$ is an appropriate kernel function. Let

$$\bar{f}_n(t) = Ef_n(t) = \frac{1}{h_n} \int K\left(\frac{t-x}{h_n}\right) dF(x). \tag{1.2}$$

Stute (1982b) proved a law of the logarithm for kernel density estimator. For each $\varepsilon > 0$ and $I = (a, b)$ with $a < b$, put $I_\varepsilon = (a + \varepsilon, b - \varepsilon)$. Assume that K is of bounded variation with $K(x) = 0$ outside some finite interval $[r, s)$. Then if $f(x)$ is uniformly continuous on I with $0 < \delta \le f(x) \le M < \infty$ for all $x \in I$, Stute showed that [Theorem 1.3 in Stute (1982b)]

$$\lim_{n\to\infty} \sqrt{\frac{nh_n}{\log h_n^{-1}}} \sup_{t \in I_\varepsilon} \frac{|f_n(t) - \bar{f}_n(t)|}{\sqrt{f(t)}} = \left(2 \int_r^s K^2(x)\, dx\right)^{1/2}. \tag{1.3}$$

Stute's result gives the best uniform convergence rate of $f_n(t)$ to $\bar{f}_n(t)$ on I_ε and can be applied to get the optimal bandwidths with respect to almost sure uniform convergence of $f_n(t)$ to $f(t)$. For instance, if $f^{(2)}(t)$ is continuous on I, the

Received February 1992; revised February 1993.

[1]This manuscript was prepared using computer facilities supported in part by NSF Grants DMS 89-05292, DMS 87-03942, DMS 86-01732 and DMS-84-04941 and by the University of Chicago Block Fund.

AMS 1991 *subject classifications.* Primary 60F15; secondary 62G05, 62G30.

Key words and phrases. Quantile density function, random censorship, Kaplan–Meier estimator, kernel quantile density estimator, optimal bandwidths, strong Gaussian approximation, oscillation modulus.

corresponding optimal bandwidth given in Stute (1982b) is

$$(1.4)\qquad h_n \sim \left(\frac{\int K^2(u)du}{10\sup_{t\in I_\varepsilon}\left[|f^{(2)}(t)|^2/f(t)\right]\left(\int K(u)u^2du\right)^2}\frac{\log n}{n}\right)^{1/5},$$

and with this optimal bandwidth,

$$\frac{f_n(t)-f(t)}{\sqrt{f(t)}} = O\left(\left(\frac{\log n}{n}\right)^{2/5}\right)$$

uniformly on I_ε.

Let $Q(t) = \inf\{x: F(x) \geq t\}, 0 < t < 1$, be the quantile function of $F(x)$ and $q(t) = Q'(t)$ be the quantile density function. The quantile density function plays an important role in the statistical data modeling [see Parzen (1979)], reliability and medical studies. Parzen (1979) first introduced a kernel quantile density estimator. One version of the kernel quantile density estimator is

$$(1.5)\qquad \widehat{q}_n^*(t) = -\frac{1}{h_n^2}\int_0^1 F_n^{-1}(x)K'\left(\frac{x-t}{h_n}\right)dx,$$

where $F_n^{-1}(x) = \inf\{u: F_n(u) \geq x\}$. Falk (1986) established the asymptotic normality of $\widehat{q}_n^*(t)$ and obtained optimal bandwidths by minimizing the mean squared error. Sheather and Marron (1990) got a similar result. Yang (1985) introduced a new kernel quantile estimator defined by

$$(1.6)\qquad \widetilde{Q}_n(t) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{i/n-t}{h_n}\right)X_{(i)},$$

where $X_{(1)}, \dots, X_{(n)}$ are the order statistics of the sample $X_1, \dots, X_n$. Equation (1.6) suggests an alternative kernel quantile density estimator:

$$(1.7)\qquad \widehat{q}_n(t) = -\frac{1}{nh_n^2}\sum_{i=1}^{n} K'\left(\frac{i/n-t}{h_n}\right)X_{(i)}.$$

This estimator is easier to calculate than $\widehat{q}_n^*(t)$.

In this paper, we assume that the data come from a randomly right-censored model, that is, associated with each X_i, there is an independent censoring time Y_i and $Y_1, \dots, Y_n$ are assumed to be i.i.d. random variables with common distribution function $G(x)$. The distribution function $F(x)$ of X_i is called the survival time distribution. The observations in this model are the pairs (T_i, δ_i), where $T_i = \min(X_i, Y_i)$ and $\delta_i = I_{(X_i \leq Y_i)}$, $i = 1, 2, \dots, n$. Clearly, the T_i are i.i.d with common distribution function $H(x) = 1 - (1 - F(x))(1 - G(x))$, and the uncensored model is the special case of the censored model with $G = 0$. Based on such right-censored data, we want to estimate the quantile density function $q(t)$ by using kernel-type estimators constructed from the Kaplan–Meier estimator. The Kaplan–Meier estimator is defined by

$$\widehat{F}_n(t) = \begin{cases} 1 - \prod_{T_{(i)} \le t} \left(\dfrac{n-i}{n-i+1} \right)^{\delta_{(i)}}, & t < T_{(n)}, \\ 1, & t \ge T_{(n)}, \end{cases}$$

where $T_{(1)} \le T_{(2)} \le \cdots \le T_{(n)}$ are the order statistics of the T_i and $\delta_{(1)}, \dots, \delta_{(n)}$ are the corresponding δ_i. Let s_j denote the jump of $\widehat{F}_n(t)$ at $T_{(j)}$, that is,

$$s_j = \begin{cases} \widehat{F}_n\big(T_{(1)}\big), & j = 1, \\ \widehat{F}_n\big(T_{(j)}\big) - \widehat{F}_n\big(T_{(j-1)}\big), & j = 2, \dots, n, \end{cases}$$

Let $\widehat{F}_n^{-1}(x) = \inf\{u: \widehat{F}_n(u) \ge x\}$. Corresponding to the kernel quantile density estimators defined by (1.5) and (1.7), in the censored model, our estimators are

$$\widehat{q}_n^*(t) = -\frac{1}{h_n^2} \int_0^1 \widehat{F}_n^{-1}(x) K'\left(\frac{x-t}{h_n}\right) dx \tag{1.8}$$

and

$$\widehat{q}_n(t) = -\frac{1}{h_n^2} \sum_{i=1}^{n} T_{(i)} s_i K'\left(\frac{\widehat{F}_n\big(T_{(i)}\big) - t}{h_n}\right). \tag{1.9}$$

The estimator in (1.9) is motivated from Padgett's estimator of the quantile function [see Padgett (1986)]. Xiang (1992) established a Bahadur representation and a law of the iterated logarithm for the kernel quantile estimator and its derivatives for each fixed $t \in (0, F(T_H))$, where $T_H = \inf\{t: H(t) = 1\}$.

The main contribution of this article is to derive a law of the logarithm for $\widehat{q}_n^*(t)$ and $\widehat{q}_n(t)$ in Stute's sense when the data come from the random right-censorship model. These results are applied to get optimal bandwidths with respect to almost sure uniform convergence.

For the kernel $K(x)$ in this paper, we require that $K(x)$ is symmetric and, for a positive integer l,

$$K(x) \in C^l(-\infty, \infty), \quad K(x) \text{ has compact support } [-1, 1], \tag{1.10}$$

where

$$C^l(-\infty, \infty) = \{f: f^{(l)} \text{ is continuous on } (-\infty, \infty)\}$$

and, for some integer $m \ge 2$,

$$\begin{aligned} \int_{-1}^{1} K(x)\, dx &= 1, \\ \int_{-1}^{1} x^j K(x)\, dx &= 0, \qquad j = 1, \dots, m-1, \\ \int_{-1}^{1} x^m K(x)\, dx &= \alpha_m \ne 0. \end{aligned} \tag{1.11}$$

We note that (1.10) implies $K^{(i)}(x)$, $i = 1, \dots, l$, have compact support $[-1, 1]$. The kernel with properties (1.10) and (1.11) was investigated by Gasser and Müller (1984).

For the sequence of bandwidths $\{h_n\}$, we require that the following hold:

$$\text{(1.12)} \qquad \text{(i) } nh_n \uparrow \infty; \qquad \text{(ii) } \frac{\log h_n^{-1}}{nh_n} \to 0; \qquad \text{(iii) } \frac{\log h_n^{-1}}{\log\log n} \to \infty.$$

These conditions are necessary in Stute (1982a) to obtain local estimates for the empirical distribution function.

We use the notation $a_n \sim b_n$ if and only if $a_n/b_n \to 1$, as $n \to \infty$.

The present paper is organized in the following manner. The main results are given in Section 2. The estimation of higher derivatives of quantile function is given in Section 3.

2. Main results. Let $B(t)$ be a Brownian bridge and let $A(x)$ be a function defined on an interval $I \subset [0, 1]$ with $0 \le A(x) \le 1$ and a uniformly continuous derivative $a(x), a(x) > \delta > 0$ on I. We claim that the results in Stute (1982a) for $\alpha_n(t)$ and $\beta_n(t)$ also hold for $B(t)$ and $B(A(t))$, respectively. For example, from Shorack and Wellner [(1986), page 532], we have

$$\text{(2.1)} \qquad \lim_{n\to\infty} \sup_{\substack{\underline{c}h_n \le t-u \le \bar{c}h_n \\ t,u \in I}} \frac{|B(t) - B(u)|}{\sqrt{2(t-u)\log h_n^{-1}}} = 1 \quad \text{a.s.}$$

where $0 < \underline{c} \le \bar{c} < \infty$ are fixed numbers. Equation (2.1) is similar to Theorem 2.10 in Stute (1982a) and the analogue of Theorem 2.13 in Stute (1982a) is

$$\text{(2.2)} \qquad \lim_{n\to\infty} \sup_{\substack{\underline{c}h_n \le t-u \le \bar{c}h_n \\ t,u \in I}} \frac{|B(A(t)) - B(A(u))|}{\sqrt{2(t-u)a(x_{u,t})\log h_n^{-1}}} = 1 \quad \text{a.s.}$$

where $x_{u,t}$ is any point between u and t. Let

$$\text{(2.3)} \qquad L_n(t) = \frac{1}{h_n} \int K\left(\frac{t-x}{h_n}\right) dB(A(x)).$$

Let $J_\varepsilon = [c - \varepsilon, d + \varepsilon] \subset (0, 1)$ for some $\varepsilon > 0$ and $J = [c, d]$ with $c < d$. Thus, with a similar argument to Stute (1982b), we have the following lemma.

LEMMA 2.1. *Suppose that $a(x) = A'(x)$ is continuous on J_ε with $0 < \delta \le a(x) \le M < \infty$ for all $x \in J_\varepsilon$. Let $K(x)$ be any kernel function of bounded variation with $K(x) = 0$ outside $[-1, 1]$. Then, with probability* 1,

$$\text{(2.4)} \qquad \lim_{n\to\infty} \sqrt{\frac{h_n}{\log h_n^{-1}}} \sup_{t \in J} \frac{|L_n(t)|}{\sqrt{a(t)}} = \left(2\int_{-1}^{1} K^2(x)\,dx\right)^{1/2}.$$

Now assume $J_\varepsilon \subset (0, F(T_H))$. Let

$$\bar{q}_n(t) = -\frac{1}{h_n^2}\int_0^1 Q(x)K'\left(\frac{x-t}{h_n}\right)dx. \tag{2.5}$$

Our main result is the following.

THEOREM 2.2. *Assume that $q(t)$ is continuous on J_ε with $0 < \delta \le q(t) \le M < \infty$ for all $t \in J_\varepsilon$. Let* (1.10) *hold for $l = 1$, let $K'(x)$ be Lipschitz of order* 1 *and let $G(x)$ be Lipschitz of order $\frac{1}{2}$ on $[Q(c-\varepsilon), Q(d+\varepsilon)]$. Then if*

$$\frac{\log_2 n}{\left(nh_n^3 \log h_n^{-1}\right)^{1/2}} \to 0, \tag{2.6}$$

$$\lim_{n\to\infty}\sqrt{\frac{nh_n}{\log h_n^{-1}}}\sup_{t\in J}\frac{\sqrt{1-G(Q(t))}\,|\hat{q}_n^*(t)-\bar{q}_n(t)|}{q(t)} = \left(2\int_{-1}^1 K^2(x)\,dx\right)^{1/2} \tag{2.7}$$

and

$$\lim_{n\to\infty}\sqrt{\frac{nh_n}{\log h_n^{-1}}}\sup_{t\in J}\frac{\sqrt{1-G(Q(t))}\,|\hat{q}_n(t)-\bar{q}_n(t)|}{q(t)} = \left(2\int_{-1}^1 K^2(x)\,dx\right)^{1/2}. \tag{2.8}$$

Our approach is based on a strong embedding result of Major and Rejtö (1988). Let

$$H^u(t) = P(T_1 \le t, \delta_1 = 1) \quad \text{and} \quad H^c(t) = P(T_1 \le t, \delta_1 = 0).$$

Major and Rejtö (1988) showed that, for $t < T_H$,

$$\hat{F}_n(t) - F(t) = \frac{1}{n}\sum_{i=1}^n \psi_i(t) + \tau_n(t) \tag{2.9}$$

and

$$\hat{F}_n(t) - F(t) = \frac{1}{\sqrt{n}}W(t) + \gamma_n(t), \tag{2.10}$$

where

$$\begin{aligned}\psi_i(t) = (1-F(t))\Bigg\{&\int_{-\infty}^t \frac{I_{(T_i\le y)} - H(y)}{(1-H(y))^2}dH^u(y) + \frac{I_{(T_i\le t,\delta_i=1)} - H^u(t)}{1-H(t)}\\ &-\int_{-\infty}^t \frac{I_{(T_i\le y,\delta_i=1)} - H^u(y)}{(1-H(y))^2}dH(y)\Bigg\}\end{aligned} \tag{2.11}$$

and $W(t)$ is a Gaussian process defined by

$$W(t) = (1 - F(t))\left\{ \int_{-\infty}^{t} \frac{B(H^u(y)) - B(1 - H^c(y))}{(1 - H(y))^2} dH^u(y) + \frac{B(H^u(t))}{1 - H(t)} - \int_{-\infty}^{t} \frac{B(H^u(y))}{(1 - H(y))^2} dH(y) \right\}. \tag{2.12}$$

The remainder terms in (2.9) and (2.10) satisfy, with probability 1,

$$\sup_{t \le T} |\tau_n(t)| = O\left(\frac{\log n}{n}\right), \qquad T < T_H \tag{2.13}$$

and

$$P\left(\sup_{t \le T} |n\gamma_n(t)| > \frac{2C}{\Delta}\log n + x\right) < 2ke^{-\lambda\Delta^2 x}, \tag{2.14}$$

for all $x > 0$, where $0 < \Delta < 1 - H(T)$ and C, k and λ are some positive universal constants.

Proof of Theorem 2.2. We first prove (2.7). Write $G_n(x,t) = \int_0^x K'((s - t)/h_n)ds$. Then from the change of variables theorem [Billingsley (1986), page 219],

$$\begin{aligned} \widehat{q}_n^*(t) - \overline{q}_n(t) &= -\frac{1}{h_n^2}\int_{-\infty}^{\infty} x d\left(G_n(\widehat{F}_n(x), t) - G_n(F(x), t)\right) \\ &= \frac{1}{h_n^2}\int_{-\infty}^{\infty} \left(G_n(\widehat{F}_n(x), t) - G_n(F(x), t)\right) dx \\ &= \frac{1}{h_n}\int_{-\infty}^{\infty} \left(\int_{[F(x) - t]/h_n}^{[\widehat{F}_n(x) - t]/h_n} K'(u)\, du\right) dx \\ &= I_{1n}(t) + I_{2n}(t), \end{aligned}$$

where

$$I_{1n}(t) = \frac{1}{h_n}\int_{-\infty}^{\infty} K'\left(\frac{F(x) - t}{h_n}\right)\left(\frac{\widehat{F}_n(x) - F(x)}{h_n}\right) dx \tag{2.15}$$

and

$$I_{2n}(t) = \frac{1}{h_n}\int_{-\infty}^{\infty} \left(\int_{[F(x) - t]/h_n}^{[\widehat{F}_n(x) - t]/h_n} \left(K'(u) - K'\left(\frac{F(x) - t}{h_n}\right)\right) du\right) dx. \tag{2.16}$$

From (2.10),

$$I_{1n}(t) = r_{1n}(t) + r_{2n}(t)$$

with

$$r_{1n}(t) = \frac{1}{\sqrt{n}h_n^2}\int_{-\infty}^{\infty} K'\left(\frac{F(x)-t}{h_n}\right)W(x)\,dx$$

and

$$r_{2n}(t) = \frac{1}{h_n^2}\int_{-\infty}^{\infty} K'\left(\frac{F(x)-t}{h_n}\right)\gamma_n(x)\,dx.$$

To make use of Lemma 2.1, we note that

$$\begin{aligned} r_{1n} &= \frac{1}{\sqrt{n}h_n^2}\int_{t-h_n}^{t+h_n} Q'(v)K'\left(\frac{v-t}{h_n}\right)W(Q(v))\,dv \\ &\sim -\frac{1}{\sqrt{n}h_n}Q'(t)\int_{t-h_n}^{t+h_n} K\left(\frac{v-t}{h_n}\right)dW(Q(v)). \end{aligned} \tag{2.17}$$

Let

$$\begin{aligned} W_1(Q(t)) &= \frac{B\big(H^u(Q(t))\big)}{1-G(Q(t))}, \\ W_2(Q(t)) &= (1-t)\int_{-\infty}^{Q(t)} \frac{B(H^u(y)) - B(1-H^c(y))}{(1-H(y))^2}\,dH^u(y), \\ W_3(Q(t)) &= -(1-t)\int_{-\infty}^{Q(t)} \frac{B(H^u(y))}{(1-H(y))^2}\,dH(y). \end{aligned}$$

Then, from (2.12),

$$W(Q(t)) = \sum_{i=1}^{3} W_i(Q(t)).$$

Let

$$\omega_i(h) = \sup_{|u-t|\le h,u,t\in J} \left|W_i(Q(t)) - W_i(Q(u))\right|, \qquad i = 1,2,3,$$

be the oscillation modulus of $W_i(Q(t))$. Thus as $h \downarrow 0$, Lévy's theorem [cf. Shorack and Wellner (1986), page 534] and the smoothness conditions imposed on $G(t)$ and $F(t)$ imply that, with probability 1,

$$\omega_1(h) = O\big(h^{1/2}(\log h^{-1})^{1/2}\big) \quad \text{and} \quad \omega_i(h) = O(h), \qquad i = 2,3.$$

Hence, it follows that

$$\begin{aligned} r_{1n}(t) &\sim -\frac{1}{\sqrt{n}h_n}Q'(t)\int_{t-h_n}^{t+h_n} K\left(\frac{v-t}{h_n}\right)dW_1(Q(v)) \\ &\sim -\frac{1}{\sqrt{n}h_n}\frac{Q'(t)}{1-G(Q(t))}\int_{t-h_n}^{t+h_n} K\left(\frac{v-t}{h_n}\right)dB(A(v)), \end{aligned}$$

with $A(t) = H^u(Q(t))$. By

$$H^u(x) = \int_{-\infty}^{x} \big(1 - G(y)\big)\, dF(y),$$

we obtain

$$a(t) = A'(t) = 1 - G\big(Q(t)\big).$$

Hence, by using Lemma 2.1, with probability 1,

$$\lim_{n\to\infty} \sqrt{\frac{nh_n}{\log h_n^{-1}}} \sup_{t\in J} \frac{\sqrt{1-G\big(Q(t)\big)}\,|r_{1n}(t)|}{q(t)} = \left(2\int_{-1}^{1} K^2(x)\,dx\right)^{1/2}. \tag{2.18}$$

To complete the proof of (2.7), it remains to show

$$\sup_{t\in J} \big|r_{2n}(t)\big| = o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right) \tag{2.19}$$

and

$$\sup_{t\in J} \big|I_{2n}(t)\big| = o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right). \tag{2.20}$$

For small enough h_n, we have $Q(t + h_n u) \in (-\infty, T_0]$ for some $T_0 < T_H$ and $t + h_n u \in J_\varepsilon$ for all $t \in J$ and $u \in [-1, 1]$. Let $a_n = ([\log h_n^{-1}]/nh_n)^{1/2}$. Then, for any $\eta > 0$, there exist positive constants C_0, C_1, C_2, such that

$$\begin{aligned} P\Big(\sup_{t\in J} \big|a_n^{-1} r_{2n}(t)\big| > \eta\Big) &\le P\Big(\sup_{t\le T_0} \big|\gamma_n(t)\big| > \frac{C_0 a_n h_n \eta}{\int_{-1}^{1} |K'(x)|dx}\Big) \\ &\le 2k \exp\big\{-\lambda\Delta^2(C_1 a_n h_n n - C_2 \log n)\big\}. \end{aligned} \tag{2.21}$$

Hence, (2.19) follows easily from (2.6) and the Borel–Cantelli lemma.

To prove (2.20), for an $\varepsilon > 0$ with $1+\varepsilon \le (1-t)/h_n$ (this holds if n is large), write

$$\begin{aligned} I_{2n}(t) &= \int_{-1-\varepsilon}^{1+\varepsilon} Q'(t + h_n x) \int_x^{[\hat{F}_n(Q(t+xh_n)) - t]/h_n} \big(K'(u) - K'(x)\big)\, du\, dx \\ &\quad + \int_{1+\varepsilon}^{(1-t)/h_n} Q'(t + h_n x) \int_x^{[\hat{F}_n(Q(t+xh_n)) - t]/h_n} \big(K'(u) - K'(x)\big)\, du\, dx \\ &\quad + \int_{-t/h_n}^{-1-\varepsilon} Q'(t + h_n x) \int_x^{[\hat{F}_n(Q(t+xh_n)) - t]/h_n} \big(K'(u) - K'(x)\big)\, du\, dx \\ &= S_{1n}(t) + S_{2n}(t) + S_{3n}(t). \end{aligned}$$

We have

$$\begin{aligned}
|S_{2n}(t)| &\le \int_{1+\varepsilon}^{(1-t)/h_n} |Q'(t+h_n x)| I\left(\frac{\widehat{F}_n(Q(t+xh_n)) - t}{h_n} < 1\right) \\
&\quad \times \left| \int_x^{[\widehat{F}_n(Q(t+xh_n))-t]/h_n} (K'(u) - K'(x))\, du \right| dx \\
&\le \int_{1+\varepsilon}^{(1-t)/h_n} |Q'(t+h_n x)| \\
&\quad \times I\left(\frac{\widehat{F}_n\big(Q(t+(1+\varepsilon)h_n)\big) - \big(t+(1+\varepsilon)h_n\big)}{h_n} < -\varepsilon\right) \\
&\quad \times \left| \int_x^{[\widehat{F}_n(Q(t+xh_n))-t]/h_n} (K'(u) - K'(x))\, du \right| dx.
\end{aligned}$$

Hence, if h_n tends to zero slower than $([\log_2 n]/n)^{1/2}$, Corollary 1 of Földes and Rejtö (1981) implies, with probability 1, $\sup_{t\in J} |S_{2n}(t)| = 0$ for large n. Similarly, $\sup_{t\in J} |S_{3n}(t)| = 0$ for large n. To estimate $S_{1n}(t)$, it follows again from Corollary 1 of Földes and Rejtö (1981) and (2.6) that

$$\begin{aligned}
\sup_{t\in J} |S_{1n}(t)| &= O\left(\sup_{t\in J, |u|\le 1+\varepsilon} \left|\frac{\widehat{F}_n(Q(t+h_n u)) - t}{h_n} - u\right|^2\right) \\
&= o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right).
\end{aligned}$$

Hence (2.20) follows.

To prove (2.8), we write

$$\widehat{q}_n(t) = -\frac{1}{h_n^2}\int_0^1 \widehat{F}_n^{-1}(x) K_n\left(\frac{x-t}{h_n}\right) dx$$

and introduce

$$\widetilde{q}_n(t) = -\frac{1}{h_n^2}\int_0^1 Q(x) K_n\left(\frac{x-t}{h_n}\right) dx,$$

where K_n is defined by

$$K_n\left(\frac{x-t}{h_n}\right) = K'\left(\frac{i/n-t}{h_n}\right), \qquad \frac{i-1}{n} < x \le \frac{i}{n}, \qquad i = 0, \pm 1, \pm 2, \dots .$$

It is easy to check that

$$K_n(u) = 0, \quad \text{if } |u| \ge 1 + \frac{1}{nh_n}$$

and

$$\sup_{-\infty<u<\infty} |K_n(u) - K'(u)| = O\left(\frac{1}{nh_n}\right).$$

Thus

$$\widehat{q}_n(t) - \overline{q}_n(t) = \widehat{q}_n^*(t) - \overline{q}_n(t) + \widetilde{q}_n(t) - \overline{q}_n(t) + \frac{1}{h_n} \int_{-\infty}^{\infty} \int_{[F(x)-t]/h_n}^{[\widehat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du\, dx.$$

To complete the proof, it suffices to show

$$\sup_{t \in J} \left|\widetilde{q}_n(t) - \overline{q}_n(t)\right| = o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right) \tag{2.22}$$

and

$$\sup_{t \in J} \frac{1}{h_n} \int_{-\infty}^{\infty} \left| \int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du \right| dx = o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right). \tag{2.23}$$

Equation (2.22) follows easily from

$$\begin{aligned} &\sup_{t \in J} \left|\widetilde{q}_n(t) - \overline{q}_n(t)\right| \\ &\quad \le \sup_{t\in J} \frac{1}{h_n} \int_{|u| \le 1 + 1/nh_n} \left|Q(t + uh_n)\right| \left|K_n(u) - K'(u)\right| du \\ &\quad = O\left(\frac{1}{nh_n^2}\right) = o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right). \end{aligned} \tag{2.24}$$

To prove (2.23), we have for a given $\varepsilon > 1/nh_n$,

$$\begin{aligned} &\int_{Q(t+h_n+2h_n\varepsilon)}^{\infty} \left| \int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du \right| dx \\ &\quad \le \int_{Q(t+h_n+2h_n\varepsilon)}^{\infty} I\left(\frac{\widehat{F}_n(x) - t}{h_n} < 1 + \varepsilon\right) \\ &\qquad \times \left| \int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du \right| dx \\ &\quad \le \int_{Q(t+h_n+2h_n\varepsilon)}^{\infty} I\left(\frac{\widehat{F}_n\big(Q(t + h_n + 2h_n\varepsilon)\big) - t}{h_n} < 1 + \varepsilon\right) \\ &\qquad \times \left| \int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du \right| dx \\ &\quad \le \int_{Q(t+h_n+2h_n\varepsilon)}^{\infty} I\left(\frac{\widehat{F}_n\big(Q(t+h_n+2h_n\varepsilon)\big) - (t + h_n + 2h_n\varepsilon)}{h_n} < -\varepsilon\right) \\ &\qquad \times \left| \int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n} \big(K_n(u) - K'(u)\big)\, du \right| dx. \end{aligned}$$

Hence, with probability 1,

$$\sup_{t\in J}\int_{Q(t+h_n+2h_n\varepsilon)}^{\infty}\left|\int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n}\big(K_n(u)-K'(u)\big)\,du\right|dx=0\quad\text{for } n \text{ large.}$$

Similarly,

$$\sup_{t\in J}\int_{-\infty}^{Q(t-h_n-2h_n\varepsilon)}\left|\int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n}\big(K_n(u)-K'(u)\big)\,du\right|dx=0\quad\text{for } n \text{ large.}$$

These together with (2.6) imply

$$\begin{aligned}(2.25)\qquad &\sup_{t\in J}\frac{1}{h_n}\int_{-\infty}^{\infty}\left|\int_{[F(x)-t]/h_n}^{[\hat{F}_n(x)-t]/h_n}\big(K_n(u)-K'(u)\big)\,du\right|dx\\ &=O\left(\frac{1}{nh_n^2}\sup_{t\in J}\sup_{Q(t-h_n-2h_n\varepsilon)\le x\le Q(t+h_n+2h_n\varepsilon)}\left|\widehat{F}_n(x)-F(x)\right|\right)\\ &=o\left(\left(\frac{\log h_n^{-1}}{nh_n}\right)^{1/2}\right).\end{aligned}$$

The proof is complete. □

REMARK. If $Q(t)$ is twice differentiable in J_ε, Major and Rejtö (1988) and Lo and Singh (1986) imply

$$\widehat{F}_n^{-1}(t)-Q(t)=-\frac{1}{\sqrt{n}}W\big(Q(t)\big)+B_n(t), \tag{2.26}$$

where $W(t)$ is defined by (2.10) and

$$\sup_{t\in J}\left|\beta_n(t)\right|=O\left(\left(\frac{\log n}{n}\right)^{3/4}\right). \tag{2.27}$$

Based on this representation, we can prove (2.7) and (2.8) under weaker conditions on kernel function $K(x)$ and bandwidth h_n. We give this result without proof.

THEOREM 2.3. *Assume that $Q(t)$ is twice differentiable in J_ε with $0<\delta\le q(t)\le M<\infty$ for all $t\in J_\varepsilon$. Let (1.10) hold for $l=1$, and let $G(x)$ be Lipschitz of order $\frac{1}{2}$ on $[Q(c-\varepsilon),Q(d+\varepsilon)]$. Then if*

$$\frac{(\log n)^3}{nh_n^2(\log h_n^{-1})^2}\to 0\quad \text{as } n\to\infty, \tag{2.28}$$

$$\lim_{n\to\infty}\sqrt{\frac{nh_n}{\log h_n^{-1}}}\sup_{t\in J}\frac{\sqrt{1-G\big(Q(t)\big)}\left|\widehat{q}_n^{*}(t)-\overline{q}_n(t)\right|}{q(t)}=\left(2\int_{-1}^{1}K^2(x)dx\right)^{1/2} \tag{2.29}$$

and

$$\text{(2.30)}\quad \lim_{n\to\infty}\sqrt{\frac{nh_n}{\log h_n^{-1}}}\sup_{t\in J}\frac{\sqrt{1-G(Q(t))}\,|\widehat{q}_n(t)-\overline{q}_n(t)|}{q(t)}=\left(2\int_{-1}^{1}K^2(x)dx\right)^{1/2}.$$

To apply Theorem 2.2 to get optimal bandwidths, we further assume that $q^{(m)}(t)$ is continuous in J_ε, with $m \geq 2$, and that $K(x)$ satisfies (1.11). From Theorem 2.2 and

$$\overline{q}_n(t)-q(t)=\frac{h_n^m}{m!}q^{(m)}(t)\alpha_m+o(h_n^m),$$

the optimal bandwidth is obtained by minimizing the term

$$\text{(2.31)}\quad \begin{aligned}&\frac{h_n^m}{m!}\sup_{t\in J}\frac{\sqrt{1-G(Q(t))}\,|q^{(m)}(t)|}{q(t)}\int_{-1}^{1}|K(u)u^m|du\\&+\left(\frac{2\log h_n^{-1}}{nh_n}\int_{-1}^{1}K^2(u)du\right)^{1/2}.\end{aligned}$$

For $m = 2$, the asymptotically optimal bandwidth is

$$\text{(2.32)}\quad \begin{aligned}&h_n\\&\sim\left(\frac{\int_{-1}^{1}K^2(u)\,du}{10\sup_{t\in J}\left[|q^{(2)}(t)|^2(1-G(Q(t)))/q^2(t)\right]\left(\int_{-1}^{1}K(u)u^2du\right)^2}\frac{\log n}{n}\right)^{1/5}\end{aligned}$$

and, with probability 1,

$$\frac{\widehat{q}_n^*(t)-q(t)}{q(t)}=O\left(\left(\frac{\log n}{n}\right)^{2/5}\right)$$

and

$$\frac{\widehat{q}_n(t)-q(t)}{q(t)}=O\left(\left(\frac{\log n}{n}\right)^{2/5}\right),$$

uniformly on J.

3. Estimation of higher derivatives of the quantile function. Assume that (1.10) holds for $l = r > 1$ and $K^{(r)}$ is Lipschitz of order 1. Define estimators

$$\text{(3.1)}\quad \widehat{q}_n^{*(r)}(t)=\frac{(-1)^r}{h_n^{r+1}}\int_0^1\widehat{F}_n^{-1}(x)K^{(r)}\left(\frac{x-t}{h_n}\right)dx$$

and

$$\widehat{q}_n^{(r)}(t) = \frac{(-1)^r}{h_n^{r+1}} \sum_{i=1}^{n} T_{(i)} s_i K^{(r)}\left(\frac{\widehat{F}_n(T_{(i)}) - t}{h_n}\right). \tag{3.2}$$

Let

$$\overline{q}_n^{(r)}(t) = \frac{(-1)^r}{h_n^{r+1}} \int_0^1 Q(t) K^{(r)}\left(\frac{x-t}{h_n}\right) dx. \tag{3.3}$$

If $q(t) = Q'(t)$ and $G(x)$ satisfy the assumptions of Theorem 2.2, we get

$$\begin{aligned} &\lim_{n\to\infty} h_n^{r-1} \sqrt{\frac{nh_n}{\log h_n^{-1}}} \sup_{t\in J} \frac{\sqrt{1-G(Q(t))}\,|\widehat{q}_n^{*(r)}(t) - \overline{q}_n^{(r)}(t)|}{q(t)} \\ &\quad = \left(2\int_{-1}^{1} [K^{(r-1)}(x)]^2 dx\right)^{1/2} \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} &\lim_{n\to\infty} h_n^{r-1} \sqrt{\frac{nh_n}{\log h_n^{-1}}} \sup_{t\in J} \frac{\sqrt{1-G(Q(t))}\,|\widehat{q}_n^{(r)}(t) - \overline{q}_n^{(r)}(t)|}{q(t)} \\ &\quad = \left(2\int_{-1}^{1} [K^{(r-1)}(x)]^2 dx\right)^{1/2}. \end{aligned} \tag{3.5}$$

Furthermore, if, for some $m \geq 2$, $Q^{(r+m)}(t)$ is continuous on J_ε for $J_\varepsilon \subset (0, F(T_H))$, the optimal h_n is of order $((\log n)/n)^{1/[2(r+m)+1]}$ and, with probability 1,

$$\frac{\widehat{q}_n^{*(r)}(t) - q^{(r)}(t)}{q(t)} = O\left(\left(\frac{\log n}{n}\right)^{m/[2(r+m)+1]}\right)$$

and

$$\frac{\widehat{q}_n^{(r)}(t) - q^{(r)}(t)}{q(t)} = O\left(\left(\frac{\log n}{n}\right)^{m/[2(r+m)+1]}\right)$$

uniformly on J.

Acknowledgments. I am very grateful to Professor S. M. Stigler for introducing me to this subject and for his encouragement and advice, and to Professor R. R. Bahadur for helpful discussions. I would also like to thank the referee and an Associate Editor for their valuable comments, which led to this greatly improved version of the manuscript.

REFERENCES

BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. Wiley, New York.

FALK, M. (1986). On the estimation of the quantile density function. *Statist. Probab. Lett.* **4** 69–73.

FÖLDES, A. and REJTÖ, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. Verw. Gebiete* **56** 75–86.

GASSER, TH. and MÜLLER, H. G. (1984). Estimating regression functions and their derivatives by the kernel method. *Scand. J. Statist.* **11** 171–185.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

LO, S. H. and SINGH, K. (1986). The product-limit estimator and the bootstrap: Some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.

MAJOR, P. and REJTÖ, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *Ann. Statist.* **16** 1113–1132.

PADGETT, W. J. (1986). A kernel-type estimator of a quantile function from right-censored data. *J. Amer. Statist. Assoc.* **81** 215–222.

PARZEN, E. (1979). Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.* **74** 105–131.

SHEATHER, S. J. and MARRON, J. S. (1990). Kernel quantile estimators. *J. Amer. Statist. Assoc.* **85** 410–416.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

STUTE, W. (1982a). The oscillation behavior of empirical processes. *Ann. Probab.* **10** 86–107.

STUTE, W. (1982b). A law of the logarithm for kernel density estimators. *Ann. Probab.* **10** 414–422.

XIANG, X. (1992). Strong Gaussian approximation of kernel estimators of a quantile function and its derivatives from randomly right-censored data. Technical Report 335, Dept. Statistics, Univ. Chicago.

YANG, S. S. (1985). A smooth nonparametric estimator of a quantile function. *J. Amer. Statist. Assoc.* **80** 1004–1011.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OREGON
EUGENE, OREGON 97403