Statistica Sinica **10**(2000), 1217-1229

QUANTILE ESTIMATION FOR LEFT TRUNCATED AND RIGHT CENSORED DATA

Xian Zhou, Liuquan Sun and Haobo Ren

Hong Kong Polytechnic University, Academia Sinica and Peking University

Abstract: In this paper we study the estimation of a quantile function based on left truncated and right censored data by the kernel smoothing method. Asymptotic normality and a Berry-Esseen type bound for the kernel quantile estimator are derived. Monte Carlo studies are conducted to compare the proposed estimator with the PL-quantile estimator.

Key words and phrases: Asymptotic normality, Berry-Esseen type bound, quantile function, kernel estimation, truncated and censored data.

1. Introduction

In medical follow-up or in engineering life testing studies one may not be able to observe the variable of interest, referred to hereafter as the lifetime. Among the different forms in which incomplete data appear, right censoring and left truncation are common. Left truncation may occur if the time origin of the lifetime precedes the time origin of the study. Only those individuals that fail after the start of the study are being followed, otherwise they are left truncated. Those that are followed are further subject to right censoring during the follow-up period. See for example, an AIDS study by Struthers and Farewell (1989) where "lifetime" is the incubation period, and the truncation variable is the time from infection until entry to the study. Formally, let (X, T, Y) denote random variables, where X is the lifetime with continuous distribution function (d.f.) F, T is the random left truncation time with arbitrary d.f. G, and Y is the random right censoring time with arbitrary d.f. H. It is assumed that X, T, Y are mutually independent and, without loss of generality, that they are nonnegative. In the random left truncation and right censoring (LTRC) model one observes (Z, T, δ) if $Z \geq T$, where $Z = X \wedge Y = \min(X, Y)$ and $\delta = I(X \leq Y)$ is the indicator of censoring status. When $Z < T$ nothing is observed. Let $\alpha \equiv P(T \leq Z) > 0$, and let W be the d.f. of Z, i.e., $1 - W = (1 - F)(1 - H)$. Let $(Z_i, T_i, \delta_i), i = 1, \ldots, n$ be an independent and identically distributed (i.i.d.) sample of (Z, T, δ) which is observed (i.e., $Z_i \geq T_i$). Let $C(x) = P(T \leq x \leq Z \mid T \leq Z) = \alpha^{-1}G(x)(1 - W(x-))$, and consider its empirical estimate $C_n(x) = n^{-1}\sum_{i=1}^{n} I(T_i \leq x \leq Z_i)$.

For LTRC data, it is important to be able to obtain nonparametric estimates of various characteristics of the distribution function F. Tsai, Jewell and Wang (1987) gave the nonparametric maximum likelihood estimator of F itself, called the product-limit (PL) estimator, as

$$\hat{F}_n(x) = \begin{cases} 1 - \prod_{Z_i \le x} \left(1 - [nC_n(Z_i)]^{-1}\right)^{\delta_i}, & x < Z_{(n)}, \\ 1, & x \ge Z_{(n)}, \end{cases} \tag{1.1}$$

where $Z_{(n)} = \max(Z_1, \ldots, Z_n)$. The properties of $\hat{F}_n$ have been studied by Wang (1987), Gu and Lai (1990), Lai and Ying (1991), Gijbels and Wang (1993), and Zhou (1996), among others. Nonparametric estimates of the density and hazard rate for F have been studied by Gijbels and Wang (1993), Sun and Zhou (1998), Uzunoḡullari and Wang (1992), Gu (1995) and Sun (1997).

One characteristic of the distribution function F that is of interest is the quantile function. It plays an important role in various statistical applications, especially in data modeling, reliability and medical studies. Let $Q(t) \equiv F^{-1}(t) = \inf\{x : F(x) \ge t\}$, $0 < t < 1$, be the quantile function of F. We focus here on estimating the quantile function based on LTRC data.

A natural estimator of Q is the PL-quantile function defined as $\hat{F}_n^{-1}(x) = \inf\{u : \hat{F}_n(u) \ge x\}, 0 < x < 1$. In case Q is a continuous function, it may be more suitable to use a smooth estimator rather than the step function $\hat{F}_n^{-1}$, since smoothing reduces random variation in the data and allows a better display of interesting features of the lifetime distribution function. Numerous smooth quantile function estimators have been proposed for complete samples. In the censored model, Padgett (1986), Lio, Padgett and Yu (1986), Xiang (1995a, b) studied kernel-type quantile function estimators. None of their results included the case where both left truncation and right censoring are involved however. On the other hand, Gürler, Stute and Wang (1993) provide a strong representation of the empirical quantile function for left truncated data, together with asymptotic normality and the Law of Iterated Logarithm (LIL), which can be easily extended to the LTRC model. In this paper, for the LTRC model, we propose and study the kernel quantile function estimator of the form

$$\hat{Q}_n(t) = h_n^{-1} \int_0^1 \hat{F}_n^{-1}(x) K\Big(\frac{x-t}{h_n}\Big) dx, \tag{1.2}$$

where K is a kernel function and $\{h_n\}_{n\ge1}$ is a sequence of positive bandwidths with $h_n \to 0$ as $n \to \infty$.

The first aim of this article is to derive the asymptotic normality of the kernel quantile function estimator. The quality of the approximation to the distribution of $\hat{Q}_n$ is an important issue of both theoretical and practical interest (for example, construction of confidence intervals based on the asymptotic normality). Thus we also establish a Berry-Esseen type bound of $\hat{Q}_n$. A small Monte Carlo simulation study shows that, for certain values of the bandwidth h_n, $\hat{Q}_n$ performs better than $\hat{F}_n^{-1}$ in the sense of having smaller estimated mean squared errors (MSE).

For any d.f. L, denote the left and right endpoints of its support by $a_L = \inf\{x : L(x) > 0\}$ and $b_L = \sup\{x : L(x) < 1\}$, respectively. In the current model, as discussed by Gijbels and Wang (1993) and Zhou (1996), we assume that $a_G \leq a_W, b_G \leq b_W$ and

$$\int_{a_W}^{\infty} \frac{dF(x)}{G^2(x)} < \infty. \tag{1.3}$$

Condition (1.3) requires that the left truncation is not too heavy. As for the kernel function $K(x)$, assume

$(K1)$ K is of bounded variation with support $[-1, 1]$;

$(K2)$ K is Lipschitz continuous of order one;

$(K3)$ For some integer $r \geq 2$,

$$\frac{1}{j!}\int x^j K(x)dx = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{if } j = 1, \ldots, r-1, \\ c_r \neq 0 & \text{if } j = r. \end{cases}$$

These assumptions are the usual ones encountered in the kernel method of curve estimation, as in Padgett (1986) and Uzunoḡullari and Wang (1992).

This article is organized as follows. Main results are stated in Section 2. Some simulation results are presented in Section 3. Proofs of the results are deferred to the Appendix.

2. Main Results

We first give the asymptotic normality of $\hat{Q}_n$.

Theorem 1. *Under $(K1)$ and $(K3)$, assume that F is r times continuously differentiable in a neighborhood of $Q(t)$ with $F'(Q(t)) = f(Q(t)) > 0$, $a_w < Q(t) < b_w$, where r is the same as that in (K3). If $n^{1/2}h_n^r \to 0$ as $n \to \infty$,*

$$n^{1/2}(\hat{Q}_n(t) - Q(t)) \to N(0, \sigma^2(t)) \text{ in distribution as } n \to \infty, \tag{2.1}$$

where the variance is

$$\sigma^2(t) = \frac{(1-t)^2}{f^2(Q(t))} \int_{a_W}^{Q(t)} \frac{dF(x)}{C(x)(1 - F(x))}. \tag{2.2}$$

To apply this result to make some standard inferences in hypothesis testing and construction of confidence intervals, an estimator of the variance in (2.2) is needed and a natural nonparametric estimator is given below. Let $W_1(x) = P(Z \le x, \delta = 1 | T \le Z) = \alpha^{-1}\int_{-\infty}^{x} G(u)(1 - H(u-))dF(u)$. Then it can be checked that

$$\sigma^2(t) = \frac{(1-t)^2}{f^2(Q(t))}\int_{a_W}^{Q(t)} \frac{dW_1(x)}{C^2(x)}. \tag{2.3}$$

Thus $\sigma^2(t)$ can be estimated consistently by replacing the density f in (2.3) by a density estimate from Gijbels and Wang (1993), and all other unknown quantities in (2.3) by their empirical estimates.

We now give a Berry-Esseen type bound for $\hat{Q}_n$ to assess the quality of the normal approximation in Theorem 1.

Theorem 2. *Under* $(K1)$, $(K2)$ *and* $(K3)$, *assume that* F *is* r*-time continuously differentiable in a neighborhood of* $Q(t)$ *with* $F'(Q(t)) = f(Q(t)) > 0$, $a_W < Q(t) < b_W$. *If* $a_G < a_W$ *and* $nh_n^2/\log n \to \infty$ *as* $n \to \infty$, *there exists a positive constant* $M = M(t)$ *such that for all* $n \ge 1$,

$$\sup_y \left| P\{n^{1/2}(\hat{Q}_n(t) - Q(t)) \le y\sigma(t)\} - \Phi(y)\right| \le M\left(n^{-1/2}h_n^{-1}\log n + n^{1/2}h_n^r + h_n^{1/2}\right), \tag{2.4}$$

where $\sigma(t)$ *is defined by* (2.2).

3. Simulation Studies

A small Monte Carlo study was performed in order to provide some small-sample comparisons of the kernel estimator $\hat{Q}_n$ with the PL-quantile $\hat{F}_n^{-1}$ in terms of mean squared error. The study also provides some insight into the choices of h_n that might be used to estimate Q with smaller mean squared error than $\hat{F}_n^{-1}$. For the LTRC model, the following cases are simulated.

(i) $F(x) = 1 - \exp(-x)$, $H(x) = 1 - \exp(-x)$ and $G(x) = 1 - \exp(-2x)$, with 50% censoring and 50% truncation respectively.

(ii) $F(x) = 1 - \exp(-x)$, $H(x) = 1 - \exp(-3x/7)$ and $G(x) = 1 - \exp(-30x/7)$, with 30% censoring and 25% truncation respectively.

The triangular density function $K(x) = (1 - |x|)I(|x| \le 1)$ was used as the kernel function for the estimator $\hat{Q}_n$. The ratio of the mean squared error of $\hat{F}_n^{-1}(t)$ to that of the smoothed estimator $\hat{Q}_n(t)$ were computed for various values of $t \in (0,1)$ and sample size $n = 100$. For each case, 1000 replications were done on S-PLUS on a PC-pentium 200 computer. The simulations were run for various values of h_n and the results show that $\hat{Q}_n$ gives reasonable performance for $h_n \le 0.40$ only. On the other hand, the results for h_n below 0.10 do not vary much. Some of the simulation results for h_n between 0.05 and 0.40 are presented in Tables 1-2 below, and these show that, for each value t listed, there is a range of window widths h_n such that the estimator $\hat{Q}_n$ has smaller estimated mean squared error than the PL-quantile estimator. In particular, this is true for the median estimators $\hat{Q}_n(0.5)$ and $\hat{F}_n^{-1}(0.5)$.

From a range of h_n values, we found that $h_n = 0.15$ comes close to giving the smallest discrepancy between the kernel estimator and the quantile. Figures 1-2 show the plots of the quantile, kernel estimator $\hat{Q}_n$ and the PL-quantile estimator $\hat{F}_n^{-1}$ with $h_n = 0.15$, from which we see that $\hat{Q}_n$ looks better than $\hat{F}_n^{-1}$ except for large values of t.

As one would expect for censoring, truncation and bound effect, the performance of either estimator ($\hat{Q}_n$ or $\hat{F}_n^{-1}$) at large values of t is not as good as that for values near 0.5.

Table 1. Ratios of mean squared errors for Case (i): 50% censoring and 50% truncation ($n = 100$).

	h_n									
t	0.05	0.07	0.09	0.12	0.15	0.20	0.25	0.30	0.35	0.40
0.10	1.056	1.086	1.095	1.128	1.185	1.243	1.225	1.144	1.023	0.886
0.25	1.054	1.065	1.076	1.094	1.114	1.149	1.183	1.216	1.236	1.224
0.50	1.077	1.082	1.098	1.103	1.129	1.140	1.151	1.138	1.046	0.997
0.75	1.126	1.149	1.153	1.145	1.132	1.101	1.044	1.134	1.410	1.654
0.90	1.353	1.433	1.517	1.724	1.897	1.431	0.926	0.749	0.594	0.401
0.95	1.138	1.040	0.907	0.642	0.499	0.383	0.326	0.292	0.269	0.254

Note: Ratio=MSE($\hat{F}_n^{-1}$)/MSE($\hat{Q}_n$).

Table 2. Ratios of mean squared errors for Case (ii): 30% censoring and 25% truncation ($n = 100$).

	h_n									
t	0.05	0.07	0.09	0.12	0.15	0.20	0.25	0.30	0.35	0.40
0.10	1.051	1.065	1.079	1.104	1.166	1.244	1.256	1.207	1.118	1.008
0.25	1.048	1.056	1.065	1.082	1.099	1.131	1.159	1.185	1.203	1.195
0.50	1.029	1.039	1.048	1.059	1.067	1.075	1.078	1.067	1.040	0.989
0.75	1.034	1.043	1.051	1.057	1.054	1.014	0.987	0.867	1.039	1.260
0.90	1.062	1.043	0.996	1.022	1.235	0.979	0.657	0.456	0.356	0.301
0.95	1.163	1.479	1.485	0.960	0.622	0.507	0.422	0.352	0.279	0.235

Note: Ratio=MSE($\hat{F}_n^{-1}$)/MSE($\hat{Q}_n$).

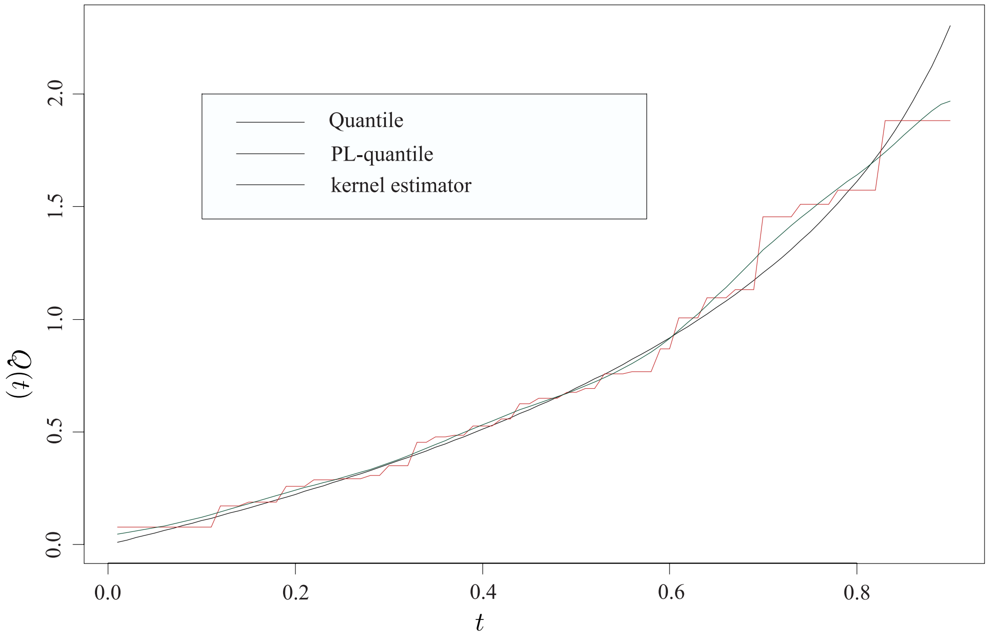

Figure 1. Plots of PL-quantile and the kernel estimator for case (i) with $n = 50$ and $h_n = 0.15$.

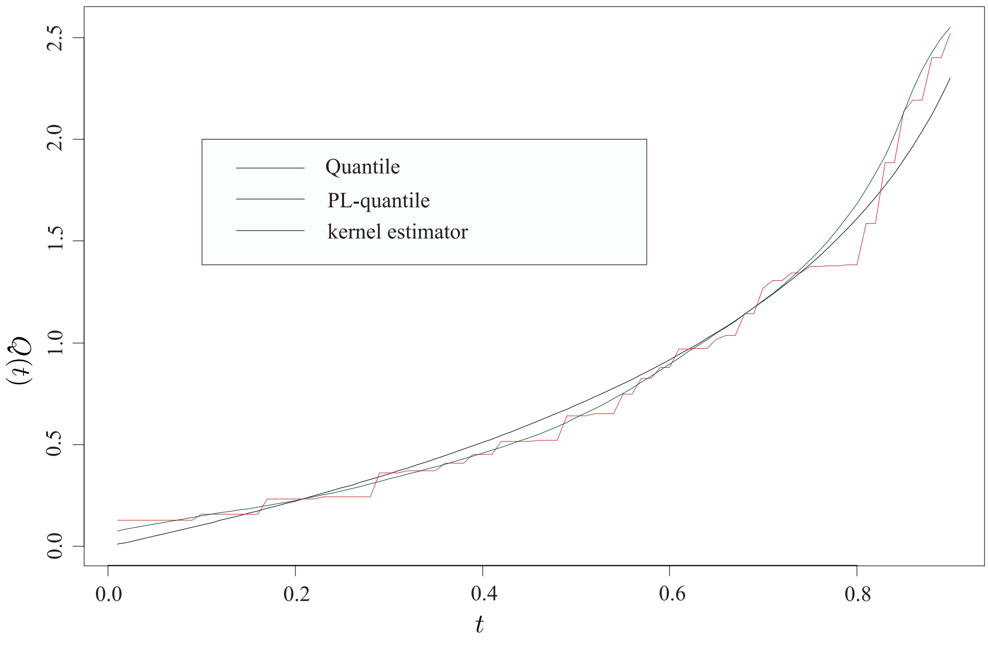

Figure 2. Plots of PL-quantile and the kernel estimator for case (ii) with $n = 50$ and $h_n = 0.15$.

Appendix. Proofs of Theorems

First we present preliminary results needed in the proofs of the theorems. Let

$$L_n(x) = \int_{a_W}^{x} \frac{d[W_{1n}(u) - W_1(u)]}{C(u)} - \int_{a_W}^{x} \frac{C_n(u) - C(u)}{C^2(u)} dW_1(u).$$

From Theorem 1 of Gijbels and Wang (1993) and Theorem 2 of Zhou (1996), we have that for $a_W \le x \le b < b_W$,

$$\hat{F}_n(x) - F(x) = (1 - F(x))L_n(x) + R_n(x), \tag{A.1}$$

where

$$\sup_{a_W \le x \le b} |R_n(x)| = O(n^{-1} \log^{1+\beta} n) \quad \text{a.s.} \tag{A.2}$$

with $\beta = 0$ when $a_G < a_W$, and $\beta > 1/2$ when $a_G = a_W$.

Introduce the PL-process $\hat{\alpha}_n(x) = n^{1/2}(\hat{F}_n(x) - F(x))$. The following lemma, of interest in its own right, provides a general version of the oscillation behavior of $\hat{\alpha}_n$.

Lemma A.1. *Assume that F is Lipschitz continuous of order one on $[c, d]$, $a_W < c \le d < b_W$. Let $\{a_n, n \ge 1\}$ be a sequence of positive constants tending to zero with $na_n / \log n \to \infty$ as $n \to \infty$. Then*

$$\sup_{c \le x, y \le d, |x-y| \le a_n} |\hat{\alpha}_n(x) - \hat{\alpha}_n(y)| = O\left(a_n^{1/2} (\log n)^{1/2}\right) + O\left(n^{-1/2} (\log n)^{1+\beta}\right) \quad \text{a.s.,} \tag{A.3}$$

where β is defined in (A.2).

Proof. It follows from (A.1) that

$$\begin{aligned} n^{-\frac{1}{2}}(\alpha_n(x) - \alpha_n(y)) &= (1 - F(x)) \int_y^x \frac{d[W_{1n}(u) - W_1(u)]}{C(u)} + (F(y) - F(x))L_n(y) \\ &\quad - (1 - F(x)) \int_y^x \frac{C_n(u) - C(u)}{C^2(u)} dW_1(u) + (R_n(x) - R_n(y)) \\ &:= \sum_{i=1}^{4} \Gamma_{ni}(x, y). \end{aligned} \tag{A.4}$$

First, by the continuity of F together with a partitioning argument similar to that in Burke, Csörgő and Horváth (1988), it can be shown that

$$\sup_{c \le x, y \le d, |x-y| \le a_n} |\Gamma_{n1}(x, y)| = O\left(n^{-1/2} a_n^{1/2} (\log n)^{1/2}\right) \quad \text{a.s.} \tag{A.5}$$

Next, in view of (1.3), the process $L_n(y)$ is the sum of two empirical processes over VC classes of functions with square integrable envelope on $[c, d]$, so it satisfies the LIL (see, e.g. Arcones and Giné (1995)), i.e., its sup over $[c,d]$ is a.s. of the order $(n^{-1}\log\log n)^{1/2}$. Moreover, note that $C_n - C$ is the difference of two empirical processes and $\inf_{c\le y\le d} C(y) > 0$. Therefore, using the LIL for empirical processes, we obtain that for all n sufficiently large,

$$\sup_{c\le x,y\le d,|x-y|\le a_n} |\Gamma_{ni}(x,y)| = O\left(n^{-1/2}a_n^{1/2}(\log n)^{1/2}\right) \quad \text{a.s.,} \quad i=2,3, \tag{A.6}$$

and by (A.2),

$$\sup_{c\le x,y\le d,|x-y|\le a_n} |\Gamma_{n4}(x,y)| = O\left(n^{-1}\log^{1+\beta} n\right) \quad \text{a.s.} \tag{A.7}$$

Lemma A.1 now follows from (A.4)–(A.7).

Proof of Theorem 1. The proof of Theorem 1 relies on the weak asymptotic representation for the PL-quantile function $\hat{F}_n^{-1}$ as given in (A.8) below.

Let $a_W < Q(t_1) \le Q(t_2) < b_W$. If F has a continuous and positive density f on $[Q(t_1)-\eta, Q(t_2)+\eta]$ for some $\eta > 0$, it follows from similar arguments as in Gürler, Stute and Wang (1993) that, as $n\to\infty$,

$$\sup_{t_1\le t\le t_2}\left|\hat{F}_n^{-1}(t) - Q(t) - \frac{t-\hat{F}_n(Q(t))}{f(Q(t))}\right| = o_p(n^{-1/2}). \tag{A.8}$$

Let $\bar{Q}_n(t) = h_n^{-1}\int_0^1 Q(x)K\left(\frac{x-t}{h_n}\right)dx$. Using (A.1) and (A.8), we get

$$\hat{Q}_n(t) - \bar{Q}_n(t) = n^{-1}\sum_{i=1}^n \int_{t-h_n}^{t+h_n} [-\varphi_i(Q(x))]\frac{1}{f(Q(x))h_n}K\Big(\frac{x-t}{h_n}\Big)dx + o_p(n^{-1/2}),$$

where

$$\varphi_i(x) = (1-F(x))\left[\frac{I(Z_i\le x,\delta_i=1)}{C(Z_i)} - \int_{a_W}^x \frac{I(T_i\le u\le Z_i)}{C^2(u)}dW_1(u)\right]. \tag{A.9}$$

Thus an application of Liapunov's form of the Central Limit Theorem gives

$$n^{1/2}(\hat{Q}_n(t) - \bar{Q}_n(t)) \to N(0,\sigma^2(t)) \text{ in distribution as } \quad n\to\infty. \tag{A.10}$$

Note that

$$\hat{Q}_n(t) - Q(t) = \left[\hat{Q}_n(t) - \bar{Q}_n(t)\right] + \left[\bar{Q}_n(t) - Q(t)\right]. \tag{A.11}$$

It follows from a Taylor expansion that

$$\bar{Q}_n(t) - Q(t) = \frac{h_n^r}{r!}\int_{-1}^1 x^r K(x) Q^{(r)}(t+h_n\theta x)dx$$

for some $0 \le \theta \le 1$, where $Q^{(r)}$ is the rth derivative of Q. Thus

$$n^{1/2}|\bar{Q}_n(t) - Q(t)| = O(n^{1/2}h_n^r) \to 0 \quad as \quad n \to \infty. \tag{A.12}$$

Theorem 1 then follows immediately from (A.10)–(A.12).

The following lemma is useful in proving Theorem 2.

Lemma A.2. Let U, V be two random variables. Then for any $a > 0$,

$$\sup_y |P(U + V \le y) - \Phi(y)| \le \sup_y |P(U \le y) - \Phi(y)| + \frac{a}{\sqrt{2\pi}} + P(|V| > a), \tag{A.13}$$

$$\sup_y |P(U \le yV) - \Phi(y)| \le \sup_y |P(U \le y) - \Phi(y)| + P(|V - 1| > a) + a. \tag{A.14}$$

Proof. The first inequality follows from Lemma 2 of Chang and Rao (1989), while the second is a consequence of Michel and Pfanzagl (1971).

In the sequel $M = M(x)$ denotes a positive constant which will vary in different contexts.

Proof of Theorem 2. Using a change of variable theorem (cf. Shorack and Wellner (1986, p.25)), we have

$$\hat{Q}_n(t) - \bar{Q}_n(t) = -\int_0^\infty \Big(\int_{(F(x)-t)/h_n}^{(\hat{F}_n(x)-t)/h_n} K(u)du \Big) dx = S_{n1} + S_{n2}, \tag{A.15}$$

where

$$S_{n1} = -\int_0^\infty K\left(\frac{F(x) - t}{h_n}\right)\left(\frac{\hat{F}_n(x) - F(x)}{h_n}\right) dx,$$

$$S_{n2} = -\int_0^\infty \left[\int_{(F(x)-t)/h_n}^{(\hat{F}_n(x)-t)/h_n} \left(K(u) - K\left(\frac{F(x) - t}{h_n}\right)\right) du \right] dx.$$

From (A.1), write

$$S_{n1} = n^{-1} \sum_{i=1}^n V_i^{(n)} + r_n \tag{A.16}$$

with

$$V_i^{(n)} = -\int_0^\infty \varphi_i(x) h_n^{-1} K\Big(\frac{F(x) - t}{h_n}\Big) dx,$$

$$r_n = -h_n^{-1} \int_0^\infty K\Big(\frac{F(x) - t}{h_n}\Big) R_n(x) dx,$$

where φ_i and R_n are defined by (A.9) and (A.1), respectively. Standard calculations yield $\mathrm{E}V_i^{(n)} = 0$,

$$\sigma_n^2 \equiv \mathrm{Var}(V_i^{(n)}) = \frac{(1-t)^2}{f^2(Q(t))} \int_{a_W}^{Q(t)} \frac{dW_1(x)}{C^2(x)} + O(h_n) = \sigma^2(t) + O(h_n) \tag{A.17}$$

and $\mathrm{E}|V_i^{(n)}|^3 \le d_1 \frac{(1-t)^3}{f^3(Q(t))} \int_{a_W}^{Q(t)} \frac{dW_1(x)}{C^3(x)} < \infty$, where $d_1 > 0$ is a constant. Thus, by the Berry-Esseen Theorem (see Chow and Teicher (1978, p.299)), for all $n \ge 1$,

$$\sup_y \left| P\{n^{-\frac{1}{2}} \sum_{i=1}^n V_i^{(n)} < y\sigma_n\} - \Phi(y) \right| \le Mn^{-1/2}. \tag{A.18}$$

It follows from (A.14) that

$$\begin{aligned} &\sup_y |P\{n^{-1/2} \sum_{i=1}^n V_i^{(n)} < y\sigma(t)\} - \Phi(y)| \\ &\le \sup_y |P\{n^{-1/2} \sum_{i=1}^n V_i^{(n)} < y\sigma_n\} - \Phi(y)| + P\{|\sigma_n^{-1}\sigma(t) - 1| > a\} + a \end{aligned} \tag{A.19}$$

for any $a > 0$. It is easy to see that for any random variable V and $0 < \varepsilon < 1/2$, $P\{|V^{-1} - 1| > 2\varepsilon, V \ne 0\} \le 2P\{|V - 1| > \varepsilon\}$. Hence, from the fact that for $y, z > 0$ the relation $|y - 1| > z$ holds if $|\sqrt{y} - 1| > \sqrt{z}$, and from (A.17), we have

$$\begin{aligned} P\{|\sigma_n^{-1}\sigma(t) - 1| > a\} \le P\{|\sigma_n^{-2}\sigma^2(t) - 1| > a^2\} &\le 2P\{|\sigma_n^2\sigma^{-2}(t) - 1| > a^2/2\} \\ &\le 2P\{d_2 h_n > a^2/2\} \end{aligned} \tag{A.20}$$

for some constant $d_2 > 0$. In view of (A.18) – (A.20) and putting $a = (2d_2h_n)^{1/2}$, we get

$$\sup_y |P\{n^{-\frac{1}{2}} \sum_{i=1}^n V_i^{(n)} \le y\sigma(t)\} - \Phi(y)| \le M(n^{-1/2} + h_n^{1/2}). \tag{A.21}$$

When $a_G < a_W$, from (1.15) of Gijbels and Wang (1993) we obtain that for each $x > 0$ and $0 \le b < b_W$,

$$P\{\sup_{0 \le y \le b} n|R_n(y)| > x + 4\theta^{-2}\} \le M(\mathrm{e}^{-\lambda x} + (x/50)^{-2n} + \mathrm{e}^{-\lambda x^3}), \tag{A.22}$$

where $\theta > 0$ and $\lambda > 0$ are some constants. Note that for large n, there exists a constant b such that $0 \le Q(t + h_n u) \le b < b_W$ for all $|u| \le 1$. Hence for all n sufficiently large, it follows from (A.22) that

$$\begin{aligned} &P\{n^{\frac{1}{2}}|r_n| \ge n^{-1/2}h_n^{-1}\log n\} \\ &\quad = P\{\Big|\int_{-1}^1 \frac{1}{f(Q(t + h_n u))} K(u) R_n(Q(t + h_n u)) du\Big| \ge (nh_n)^{-1}\log n\} \\ &\quad \le P\{\sup_{0 \le y \le b} n|R_n(y)| \ge d_3 h_n^{-1}\log n\} \le Mn^{-1}, \end{aligned} \tag{A.23}$$

where $d_3 > 0$ is a constant. For some $\tau > 0$ with $1 + \tau \le (1 - t)/h_n$ (this holds if n is large), write

$$S_{n2} = \Delta_{n1} + \Delta_{n2} + \Delta_{n3}, \tag{A.24}$$

where $\Delta_{n1} = -h_n \int_{-1-\tau}^{1+\tau} \mathcal{K}(x)dx$, $\Delta_{n2} = -h_n \int_{1+\tau}^{(1-t)/h_n} \mathcal{K}(x)dx$ and $\Delta_{n3} = -h_n \int_{-t/h_n}^{-1-\tau} \mathcal{K}(x)dx$ with

$$\mathcal{K}(x) = \frac{1}{f(Q(t + h_n x))} \int_x^{(\hat{F}_n(Q(t+h_n x))-t)/h_n} (K(u) - K(x))du.$$

Since K has bounded support $[-1, 1]$,

$$\begin{aligned}
|\Delta_{n2}| &\le h_n \int_{1+\tau}^{(1-t)/h_n} \frac{1}{f(Q(t + h_n x))} I\left(\frac{\hat{F}_n(Q(t + h_n x)) - t}{h_n} < 1\right) \\
&\qquad \times \left| \int_x^{(\hat{F}_n(Q(t+h_n x))-t)/h_n} (K(u) - K(x))du \right| dx \\
&\le h_n \int_{1+\tau}^{(1-t)/h_n} \frac{1}{f(Q(t+h_n x))} I\left(\frac{\hat{F}_n(Q(t+(1+\tau)h_n))-(t+(1+\tau)h_n)}{h_n} < -\tau\right) \\
&\qquad \times \left| \int_x^{(\hat{F}_n(Q(t+h_n x))-t)/h_n} (K(u) - K(x))du \right| dx.
\end{aligned} \tag{A.25}$$

Using Theorem 1 of Zhu (1996), we have that for $a_G < a_W \le b < b_W$ and $\varepsilon > 0$,

$$P\{\sup_{a_W \le x \le b} |\hat{F}_n(x) - F(x)| > \varepsilon\} \le d_4 \exp(-n d_5 \varepsilon^2), \tag{A.26}$$

where d_4 and d_5 are absolute constants. Therefore, (A.25) and (A.26) imply that for all n sufficiently large,

$$\begin{aligned}
P\{n^{1/2}|\Delta_{n2}| \ge n^{-1/2}h_n^{-1}\log n\} &\le P\{|\hat{F}_n(Q(t+(1+\tau)h_n)) - (t+(1+\tau)h_n)| \ge \tau h_n\} \\
&\le d_4 \exp(-n d_5 \tau^2 h_n^2) \le M n^{-1}.
\end{aligned} \tag{A.27}$$

(Recall that $nh_n^2/\log n \to \infty$, so that $nd_5\tau^2 h_n^2 > \log n$ for large n.) Similarly,

$$P\{n^{1/2}|\Delta_{n3}| \ge n^{-1/2}h_n^{-1}\log n\} \le M n^{-1}. \tag{A.28}$$

Since K is Lipschitz continuous of order one, for some constant $d_6 > 0$, $|\Delta_{n1}| \le d_6 h_n^{-1} \sup_{|x|\le 1+\tau} \left|\hat{F}_n(Q(t + h_n x)) - (t + h_n x)\right|^2$. Hence, it follows from (A.26) that

$$\begin{aligned}
&P\{n^{1/2}|\Delta_{n1}| \ge n^{-1/2}h_n^{-1}\log n\} \\
&\le P\left\{\sup_{|x|\le 1+\tau} |\hat{F}_n(Q(t+h_n x)) - (t+h_n x)|^2 \ge d_6^{-1/2} n^{-1/2} (\log n)^{1/2}\right\} \le M n^{-1}.
\end{aligned} \tag{A.29}$$

Take $U = n^{-1/2}\sum_{i=1}^{n} V_i^{(n)}$, $V = n^{1/2}(r_n + S_{n2})$, and $a_n = n^{-1/2}h_n^{-1}\log n$ in (A.13), so that $n^{\frac{1}{2}}(\hat{Q}_n(t) - \bar{Q}_n(t)) = U + V$ by (A.15)–(A.16). It follows from (A.21), (A.23), (A.24) and (A.27)–(A.29) that $\sup_y |P\{n^{1/2}(\hat{Q}_n(t) - \bar{Q}(t)) \leq y\sigma(t)\} - \Phi(y)| \leq M(n^{-1/2}h_n^{-1}\log n + h_n^{1/2})$. Note that $|\Phi(x) - \Phi(y)| \leq |x - y|$, $-\infty < x, y < \infty$. Now, using (A.11), (A.12) and Theorem 3, we get

$$\begin{aligned}
&\sup_y \Big|P\{n^{\frac{1}{2}}(\hat{Q}_n(t) - Q(t)) \leq y\sigma(t)\} - \Phi(y)\Big| \\
&= \sup_y \Big|P\{n^{\frac{1}{2}}(\hat{Q}_n(t) - \bar{Q}_n(t)) \leq y'\sigma(t)\} - \Phi(y') + \Phi(y') - \Phi(y)\Big| \\
&\leq \sup_y |P\{n^{\frac{1}{2}}(\hat{Q}_n(t) - \bar{Q}_n(t)) \leq y\sigma(t)\} - \Phi(y)| + n^{\frac{1}{2}}|\bar{Q}_n(t) - Q(t)|/\sigma(t) \\
&\leq M(n^{-1/2}h_n^{-1}\log n + n^{1/2}h_n^r + h_n^{1/2}),
\end{aligned}$$

where $y' = y - n^{1/2}(\bar{Q}_n(t) - Q(t))/\sigma(t)$. This completes the proof of Theorem 2.

Acknowledgement

This work was partially supported by an RGC Earmarked Grant No. B-Q175 and a HKPOLYU Internal Research Grant No. G-S598. We are grateful to the Editor, Associate Editor, and an anonymous referee for their valuable comments and suggestions which led to the improvement of this paper.

References

Arcones, M. A. and Giné, E. (1995). On the law of the iterated logarithm for canonical U-statistics and processes. *Stochastic Process. Appl.* **58**, 217-245.

Burke, M. D., Csörgő, S. and Horváth, L. (1988). A correction to and improvement of 'Strong Approximations of some biometric estimates under random censorship'. *Probab. Theory Related Fields* **71**, 455-465.

Chang, M. N. and Rao, P. V. (1989). Berry-Esseen bound for the Kaplan-Meier estimator. *Commun. Statist. Theory Methods* **18**, 4647-4664.

Chow, Y. S. and Teicher, H. (1978). *Probab. Theory.* Springer-Verlag, New York.

Gijbels, I. and Wang, J. L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *J. Multivariate Anal.* **47**, 210-229.

Gu, M. G. (1995). Convergence of increments for cumulative hazard function in mixed censorship-truncation model with application to hazard estimators. *Statist. Probab. Lett.* **23**, 135-139.

Gu, M. G. and Lai, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18**, 160-189.

Gürler, U., Stute, W. and Wang, J. L. (1993). Weak and strong quantile representations for randomly truncated data with applications. *Statist. Probab. Lett.* **17**, 139-148.

Lai, T. L. and Ying, Z. (1991). Estimating a distribution with truncated and censored data. *Ann. Statist.* **19**, 417-442.

Lio, Y. L., Padgett, W. J. and Yu, K. F. (1986). On the asymptotic properties of a kernel type quantile estimator from censored samples. *J. Statist. Plann. Inference* **14**, 169-177.

Michel, R. and Pfanzagl, J. (1971). The accuracy of the normal approximation for minimum constrast estimates. *Z. Wahrsch. Verw. Gebiete.* **18**, 73-84.

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

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

Struthers, C. A. and Farewell, V. T. (1989). A mixture model for times to AIDS data with left truncation and an uncertain origin. *Biometrika.* **76**, 814-817.

Sun, L. (1997). Bandwidth choice for hazard rate estimators from left truncated and right censored data. *Statist. Probab. Lett.* **36**, 101-114.

Sun, L. and Zhou, Y. (1998). Sequential confidence bands for densities under truncated and censored data. *Statist. Probab. Lett.* **40**, 31-41.

Tsai, W. Y., Jewell, N. P. and Wang, M. C. (1987). A note on the product limit estimator under right censoring and left truncation. *Biometrika* **74**, 883-886.

Uzunoḡullari, U. and Wang, J. L. (1992). A comparison of hazard rate estimators for left truncated and right censored data. *Biometrika* **79**, 293-310.

Wang, M. C. (1987). Product limit estimates: a generalized maximum likelihood study. *Commun. Statist. Theory Methods.* **16**, 3117-3132.

Xiang, X. (1995a). Bahadur representation of the kernel quantile estimator under random censorship. *J. Multivariate Anal.* **54**, 193-209.

Xiang, X. (1995b). Deficiency of the sample quantile estimator with respect to kernel quantile estimators for censored data. *Ann. Statist.* **23**, 836-854.

Zhou, Y. (1996). A note on the TJW product-limit estimator for truncated and censored data. *Statist. Probab. Lett.* **26**, 381-387.

Zhu, Y. (1996). The exponential bound of the survival function estimator for randomly truncated and censored data. *Systems Sci. Math. Sci.* **9**, 175-181.

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

E-mail: maxzhou@polyu.edu.hk

Institute of Applied Mathematics, Academia Sinica, Beijing 100080, China.

Dept of Probability and Statistics, Peking University, Beijing 100871, China.

(Received February 1999; accepted March 2000)