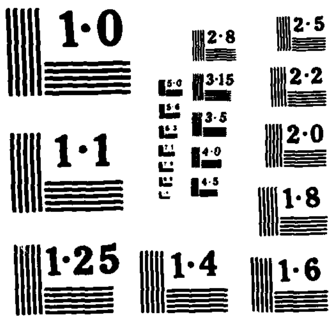
AD-A162 837 SOME CONVERGENCE RESULTS FOR KERNEL-TYPE QUANTILE ESTIMATORS UNDER CENSOR..(U) SOUTH CAROLINA UNIV COLUMBIA DEPT OF MATHEMATICS AND STATISTI.. 1/1

UNCLASSIFIED Y L LIO ET AL. NOV 85 TR-108 F/G 12/1 NL

END
FILMED
DTIC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 22

AD-A162 837

SOME CONVERGENCE RESULTS FOR KERNEL-TYPE
QUANTILE ESTIMATORS UNDER CENSORING*

by

Y. L. Lio and W. J. Padgett

Department of Statistics
University of South Carolina
Columbia, SC 29208

University of South Carolina
Statistics Technical Report No. 108
62G05-14

November, 1985

* Research supported by the U.S. Air Force Office of Scientific Research under grant number AFOSR 84-0156 and the U.S. Army Research Office under grant number MIPR ARO 139-85.

ABSTRACT

Based on right-censored data from a lifetime distribution F_o, a kernel-type estimator of ~~the~~ a certain quantile function $Q^o(p)$, defined by $Q_n(p) = h_n^{-1}\int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$, is studied. This estimator is smoother than ~~the~~ a product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$, where $\hat{F}_n$ denotes the product-limit estimator of F_o from the censored sample. Under the random censorship model and general conditions on h_n, K, and F_o, asymptotic normality of $Q_n(p)$ and a simpler approximation to it, $Q_n^*(p)$, is shown, and mean square convergence of Q_n is proven. Also, the asymptotic mean equivalence of Q_n and Q_n^* is shown.

This document also covers: asymptotic normality; asymptotic mean equivalence; and mean square convergence. Keywords: probability distribution; kernel functions.

1. INTRODUCTION AND PRELIMINARIES

For any probability distribution G, denote the quantile function by $Q(p) \equiv G^{-1}(p) = \inf\{x: G(x) \geq p\}$, $0 \leq p \leq 1$. From a random (uncensored) sample from G, the sample quantile function $G_n^{-1}(p) = \inf\{x: G_n(x) \geq p\}$, $0 \leq p \leq 1$, has been used to estimate $Q(p)$, where G_n denotes the sample distribution function. Csörgő (1983) gave many of the known results concerning $G_n^{-1}(p)$. Also, Falk (1984) studied the relative deficiency of the sample quantile with respect to kernel-type estimators, and Falk (1985) obtained asymptotic normality for kernel estimators. Yang (1985) obtained some convergence properties of kernel estimators of $Q(p)$ and gave simulation results comparing kernel-type estimators with other estimators. For arbitrarily right-censored data, Sander (1975) proposed estimation of $Q(p)$ by the quantile function of the product-limit estimator, and she and Cheng (1984) derived asymptotic properties while Csörgő (1983) presented strong approximation results for that estimator.

For randomly right-censored data, Padgett (1985) proposed a smooth non-parametric estimator of the quantile function, defined by $Q_n(p) = h_n^{-1}\int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$, where $\hat{Q}_n$ denotes the product-limit quantile function and K is an appropriate kernel function. An approximation to $Q_n(p)$, denoted by $Q_n^*(p)$, which is somewhat easier to compute was also studied. The estimator Q_n, mentioned briefly by Parzen (1979), was shown to be strongly consistent, and Q_n and Q_n^* were shown to be almost surely asymptotically equivalent. Lio, Padgett, and Yu (1985) obtained an asymptotic normality result for Q_n and showed that Q_n and Q_n^* are asymptotically uniformly mean square equivalent under certain conditions.

In this paper, some further asymptotic normality results for Q_n and Q_n^* will be given. Also, their asymptotic mean equivalence and the mean square convergence of Q_n will be shown. To define these estimators, let $X_1^o,\ldots,X_n^o$ denote the true survival times of n items or individuals that are censored on the right by a sequence $U_1,U_2,\ldots,U_n$, which in general may be either constants or random variables. It is assumed that the X_i^o's are nonnegative independent identically distributed random variables with common unknown distribution function F_o and unknown quantile function $Q^o(p) = \xi_p^o$. The observed right-censored data are denoted by the pairs (X_i,Δ_i), $i=1,\ldots,n$, where

$$X_i = \min\{X_i^o,U_i\}, \qquad \Delta_i = \begin{cases} 1 \text{ if } X_i^o \le U_i \\ 0 \text{ if } X_i^o > U_i . \end{cases}$$

Let (Z_i,Λ_i), $i=1,\ldots,n$, denote the ordered X_i's along with their corresponding Δ_i's. A popular estimator of the survival function $S_o = 1-F_o$ is the product-limit estimator of Kaplan and Meier (1958), shown to be "self-consistent" by Efron (1967) and defined by

$$\hat{P}_n(t) = \begin{cases} 1, & 0 \le t \le Z_1, \\ \prod_{i=1}^{k-1} \left(\frac{n-i}{n-i+1}\right)^{\Lambda_i}, & Z_{k-1} < t \le Z_k, \quad k=2,\ldots,n \\ 0, & t > Z_n . \end{cases}$$

Denote the product-limit estimator of $F_o(t)$ by $\hat{F}_n(t) = 1 - \hat{P}_n(t)$, and let s_j denote the jump of $\hat{P}_n$ at Z_j, that is,

$$s_j = \begin{cases} 1 - \hat{P}_n(Z_2), & j = 1 \\ \hat{P}_n(Z_j) - \hat{P}_n(Z_{j+1}), & j = 2,\ldots,n-1 \\ \hat{P}_n(Z_n), & j = n. \end{cases}$$

Note that $s_j=0$ if and only if $\Lambda_j = 0$, $j < n$, i.e. whenever Z_j is a censored observation. Also, denote $S_i = \hat{F}_n(Z_{i+1}) = \sum_{j=1}^{i} s_j$, $i=1,\ldots,n$, with $S_0 = 0$, $Z_0 = 0$, and $Z_{n+1} = Z_n + \varepsilon$, for some positive constant ε.

It is natural to estimate ξ_p^o by the product-limit quantile function $\hat{Q}_n(p) \equiv \hat{\xi}_p^o = \inf\{t: \hat{F}_n(t) \geq p\}$. Then the kernel-type estimator $Q_n(p)$ studied by Padgett (1985) is written as

$$Q_n(p) = h_n^{-1} \int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$$

$$= h_n^{-1} \sum_{i=1}^{n} Z_i \int_{S_{i-1}}^{S_i} K((t-p)/h_n)dt, \tag{1.1}$$

for kernel function K and bandwidth sequence $\{h_n\}$. Also, the simpler kernel-type estimator Q_n^* which is an approximation to (1.1) is defined by

$$Q_n^*(p) = h_n^{-1} \sum_{i=1}^{n} Z_i s_i \, K((S_i - p)/h_n). \tag{1.2}$$

For the results here, the random right-censorship model will be assumed; that is, $U_1,\ldots,U_n$ constitute a random sample from a distribution H (usually unknown) and are independent of $X_1^o,\ldots,X_n^o$. The distribution function of each X_i, $i=1,\ldots,n$, is then $F = 1 - (1-F_o)(1-H)$. In addition, some or all of the following conditions will be assumed for the kernel function, bandwidth sequence, and lifetime and censoring distributions:

(h.1) $h_n \to 0$ as $n \to \infty$;

(K.1) $K(x)$ is a bounded probability density function which has finite support, i.e. $K(x) = 0$ for $|x| > c$ for some $c > 0$;

(K.2) K is symmetric about zero;

(K.3) K satisfies a Lipschitz condition, i.e. there exists a constant Γ such that for all x,y,

$$|K(x) - K(y)| \leq \Gamma|x - y|;$$

(F.1) F_o is continuous with density function f_o.

These conditions are not prohibitive and the conditions on F_o are similar to conditions required by Cheng (1984).

2. ASYMPTOTIC NORMALITY

Lio, Padgett, and Yu (1985) showed that under conditions (h.1), (K.1)-(K.2), and (F.1), if the derivative f_o' is continuous at ξ_p^o, $f_o(\xi_p^o) > 0$, and $n^{\frac{1}{4}}h_n \to 0$, then for $0 < p < T$, where $T < \min\{1, T_{H(F_o^{-1})}\}$ with $T_{H(F_o^{-1})} = \sup\{t: H(F_o^{-1}(t))<1\}$, $n^{\frac{1}{2}}[Q_n(p) - Q^o(p)] \to Z$ in distribution as $n \to \infty$, where Z is a normally distributed random variable with mean zero and variance $\sigma_p^2 = (1-p)^2 \int_0^{\xi_p^o}[1-F(u)]^{-2}dF_o^*(u)/f_o^2(\xi_p^0)$. Here $F_o^*(u) = P(X_i \leq u, \Delta_i = 1)$ is the subdistribution function of the uncensored observations. The condition $n^{\frac{1}{4}}h_n \to 0$ can be replaced by $n^{\frac{1}{2}}h_n \to 0$ by using a slightly different proof than that of Lio, Padgett, and Yu (1985).

Define $Q^o(p,h_n) = h_n^{-1}\int_0^1 Q^o(t)K((t-p)/h_n)dt$ for $0 \leq p \leq 1$. An asymptotic normality result for (1.1) can be obtained without the condition on the rate of convergence of h_n to zero by centering with $Q^o(p,h_n)$ instead of $Q^o(p)$. This type of centering seems to be required for asymptotic normality of $Q_n^*(p)$.

THEOREM 2.1. Suppose (h.1), (K.1), (K.2), and (F.1) hold and f_o' is continuous at ξ_p^o with $f_o(\xi_p^o) > 0$. Then for $0 < p < T$, $n^{\frac{1}{2}}[Q_n(p) - Q^o(p,h_n)] \to Z$ in distribution as $n \to \infty$, where Z is normally distributed with mean zero and variance σ_p^2.

The proof of Theorem 2.1 follows from the following lemma proven by Lio, Padgett, and Yu (1985) and from Corollary 1 of Cheng (1984).

LEMMA 2.1. Under the conditions of Theorem 2.1

$$\left|\int_0^1[q_n(t) - q_n(p)]h_n^{-1}K((t-p)/h_n)dt\right| \to 0 \text{ in probability}$$

as $n \to \infty$, where $q_n(t) = n^{\frac{1}{2}}[\hat{Q}_n(t) - Q^o(t)]$ denotes the product-limit quantile process.

The asymptotic normality of (1.2) follows from Theorem 2.1 and the next lemma.

LEMMA 2.2. In addition to the conditions of Theorem 2.1, suppose (K.3) holds and $E(X^{o2}) < \infty$. Then as $n \to \infty$,

$$\sup_{0 \le p \le T} n[Q_n^*(p) - Q_n(p)]^2 \to 0 \text{ in probability,}$$

provided $h_n^{-4}\sqrt{\log\log n/n} \to 0$.

PROOF. For $0 \le p \le T$, when $s_i > 0$, i.e. when Z_i is uncensored, let S_i^* be an interior point of the interval (S_{i-1}, S_i) with probability one so that

$$s_i K\left(\frac{S_i^*-p}{h_n}\right) = \int_{S_{i-1}}^{S_i} K\left(\frac{t-p}{h_n}\right)dt \quad \text{a.s.}$$

Let I_A be the indicator function of the set A and let i^* be the smallest $i \le n$ such that $S_{i+1}-T > h_n c$, where c is the constant in (K.1). If no such i exists, then let $i^* = n$. By (K.3),

$$n[Q_n^*(p) - Q_n(p)]^2 I_{[0,T]}(p)$$

$$= nh_n^{-2}\left\{\sum_{i=1}^{n} Z_i s_i \left[K\left(\frac{S_i-p}{h_n}\right) - K\left(\frac{S_i^*-p}{h_n}\right)\right] I_{[0,T]}(p) I_{[S_i^*-ch_n,1]}(p)\right\}^2$$

$$\le \Gamma^2\, nh_n^{-2}\left\{\sum_{i=1}^{n} Z_i s_i h_n^{-1}\left|S_i - S_i^*\right| I_{[0,T]}(p) I_{[S_i^*-ch_n,1]}(p)\right\}^2$$

$$\le \Gamma^2\, h_n^{-4} \sum_{i=1}^{n} Z_i^2 s_i^3\, n\, I_{[0,i^*]}(i)$$

$$\le \Gamma^2\, h_n^{-4}\, 2 \sup_{0 \le x \le T_{F_o}} \left|\hat{F}_n(x) - F_o(x)\right| \sum_{i=1}^{n} Z_i^2 d\hat{F}_n(Z_i) n s_i I_{[0,i^*]}(i)$$

where $T_{F_o} = \sup\{t: F_o(t) < 1\}$.

From Sander (1975), for $i \le i^*$, $0 \le ns_i \le [1-H(F_o^{-1}(T))]^{-1} + o_p(1)$, where $o_p(1)$ denotes a term converging to zero in probability as $n \to \infty$. By the results of Földes and Rejtö (1981), $\sup_{0 \le x \le T_{F_o}} \left|\hat{F}_n(x) - F_o(x)\right| = O((\log \log n/n)^{1/2})$ with probability one, and by a proof similar to that for Theorem 4.1 of Mauro (1985), $\sum_{i=1}^{n} Z_i^2 d\hat{F}_n(Z_i) \to E(X^{o^2})$ in probability. Therefore,

$$n[Q_n^*(p) - Q_n(p)]^2 I_{[0,T]}(p) = O_p(h_n^{-4}(\log \log n/n)^{1/2}),$$

completing the proof.

The following result for asymptotic normality of Q_n^* is therefore obtained.

THEOREM 2.2. Under the same conditions as in Lemma 2.2, for $0 < p < T$,

$$n^{1/2}[Q_n^*(p) - Q^o(p,h_n)] \to Z \text{ in distribution}$$

as $n \to \infty$, where Z is a normally distributed random variable with mean zero and variance σ_p^2.

3. ASYMPTOTIC MEAN EQUIVALENCE OF Q_n AND Q_n^*

It is shown in this section that Q_n and Q_n^* are equivalent in the mean and mean squared convergence sense. First, it is proven that $Q_n^*(p) - Q_n(p)$ converges to zero in the mean uniformly in p for certain choices of h_n.

For any distribution function G, define $T_G = \sup\{t: G(t) < 1\}$.

THEOREM 3.1. Assume that (h.1), (K.1) and (K.3) hold, H and F_o are continuous with $E|X^o| < \infty$, and $T_{F_o} \le T_H \le \infty$. Let ϕ be such a function on $[0,1-F(T^*))$ that $\phi(x) \ge x$, $\phi(x) \to \phi(0) = 0$ as $x \to 0^+$, and $1 - F_o(t) \le \phi(1 - F(t))$ for $t \in (T^*, T_{F_o})$, where T^* is arbitrary with $T^* < T_{F_o}$. Then

$$E[|Q_n^*(p) - Q_n(p)|] = O(\phi(d(\frac{\log \log n}{2n})^{1/2})h_n^{-2}),$$

where $d > 1$ is some constant.

PROOF. As in the proof of Lemma 2.2, using condition (K.3), with probability one

$$|Q_n^*(p) - Q_n(p)| \le h_n^{-1} \sum_{i=1}^{n} Z_i s_i \left|K\left(\frac{S_i-p}{h_n}\right) - K\left(\frac{S_i^*-p}{h_n}\right)\right|$$

$$\le h_n^{-2}\Gamma \sum_{i=1}^{n} Z_i s_i^2. \tag{3.1}$$

By continuity of F_o and using the definitions of S_i and s_i, (3.1) is less than or equal to

$$\Gamma h_n^{-2} \int_0^{T_{F_o}} x|\hat{F}_n(x) - \hat{F}_n(x^+)|d\hat{F}_n(x)$$

$$\le 2\Gamma h_n^{-2} \int_0^{T_{F_o}} x d\hat{F}_n(x) \sup_{0\le t\le T_{F_o}} |\hat{F}_n(t) - F_o(t)|.$$

Thus, by Corollary 2(v) of Csörgő and Horváth (1983) and Theorem 3.1 of Mauro (1985), the conclusion of the theorem follows.

Under similar conditions to those of Theorem 3.1, the asymptotic mean square equivalence of Q_n and Q_n^* can be obtained for some useful choices of the bandwidth sequence $\{h_n\}$.

THEOREM 3.2. Suppose that the conditions of Theorem 3.1 hold, replacing $E|X^o| < \infty$ by $E(X^{o2}) < \infty$. Then

$$E[(Q_n^*(p) - Q_n(p))^2] = O[\phi^2(d(\frac{\log\log n}{n})^{1/2}h_n^{-4}].$$

PROOF. By an argument similar to the proof of Theorem 3.1, with probability one

$$(Q_n^*(p) - Q_n(p))^2 \le 4\Gamma^2 h_n^{-4} \int_0^{T_{F_o}} x^2 d\hat{F}_n(x) \sup_{0\le t\le T_{F_o}} |\hat{F}_n(t) - F_o(t)|^2.$$

Again, the conclusion follows from Corollary 2(v) of Csörgő and Horváth (1983) and Theorem 3.1 of Mauro (1985).

From Theorems 3.1 and 3.2, if $\phi(d(\log\log n/2n)^{1/2})h_n^{-2} \to 0$ as $n \to \infty$, then $Q_n^*(p)$ and $Q_n(p)$ are asympotically equivalent in the mean. If ϕ is chosen so that $\phi(x) = (x/k)^{1/(1+\gamma)}$, for some $0 < k \leq 1$ and $\gamma \geq 0$, as in example (1) of Csörgő and Horváth (1983, p. 416), then (for $\gamma=0$ and $k=1$) the condition above becomes $(\log\log n/n)^{1/2}h_n^{-2} \to 0$, which is satisfied by $h_n = Dn^{-b}$ for $0 < b < 1/4$ and some positive constant D, for example.

4. MEAN SQUARE COVERGENCE

The following theorem yields the mean square convergence of $Q_n(p)$ to $Q^o(p)$ for appropriate choices of h_n. Also, combining Theorem 3.2 with Theorem 4.1 below gives conditions under which $Q_n^*(p)$ converges in mean square to $Q^o(p)$.

THEOREM 4.1. Let p_o be such that $0 \leq p_o < \min\{1, T_{H(F_o^{-1})}\}$. Suppose (h.1) and (K.1) hold, F_o is differentiable on some neighborhood of ξ_p^o, f_o is continuous at ξ_p^o with $f_o(\xi_p^o) > 0$, and $E(X^{o4}) < \infty$. Then for $0 \leq p < p_o$, $E\{[Q_n(p) - Q^o(p)]^2\} \leq g(n,h_n)$, where $g(n,h_n) = O(h_n^2) + O(h_n^{1/2}n^{-5/4}) + O(h_n^{-1/2}n^{-5/4}) + O(h_n n^{-3/4}(\log n)^{3/4}) + O(n^{-5/6}(\log n)^{5/6}) + O(n^{-1/4}(c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2))$ for some positive constants c_1, c_2, and c_3.

It should be noted that an example of a bandwidth sequence $\{h_n\}$ which will give $g(n,h_n) \to 0$ as $n \to \infty$ is $h_n = c_n n^{-\delta}$, $0 < \delta < 5/2$, where $c_n > 0$ is bounded by some positive constant d.

The proof of Theorem 4.1 is obtained from the following three lemmas.

LEMMA 4.1. Suppose (K.1) holds, F_o is differentiable on some neighborhood of ξ_p^o with $f_o(\xi_p^o) > 0$. Then

$$\{\int_{-c}^{c}[Q^o(p+h_n u) - Q^o(p)]\, K(u)du\}^2 = O(h_n{}^2).$$

PROOF. By the conditions on F_o and f_o, there exists a neighborhood A(p) of ξ_p^o so that $A = \sup_{t\varepsilon A(p)} [f_o(Q^o(t))]^{-1} < \infty$. Hence, by condition (K.1), the conclusion of the lemma follows.

LEMMA 4.2. Under the conditions of Theorem 4.1,

$$|E\{\int_{-c}^{c}[\hat{Q}_n(p+h_n u) - Q^o(p+h_n u)]K(u)du|$$
$$\leq O(h_n^{-\frac{1}{2}} n^{-5/4}) + O((\log n/n)^{3/4}).$$

PROOF. Let U_n denote the product-limit (PL) quantile process (Csorgő, 1983, Eq. (8.1.18), p. 118), and for each n choose $\varepsilon_n = (\log n/n)^{\frac{1}{2}}$. Define the events $A_n = [\sup_{-c\leq u\leq c} |U_n(p+h_n u) - (p+h_n u)| > \varepsilon_n]$. By Cheng (1984), for large n, $P[A_n] = O(n^{-5/2})$.

Write $E\{\int_{-c}^{c}[\hat{Q}_n(p+h_n u) - Q^o(p+h_n u)]K(u)du\} = E_1 + E_2$, where

$$E_1 = E\{\int_{-c}^{c}[Q^o(U_n(p+h_n u) - Q^o(p+h_n u)]K(u)du \cdot I_{A_n}\}$$

and

$$E_2 = E\{\int_{-c}^{c}[Q^o(U_n(p+h_n u)) - Q^o(p+h_n u)]K(u)du \cdot I_{A_n^c}\}.$$

Using the Holder inequality,

$$|E_1| \leq \{E[\int_{-c}^{c}Q^o(U_n(p+h_n u))K(u)du]^2\ P[A_n]\}^{\frac{1}{2}}$$
$$\leq \{\sup_u K(u)\ n^{-5/2}\ E[h_n^{-1}\int_{-c}^{c}\hat{Q}_n^2(t)dt]\}^{\frac{1}{2}}.$$

Thus, by Theorem 3.1 of Mauro (1985), $|E_1| = O(h_n{}^{-\frac{1}{2}}n^{-5/4})$.

Now, for E_2, using Taylor's expansion, there exists a ξ between $U_n(p+h_n u)$ and $p+h_n u$ such that

$$|E_2| = |E\{\int_{-c}^{c}[U_n(p+h_n u) - (p+h_n u)]K(u)du[f_o(Q^o(\xi))]^{-1}I_{A_n^c}\}|$$
$$= |E\{\int_{-c}^{c}[f_o(Q^o(\xi))]^{-1}\ [U_n(p+h_n u) - (p+h_n u)$$

$$+ n^{-1/2}\tilde{\alpha}_n(p+h_nu)]K(u)du \cdot I_{A_n^c}$$

$$- \int_{-c}^{c}[f_o(Q^o(\xi))]^{-1}\ n^{-1/2}\tilde{\alpha}_n(p+h_nu)K(u)du\ I_{A_n^c}\}|,$$

where $\tilde{\alpha}_n$ denotes the uniform PL - empirical process (Csörgő, 1983, p. 117). For n sufficiently large, $|1/f_o(Q^o(\xi))| \le A$, where A is defined in the proof of Lemma 4.1, and by Cheng's (1984) Theorem 2, for large n,

$$\sup_{-c\le u\le c}[U_n(p+h_nu) - (p+h_nu) + n^{-1/2}\tilde{\alpha}_n(p+h_nu)]$$
$$= O((\log n/n)^{3/4}). \tag{4.1}$$

Next,

$$|E\int_{-c}^{c}n^{-1/2}\tilde{\alpha}_n(p+h_nu)K(u)du/f_o(Q^o(\xi))I_{A_n^c}|$$

$$\le n^{-1/2}A\ E\{\int_{-c}^{c}|\tilde{\alpha}_n(p+h_nu)-n^{-1/2}K^*(p+h_nu,n)|K(u)du\}$$
$$+ n^{-1/2}|E\int_{-c}^{c}n^{-1/2}K^*(p+h_nu,n)K(u)duI_{A_n^c}|, \tag{4.2}$$

where $K^*(t,s)$ denotes the generalized Kiefer process (Csörgő, 1983, p. 118). By Theorem 8.1.1 of Csörgő (1983) (or Burke, Csörgő, and Horváth, 1981),

$$\sup_{-c\le u\le c}|\tilde{\alpha}(p+h_nu) - n^{-1/2}K^*(p+h_nu,n)| \overset{a.s.}{=} O(n^{-1/3}(\log n)^{5/2}). \tag{4.3}$$

Therefore, from (4.1)-(4.3), for large n,

$$|E_2| \le E\{A\int_{-c}^{c}|[U_n(p+h_nu) - (p+h_nu) + n^{-1/2}\tilde{\alpha}_n(p+h_nu)]K(u)|du$$
$$+ n^{-1/2}|E\int_{-c}^{c}n^{-1/2}K^*(p+h_nu,n)K(u)duI_{A_n^c}|$$
$$+ O(n^{-1/3}(\log n)^{5/2})n^{-1/2}$$
$$\le O((\log n/n)^{3/4}) + O(n^{-5/6}(\log n)^{5/2}$$
$$= O((\log n/n)^{3/4}),$$

since $E\int_{-c}^{c}n^{-1/2}|K^*(p+h_nu,n)|K(u)du < \infty$, and the proof is completed.

LEMMA 4.3. Suppose the conditions of Theorem 4.1 hold. Then

$$E\{[\int_{-c}^{c}(\hat{Q}_n(p+h_nu) - Q^o(p+h_nu)K(u)du]^2\}$$
$$\le O(n^{-5/6}(\log n)^{5/2}) + O(h_n^{-1/2}n^{-5/4})$$
$$+ O(n^{-1/4}(c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2)).$$

PROOF. As in the proof of Lemma 4.2, write

$$E\{[\int_{-c}^{c}(\hat{Q}_n(p+h_nu) - Q^o(p+h_nu))K(u)du]^2\} = E_3 + E_4,$$

where

$$E_3 = E\{[\int_{-c}^{c}(\hat{Q}_n(p+h_nu) - Q^o(p+h_nu))K(u)du]^2 I_{A_n^c}\}$$

and

$$E_4 = E\{[\int_{-c}^{c}(\hat{Q}_n(p+h_nu) - Q^o(p+h_nu))K(u)du]^2 I_{A_n}\}.$$

Now,

$$|E_4| \le E\{[\int_{-c}^{c}\hat{Q}_n^2(p+h_nu)K(u)du)]I_{A_n}\}$$
$$\le \{E[(\int_{-c}^{c}\hat{Q}_n^2(p+h_nu)K(u)du)^2] \cdot P(A_n)\}^{1/2}$$
$$\le \{E[\int_{-c}^{c}\hat{Q}_n^4(p+h_nu)K(u)du]\}^{1/2} n^{-5/4}$$
$$\le [\sup_u K(u)]^{1/2} \{h_n^{-1}E\int\hat{Q}_n^4(t)dt\}^{1/2} n^{-5/4}$$
$$\le O(h_n^{-1/2}n^{-5/4}) \tag{4.4}$$

by Hölder's inequality and Mauro's (1985) Theorem 3.1.

Next, using the Taylor expansion as in the proof of Lemma 4.2,

$$|E_3| \le |E\{[\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}(U_n(p+h_nu) - (p+h_nu)$$
$$+ n^{-1/2}\tilde{\alpha}_n(p+h_nu))K(u)du]^2\}|$$
$$+ |E\{[\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}n^{-1/2}\tilde{\alpha}_n(p+h_nu)K(u)du]^2\}|$$
$$+ 2|E\{\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}[U_n(p+h_nu) - (p+h_nu)$$
$$+ n^{-1/2}\tilde{\alpha}_n(p+h_nu)]K(u)du$$
$$\times \int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}n^{-1/2}\tilde{\alpha}_n(p+h_nu)K(u)du\}|$$
$$\equiv E_{31} + E_{32} + E_{33}.$$

Now, as in the proof of Lemma 4.2,

$$|E_{31}| \le [O(\log n/n)^{3/4})]^2 = O((\log n/n)^{3/2}).$$

Also,

$$
\begin{aligned}
E_{32} &\le n^{-\frac14}\left|E\{[\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}(\tilde{\alpha}_n(p+h_nu) \right.\\
&\quad \left. - n^{-\frac12}K^*(p+h_nu,n))K(u)du]^2\}\right| \\
&\quad + n^{-\frac14}\left|E\{\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}n^{-\frac12}K^*(p+h_nu)K(u)du\}^2\right| \\
&\quad + n^{-\frac14}2\left|E\{\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}[\tilde{\alpha}_n(p+h_nu)\right. \\
&\quad - n^{-\frac12}K^*(p+h_nu,n)]K(u)du \\
&\quad \left.\times\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}n^{-\frac12}K^*(p+h_nu,n)K(u)du\}\right| \\
&\le n^{-\frac14}[O(n^{-1/3}(\log n)^{5/2})]^2 \\
&\quad + A^2n^{-\frac14}E\{\int_{-c}^{c}n^{-1}K^{*2}(p+h_nu,n)K^2(u)du\} \\
&\quad + 2A^2n^{-\frac14}O(n^{-1/3}(\log n)^{5/2})E\int_{-c}^{c}n^{-\frac12}|K^*(p+h_nu,n)|K(u)du. \qquad (4.5)
\end{aligned}
$$

Since $n^{-\frac12}K^*(t,n)$ is a Gaussian process in t, the last expectation in (4.5) is finite, and by Fubini's theorem and section 8.2 of Csörgő (1983), for n sufficiently large,

$$
\begin{aligned}
&E\{\int_{-c}^{c}n^{-1}K^{*2}(p+h_nu,n)K(u)du\} \\
&= \int_{-c}^{c}[1-(p+h_nu)]^2\left(\int_0^{p+h_nu}\frac{dx}{(1-x)^2(1-H(Q^o(x)))}\right)K(u)du \\
&\le \left[\int_0^{p_o}\frac{dx}{(1-x)^2(1-H(Q^o(x)))}\right][\int_{-c}^{c}(1-p)^2K(u)du \\
&\quad +2(1-p)h_n\int_{-c}^{c}uK(u)du+h_n^2\int_{-c}^{c}u^2K(u)du]. \qquad (4.6)
\end{aligned}
$$

Therefore, from (4.6) and condition (K.1), (4.5) yields for appropriate constants c_1, c_2, and c_3

$$|E_{32}| \le O(n^{-7/2}(\log n)^{5/2}) + n^{-\frac14}[c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2].$$

Similarly, since $|E\int_{-c}^{c}(f_o(Q^o(\xi)))^{-1}n^{-\frac12}\tilde{\alpha}_n(p+h_nu)K(u)du| < \infty$,

$$|E_{33}| \le O(n^{-5/6}(\log n)^{5/2}).$$

Therefore, combining these results for E_3,

$$
\begin{aligned}
&E\{\int_{-c}^{c}(\dot{Q}_n(p+h_nu) - Q^o(p+h_nu))K(u)du]^2\} \\
&\quad \le O(h_n^{-\frac12}n^{-5/4}) + O((\log n/n)^{3/2}) + O(n^{-7/2}(\log n)^{-5/2}) \\
&\quad + O(n^{-\frac14}[c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2]) + O(n^{-5/6}(\log n)^{5/2})
\end{aligned}
$$

$$= O(h_n^{-1/4} n^{-5/4}) + O(n^{-5/6}(\log n)^{5/2})$$
$$+ O(n^{-1/4}[c_1(1-p)^2 + c_2(1-p)h_n + c_3h_n^2]),$$

completing the proof.

The proof of Theorem 4.1 is then obtained from Lemmas 4.1-4.3 by writing, for sufficiently large n,

$$E\{[Q_n(p) - Q^o(p)]^2\}$$
$$= E\{[\int_{-c}^{c}(\hat{Q}_n(p+h_nu) - Q^o(p+h_nu))K(u)du]^2\}$$
$$+ [\int_{-c}^{c}(Q^o(p+h_nu) - Q^o(p))K(u)du]^2$$
$$+ 2\int_{-c}^{c}[Q^o(p+h_nu) - Q^o(p)]K(u)du$$
$$\times E\{\int_{-c}^{c}[\hat{Q}_n(p+h_nu) - Q^o(p+h_nu)]K(u)du\}.$$

It should be noted that the conditions of Theorem 4.1, as well as those of the other theorems in this paper, are not restrictive and are similar to conditions for results for right-censored data obtained by previous authors. See Chapter 8 of Csörgő (1983) for various references and Cheng (1984), for example.

REFERENCES

BURKE, M.D., CSÖRGŐ, S. and HORVÁTH, L. (1981). Strong Approximations of Some Biometric Estimates under Random Censorship. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwande Gebiete, 56, 87-112.

CHENG, K.F. (1984). On Almost Sure Representations for Quantiles of the Product Limit Estimator with Applications. Sankhya, Ser. A 46, 426-443.

CSÖRGŐ, M. (1983). Quantile Processes with Statistical Applications. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA.

CSÖRGŐ, S. and HORVÁTH, L. (1983). The Rate of Strong Uniform Consistency for the Product-Limit Estimator. Zeitschrift für Wahrscheinlichkeitstheorie und Verwande Gebiete, 62, 411-426.

EFRON, B. (1967). The Two-Sample Problem with Censored Data. Proceedings of the Fifth Berkeley Symposium, 4, 831-853.

FALK, M. (1984). Relative Deficiency of Kernel Type Estimators of Quantiles. Annals of Statistics, 12, 261-268.

FALK, M. (1985). Asymptotic Normality of the Kernel Quantile Estimator. Annals of Statistics, 13, 428-433.

FÖLDES, A., and REJTÖ, L. (1981). A LIL Type Result for the Product Limit Estimator. Zeitschrift für Wahrscheinlichkeitstheorie und Verwande Gebiete, 56, 76-86.

KAPLAN, E.L., and MEIER, P. (1958). Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association, 53, 457-481.

LIO, Y.L., PADGETT, W.J., and YU, K.F. (1985). On the Asymptotic Properties of a Kernel-Type Quantile Estimator from Censored Samples. Journal of Statistical Planning and Inference, in press.

MAURO, D. (1985). A Combinatoric Approach to the Kaplan-Meier Estimator. Annals of Statistics, 13, 142-149.

PADGETT, W.J. (1985). A Kernel-Type Estimator of a Quantile Function from Right-Censored Data. Journal of the American Statistical Association, in press.

PARZEN, E. (1979). Nonparametric Statistical Data Modeling. Journal of the American Statistical Association, 74, 105-121.

SANDER, J. (1975). The Weak Convergence of Quantiles of the Product Limit Estimator. Technical Report 5, Stanford University, Dept. of Statistics.

YANG, S.S. (1985). A Smooth Nonparametric Estimator of a Quantile Function. Journal of the American Statistical Association, in press.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A162 837

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Stat. Tech Rep. No. 108 (62G05-14)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOS[illegible]

6a. NAME OF PERFORMING ORGANIZATION
Department of Statistics

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS *(City, State and ZIP Code)*
University of South Carolina
Columbia, SC 29208

7b. ADDRESS *(City, State and ZIP Code)*
Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR, ARO

8b. OFFICE SYMBOL *(If applicable)*
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-84-0156

8c. ADDRESS *(City, State and ZIP Code)*
Bolling AFB DC 20332

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE *(Include Security Classification)*
Some Convergence Results for Kernel-Type Quantile Estimators under Censoring

12. PERSONAL AUTHOR(S)
Y. L. Lio and W. J. Padgett

13a. TYPE OF REPORT
Technical

13b. TIME COVERED
FROM ________ TO ________

14. DATE OF REPORT *(Yr., Mo., Day)*
1985 November

15. PAGE COUNT
15

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Random right-censorship; Kernel estimation; Product-limit quantile function; Asymptotic normality; Mean convergence; Mean-squared error.

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Based on right-censored data from a lifetime distribution F_o, a kernel-type estimator of the quantile function $Q^o(p)$, defined by $Q_n(p) = h_n^{-1}\int_0^1 \hat{Q}_n(t)K((t-p)/h_n)dt$, is studied. This estimator is smoother than the product-limit quantile function $\hat{Q}_n(p) = \inf\{t: \hat{F}_n(t) \geq p\}$, where $\hat{F}_n$ denotes the product-limit estimator of F_o from the censored sample. Under the random censorship model and general conditions on h_n, K, and F_o, asymptotic normality of $Q_n(p)$ and a simpler approximation to it, $Q_n^*(p)$, is shown, and mean square convergence of Q_n is proven. Also, the asymptotic mean equivalence of Q_n and Q_n^* is shown.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED XX SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Maj. Brian W. Woodruff

22b. TELEPHONE NUMBER *(Include Area Code)*
(202) 767-5027

22c. OFFICE SYMBOL
NM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

END

FILMED

2-86

DTIC