ASYMPTOTIC REPRESENTATIONS FOR KERNEL DENSITY AND HAZARD FUNCTION ESTIMATORS WITH LEFT TRUNCATION

Yong Zhou

Academia Sinica, Beijing and The University of Hong Kong

Abstract: Kernel estimators of the density function and the hazard function based on the product-limit estimator are considered when the data are subject to left truncation. Several asymptotic uniformly strong and weak representations for these estimators are established. Making use of these results we obtain the large sample properties of the kernel density and hazard function estimators. The results can be extended to the case of left truncated and right censored data.

Key words and phrases: Density estimator, hazard function, product-limit estimator, truncated data.

1. Introduction

In many survival studies a subject may not be included in the study if the time origin of its lifetime, called the onset time, precedes the starting time of the study. Such subjects are called left truncated. In this paper we study kernel estimators of density and hazard functions based on the product-limit estimator (Lynden-Bell (1971)) when the data is subject to random truncation. Let $(X_i, Y_i), 1 \leq i \leq N$, be a sequence of independent identically distributed random vectors in the plane such that X_i is independent of Y_i. The marginal distribution functions of X_i and Y_i are given by $F(t) = P(X \leq t) \quad$ and $\quad G(t) = P(Y \leq t)$, respectively. Let f denote the density function of F.

In the random left truncation model one observes only those pairs (X_i, Y_i) for which $X_i \geq Y_i$ but nothing is observed otherwise. Left truncation is a frequent cause of incomplete data. It may occur if the time origin, X^0, of the lifetime precedes the time origin of the study, X^1. To be precise, when $Y' = X^1 - X^0 > X$, where X is the lifetime of interest, the case is not observed at all (we do not even know its existence). An important example of such a model arises in the analysis of survival data of patients infected by the AIDS virus from contaminated blood transfusions (Chen, Chao and Lo (1995) and Lagakos, Barraj and De Gruttola (1988)). A feature of HIV (AIDS) development is the induction period between infection with the AIDS virus and the onset of clinical AIDS. The data collected on persons infected by contaminated blood transfusions provide a unique source

of information on the induction period. Of persons infected in this way, only those who have developed AIDS can be identified. Let Y denote the chronological time of infection and X^* denote the induction period and assume they are independent. Suppose that one can only observe a random sample of patients who are infected and develop AIDS in some chronological time interval $[0, b]$. Let $X = b - X^*$. Then the pair (X, Y) is observed if and only if $0 \leq Y \leq X (\leq b)$.

Another example arises in astronomy. As briefly explained in Woodroofe (1985) and Chen, Chao and Lo (1995), the absolute and apparent luminosities of an astronomical object are defined in terms of its brightness at a fixed distance from Earth. The magnitude is defined to be the negative logarithm of luminosity. Since one can only observe those objects which are bright enough the apparent magnitude must be small enough, say less than or equal to some constant a. It is well accepted in cosmology that the apparent magnitude can be expressed as the sum of a function of redshift, denoted by X', and the absolute magnitude, denoted by Y', and that the redshift and the absolute magnitude may be assumed to be independent. If we let $X = -X'$ and $Y = Y' - a$, the condition for the observability of a celestial object becomes $X \geq Y$. This yields the model described above.

Many other important examples can be found in econometrics (Tobin (1958), Amemiya (1985) and Tsui, Jewell and Wu (1988)) and in astronomy (Lynden-Bell (1971) and Jackson (1974)).

We regard the observed samples $X_i, Y_i, i = 1, \ldots, n$ as being generated by a large sample of independent random variables $X_j, Y_j, j = 1, \ldots, N$, where $n = n(N)$ is given by $\sum_{j=1}^{N} I\{X_j \geq Y_j\}$. We assume that $\alpha = P(Y \leq X) > 0$. Now, given n, we may consider the observed data to be the outcome of an i.i.d. sample with distribution function

$$H^*(x, y) = P(X \leq x, Y \leq y | Y \leq X) = \alpha^{-1} \int_{-\infty}^{x} G(y \wedge z) dF(z),$$

where $a \wedge b = \min(a, b) (a \vee b = \max(a, b)$ below). It follows from the SLLN that $n/N \to \alpha$ $a.s.$ as $N \to \infty$. Let

$$F^*(x) = H^*(x, \infty) = \alpha^{-1} \int_{-\infty}^{x} G(z) dF(z)$$

and

$$G^*(y) = H^*(\infty, y) = \alpha^{-1} \int_{-\infty}^{\infty} G(y \wedge z) dF(z)$$

be the marginals of H^*. The observed X's and Y's thus have distribution functions F^* and G^*, respectively. For simplicity of presentation, we assume that X and Y are nonnegative random variables.

The quantity $\Lambda(x) = -\log(1 - F(t))$ is the cumulative hazard function. It plays an important role in the truncation model. For any distribution function L, take the left and right endpoints of its support to be $a_L = \inf\{x : L(x) > 0\}$ and $T_L = \sup\{x : L(x) < 1\}$.

Woodroofe (1985) pointed out that F can be estimated only if $a_G \leq a_F$ and showed that

$$\Lambda(x) = \int_{a_F}^x \frac{dF^*(z)}{C(z)}, \ a_F < x < T_F,$$

with $C(z) = G^*(z) - F^*(z) = \alpha^{-1}G(z)(1 - F(z)), \ a_F < x < T_F$. Note that $C(z)$ is consistently estimated by $C_n(z) = n^{-1}\sum_{i=1}^n I(Y_i \leq z \leq X_i)$, while the above representation of Λ in terms of F^* and C suggests estimation of Λ by

$$\Lambda_n(x) = \int_{a_F}^x \frac{dF_n^*(z)}{C_n(z)} = n^{-1}\sum_{i:X_i\leq x} C_n^{-1}(X_i).$$

The nonparametric maximum likelihood estimator (MLE) of F was derived by Lynden-Bell (1971) and is given by

$$1 - \hat{F}_n(x) = \prod_{\{i:X_i\leq x\}}' \left[1 - \frac{r_n(X_i)}{nC_n(X_i)}\right], \tag{1.1}$$

where $r_n(x) = \sum_{i=1}^n I(j \leq n : X_j = x), C_n(x) = n^{-1}\sum_{j=1}^n I(Y_j \leq x \leq X_j)$ and $\prod'$ extends over all pointwise distinct $X_1, \ldots, X_n$. Nicoll and Segal (1980) derived MLEs for grouped data; and Bhattacharya, Chernoff and Yang (1983) derived MLEs from a conditional likelihood function of certain counts, given the observed X-values. Bhattacharya, Chernoff and Yang (1983) constructed nonparametric estimators of regression parameters in linear models, and showed the asymptotic normality of the estimation error, properly normalized. Bhattacharya (1983) also considered the asymptotic distribution of a goodness-of-fit statistic with a view towards testing hypotheses about regression parameters.

In the left truncation model, $\hat{F}_n(x)$ has been generally accepted as a substitute for the empirical distribution function. Woodroofe (1985) proved the weak convergence of $\widehat{F}_n(t)$ to a certain Gaussian process in the space $D[a, b]$ and the convergence of $\widehat{F}_n(t)$ in probability under the integral condition $\int_{a_F}^\infty dF/G < \infty$. It should be noticed that the only condition for the uniform weak consistency of the PL-estimator $\widehat{F}_n$ over $[a_F, \infty)$ to hold is that $a_F \geq a_G$. Strong uniform consistency has been proved by Wang, Jewell and Tsai (1986) and Wellek (1990) in case of $a_F > a_G$. In the case $a_G = a_F$, Woodroofe (1985) and Keiding and Gill (1990) obtained the weak uniform consistency of the PL-estimator $\widehat{F}_n(x)$. Chen Chao and Lo (1995) also considered the strong uniform consistency of $\widehat{F}_n(t)$ in the case $a_F = a_G$ without any integral conditions.

Chao and Lo (1988) and Stute (1993) derived an almost sure representation of $\hat{F}_n$ under the different integral conditions $\int_{a_F}^{\infty} dF/G < \infty$ and $\int_{a_F}^{\infty} dF/G^2 < \infty$, respectively.

In fact, the left truncation model can be extended to the left truncation and right censorship model (LTRC). Gijbels and Wang (1993) and Zhou (1996) obtained a strong i.i.d. representation for the product-limit estimator (Tsai, Jewell and Wang (1987)), an extension of Lynden-Bell product-limit estimator, under different conditions on the support of the distributions. Hence the results of this paper can be extended to the left truncation and right censorship model.

In this paper, we write

$$\int_s^t f(x)dg(x) = \int_{(s,t]} f(x)dg(x) \text{ and } \int_a^t f(x)dg(x) = 0, \text{ if } t \le a.$$

We can construct a kernel estimator of f based on $\hat{F}_n(x)$. Thus

$$f_n(t) = a_n^{-1} \int_{a_F}^{\infty} k\left(\frac{t-x}{a_n}\right) d\hat{F}_n(x), \qquad (1.2)$$

where $\{a_n, n \ge 1\}$ is a sequence of bandwidth tending to zero at appropriate rates and k is a smooth probability kernel. The hazard rate function $\lambda(t) = \Lambda'(t)$ is defined by

$$\lambda(t) = \frac{f(t)}{1 - F(t)}, \text{ for } F(t) < 1. \qquad (1.3)$$

An estimate for the hazard function $\lambda(t)$ for an i.i.d. sample $X_1, \ldots, X_n$, at risk of being truncated from the left, is defined by

$$\lambda_n(t) = a_n^{-1} \int_{a_F}^{\infty} k\left(\frac{t-x}{a_n}\right) d\Lambda_n(x). \qquad (1.4)$$

We establish convergence properties of the kernel estimators $f_n(x)$ and $\lambda_n(x)$. To formulate our results, let

$$\bar{f}_n(t) = a_n^{-1} \int_{a_F}^{\infty} k\left(\frac{t-x}{a_n}\right) dF(x), \Delta_n(t) = f_n(t) - \bar{f}_n(t) - \frac{\alpha[f_n^*(t) - Ef_n^*(t)]}{G(t)},$$

$$\bar{\lambda}_n(t) = a_n^{-1} \int_{a_F}^{\infty} k\left(\frac{t-x}{a_n}\right) d\Lambda(x), \ \Theta_n(t) = \lambda_n(t) - \bar{\lambda}_n(t) - \frac{f_n^*(t) - Ef_n^*(t)}{C(t)},$$

$$f_n^*(t) = a_n^{-1} \int_{a_F}^{\infty} k\left(\frac{t-x}{a_n}\right) dF_n^*(x),$$

where $F_n^*(x)$ is the empirical distribution function of $X_1, \ldots, X_n$ and $EF_n^*(x) = F^*(x)$.

To prove the consistency of the kernel density and hazard function estimators, we only consider the situation $a_G \le a_F$.

Theorem 1.1. *Let $a_G < a_F$. Suppose k is a bounded variation probability kernel vanishing outside some finite interval $-\infty < r < 0 < s < \infty$. Let $f = F'$ and $g = G'$ be bounded on $[a_F, T]$ for some $a_F < T < T_F$. Then*

$$\sup_{a_F < t \leq T} |\Delta_n(t)| = O_p\Big(\frac{\log n}{n a_n}\Big) + O_p(n^{-\frac{1}{2}}),$$

$$\sup_{a_F < t \leq T} |\Delta_n(t)| = O\Big(\frac{\log n}{n a_n}\Big) + O((n^{-1} \log \log n)^{\frac{1}{2}}) \quad a.s.$$

Theorem 1.2. *Under the assumptions of Theorem* 1.1.

$$\sup_{a_F < t \leq T} |\Theta_n(t)| = O_p\Big(\frac{\log n}{n a_n}\Big) + O_p(n^{-\frac{1}{2}}),$$

$$\sup_{a_F < t \leq T} |\Theta_n(t)| = O\Big(\frac{\log n}{n a_n}\Big) + O((n^{-1} \log \log n)^{\frac{1}{2}}) \quad a.s.$$

Stute (1993) found asymptotic representations of $\Lambda_n(x) - \Lambda(x)$ and $\hat{F}_n(x) - F(x)$ under the condition (apart from $a_G \leq a_F$)

$$\int_{a_F}^{\infty} \frac{dF(z)}{G^2(z)} < \infty. \tag{1.5}$$

Obviously, when $a_G < a_F$, (1.5) is satisfied. The more interesting situation is $a_F = a_G = a$, that is, $F(a) = G(a) = 0$ but $F, G > 0$ on (a, ∞). In this case we obtain weaker rates of convergence for kernel density and hazard function estimators than those of Theorem 1.1 and Theorem 1.2. Note that rates are derived under (1.5) and the condition

$$\int_{a_F}^{\infty} \frac{dF(z)}{G^p(z)} < \infty \tag{1.6}$$

for $p \geq 2$.

Theorem 1.3. *Suppose that the conditions of Theorem* 1.1 *hold. If $a_G = a_F$ and (1.6) is satisfied, then for any $\varepsilon > 0$ and $\delta > 0$,*

$$\sup_{a_F < t \leq T} |G(t)\Delta_n(t)| = \begin{cases} o\Big(a_n^{-\frac{1}{2p}}\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) + O\Big(\frac{(\log n)^{1+\delta}}{n a_n}\Big) & a.s. \text{ if } p \geq 3, \\ o\Big(a_n^{-1}\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) + O\Big(\frac{(\log n)^{1+\delta}}{n a_n}\Big) & a.s. \text{ if } 2 \leq p < 3, \end{cases}$$

and

$$\sup_{a_F < t \leq T} |C(t)\Theta_n(t)| = \begin{cases} o\Big(a_n^{-\frac{1}{2p}}\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) + O\Big(\frac{(\log n)^{1+\delta}}{n a_n}\Big) & a.s. \text{ if } p \geq 3, \\ o\Big(a_n^{-1}\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) + O\Big(\frac{(\log n)^{1+\delta}}{n a_n}\Big) & a.s. \text{ if } 2 \leq p < 3. \end{cases}$$

Remark 1.1. If "$\sup_{a_F < t < T}$" in Theorem 1.3 is replaced by "$\sup_{c \le t \le T}$", for some constant $c > a_F$, then the orders of convergence in the results of Theorem 1.3 are the same as those in Theorems 1.1 and 1.2 under the integral condition (1.5).

Using Theorem 1.1 and Theorem 1.3, we can derive many of the asymptotic properties of the kernel density estimator $f_n(x)$. Moreover, we can obtain similar results for the hazard function estimator λ_n from Theorems 1.2 and 1.3. We need the following assumptions for the bandwidth $\{a_n, n \ge 1\}$.

Assumptions. Let $\{a_n, n \ge 1\}$ be a sequence of positive constants such that

(i) $\dfrac{(\log n)^2}{n a_n \log \log n} \to 0$,

(ii) $\dfrac{\log a_n^{-1}}{n a_n} \to 0$, $\dfrac{\log a_n^{-1}}{\log \log n} \to \infty$, and $\dfrac{(\log n)^2}{n a_n \log a_n^{-1}} \to 0$,

(iii) $\dfrac{a_n^{\frac{p-1}{p}} \log^{1+\varepsilon} n}{\log \log n} \to 0$ and $\dfrac{(\log n)^{3+\varepsilon}}{n a_n \log \log n} \to 0$ for any $\varepsilon > 0$ and some $p \ge 3$,

(iv) $\dfrac{a_n^{\frac{p-1}{p}} \log^{1+\varepsilon} n}{\log a_n^{-1}} \to 0$ and $\dfrac{(\log n)^{3+\varepsilon}}{n a_n \log a_n^{-1}} \to 0$ for any $\varepsilon > 0$ and some $p \ge 3$.

As a first application of the theorems, we determine the pointwise almost sure convergence rate.

Corollary 1. *Assume that $a_n \to 0$ and $n a_n \to \infty$ in such a way that*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup_m \left| \frac{a_m}{a_n} - 1 \right| = 0,$$

where the supremum is taken over values of m with $|m - n| \le n\varepsilon$, and (a) if $a_G < a_F$ and $\{a_n, n \ge 1\}$ satisfies assumption (i), or (b) if $a_G = a_F$, (1.6) is satisfied for $p \ge 3$ and $\{a_n, n \ge 1\}$ satisfies assumption (iii), we have

$$\limsup_{n \to \infty} \pm \left(\frac{n a_n}{2 \log \log n} \right)^{\frac{1}{2}} (f_n(t) - \bar{f}_n(t)) = \left[\frac{\alpha f(t)}{G(t)} \int k^2(y) dy \right]^{\frac{1}{2}} \quad a.s.$$

Proof. After Theorem 1.1 and Theorem 1.3, we obtain Corollary 1 by Theorem 2 of Hall (1981).

Corollary 2. *Assume that on $[T_1, T_2]$, $a_F < T_1 < T_2 < T$, we have $f \ge m > 0$. Let $a_n \downarrow 0$ and $n a_n \uparrow \infty$. (a) If $a_G < a_F$, and $\{a_n, n \ge 1\}$ satisfies assumption (ii), or (b) if $a_G = a_F$, (1.6) is satisfied for $p \ge 3$ and $\{a_n, n \ge 1\}$ satisfies assumption (ii) and (iv), we have*

$$\lim_{n \to \infty} \left(\frac{n a_n}{2 \log a_n^{-1}} \right)^{\frac{1}{2}} \sup_{T_1 \le t \le T_2} \left(\frac{G(t)}{f(t)} \right)^{\frac{1}{2}} |f_n(t) - \bar{f}_n(t)| = \left[\alpha \int k^2(y) dy \right]^{\frac{1}{2}} \quad a.s.$$

The result is an easy consequence of Theorem 1.1, Theorem 1.2 and Theorem 1.3 of Stute (1982).

2. Proofs of Theorems

We prove only Theorem 1.1 and the first result of Theorem 1.3 for the kernel density estimator f_n, since the arguments are similar for the hazard function estimator λ_n.

Lemma 2.1.
(a) *Assume $a_G < a_F$. Then, uniformly in $a_F < x \leq T < T_F$,*

$$\Lambda_n(x) - \Lambda(x) = L_n(x) + R_n(x), \tag{2.1}$$
$$F_n(x) - F(x) = (1 - F(x))L_n(x) + R_n^0(x), \tag{2.2}$$

with $\sup_{a_F < x \leq T} |R_n(x)| = \sup_{a_F < x \leq T} |R_n^0(x)| = O(n^{-1} \log n)$ a.s., where

$$L_n(x) = \int_{a_F}^x C^{-1}(z)d[F_n^*(z) - F^*(z)] - \int_{a_F}^x \frac{C_n(x) - C(x)}{C^2(x)}dF^*(x).$$

(b) *If $a_G = a_F = 0$, and (1.5) hold, then (2.1) and (2.2) hold with*

$$\sup_{a_F < x \leq T} |R_n(x)| = \sup_{a_F < x \leq T} |R_n^0(x)| = o(n^{-1}(\log n)^\delta) \text{ a.s. for } \delta > 3/2.$$

Proof. Assertion (a) follows from (b) and (c) of Theorem 1 of Gijbels and Wang (1993). The result (b) follows from Theorem 1 of Stute (1993) and Theorem 2 of Zhou (1996).

Remark 2.1. According to results of Zhou (1996) and Gijbels and Wang (1993), Lemma 2.1 still holds when the left truncation model is extended to the left truncation and right censorship model. Hence corresponding results for truncated and censored data hold.

Proof of Theorem 1.1 and Theorem 1.3. By (2.2) and integration by parts, we have

$$f_n(x) - \bar{f}_n(x) = n^{-1} \int_0^\infty k\left(\frac{x - t}{a_n}\right)d[F_n(t) - F(t)]$$
$$= a_n^{-1} \int_0^\infty (1 - F(t)) \int_{a_F}^t C^{-1}(z)d[F_n^*(z) - F^*(z)]dk\left(\frac{x - t}{a_n}\right)$$
$$-a_n^{-1} \int_0^\infty (1 - F(t)) \int_{a_F}^t \frac{C_n(z) - C(z)}{C^2(z)}dF^*(z)dk\left(\frac{x - t}{a_n}\right)$$
$$+a_n^{-1} \int_0^\infty R_n^0(t)dk\left(\frac{x - t}{a_n}\right)$$
$$= I_1 + I_2 + I_3. \tag{2.3}$$

When $a_G < a_F$, put $u = (t - x)/a_n$. It follows from a change of variables that for sufficiently large n

$$|I_2| \leq a_n^{-1} \Big| \int_r^s [1 - F(x - a_n u)] \int_x^{x - a_n u} \frac{C_n(z) - C(z)}{C^2(z)} dF^*(z) dk(u) \Big|$$

$$+ a_n^{-1} \Big| \int_{a_F}^x \frac{C_n(z) - C(z)}{C^2(z)} dF^*(z) \Big| \Big| \int_r^s [1 - F(x - a_n u)] dk(u) \Big|$$

$$\leq \sup_{0 \leq x \leq T'} |C_n(x) - C(x)| \sup_{0 < x < T'} f(x) \Big[\frac{\alpha}{G(a)(1 - F(T'))^2} \int |dk(u)| + M \Big],$$

uniformly on $a_F < x \leq T$, where $\alpha = P(X \geq Y)$, $M = |r| \vee s$ and $a_G < a \leq a_F$, $T \leq T' < T_F$.

Thus we have

$$\sup_{a_F < x \leq T} |I_2| = \begin{cases} O_p(n^{-\frac{1}{2}}), \\ O((n^{-1} \log \log n)^{\frac{1}{2}}) \quad \text{a.s.} \end{cases} \tag{2.4}$$

From (2.2),

$$\sup_{a_F < x \leq T} |I_3| \leq a_n^{-1} \sup_{a \leq x \leq T'} |R_n^0(x)| \int \Big| dk(u) \Big| = O\Big(\frac{\log n}{n a_n} \Big) \quad \text{a.s.} \tag{2.5}$$

where $a_G < a < a_F$ and k is of bounded variation.

For I_1, integration by parts yields

$$I_1 = a_n^{-1} \int_0^\infty [1 - F(t)] \frac{F_n^*(t) - F^*(t)}{C(t)} dk\Big(\frac{x - t}{a_n} \Big)$$

$$- a_n^{-1} \int_0^\infty [1 - F(t)] \int_{a_F}^t \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) dk\Big(\frac{x - t}{a_n} \Big)$$

$$= I_1' + I_2'.$$

Since $a_G < a_F$, $G(a) > 0$ if $a_G < a \leq a_F$. Put $u = (x - t)/a_n$. For sufficiently large n we have

$$|I_2'| \leq a_n^{-1} \Big| \int_r^s [1 - F(x - a_n u)] \int_x^{x - a_n u} \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) dk(u) \Big|$$

$$+ a_n^{-1} \Big| \int_{a_F}^x \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) \Big| \Big| \int_r^s [1 - F(x - a_n u)] dk(u) \Big|$$

$$\leq \sup_{0 \leq x \leq T'} |F_n^*(x) - F^*(x)| \Big\{ \frac{\alpha^2 M}{G^2(a)[1 - F(T')]^2} \sup_{0 \leq x \leq T'} |C'(x)| \int |dk(u)|$$

$$+ \Big| \frac{1}{C(T)} - \frac{1}{C(a_F)} \Big| \sup_{0 \leq x \leq T'} f(x) \Big\}$$

$$= \begin{cases} O_p(n^{-\frac{1}{2}}) \\ O((n^{-1} \log \log n)^{\frac{1}{2}}) \quad \text{a.s.} \end{cases} \tag{2.6}$$

uniformly in $a_F < x \le T$, where $\alpha = P(Y \le X) > 0$, $M = |r| \vee s$ and $a_G < a \le a_F$, $T \le T' < T_F$.

Another integration by parts, gives

$$
\begin{aligned}
I_1' &= \alpha a_n^{-1} \int_0^\infty \frac{F_n^*(t) - F^*(t)}{G(t)} dk\Big(\frac{x-t}{a_n}\Big) \\
&= \frac{\alpha}{a_n G(x)} \int_0^\infty [F_n^*(t) - F^*(t)] dk\Big(\frac{x-t}{a_n}\Big) \\
&\quad + \frac{\alpha}{a_n G(x)} \int_0^\infty \frac{1}{G(t)} [F_n^*(t) - F^*(t)][G(x) - G(t)] dk\Big(\frac{x-t}{a_n}\Big) \\
&= \frac{\alpha}{G(x)} [f_n^*(x) - E f_n^*(x)] + I_1''.
\end{aligned}
$$

It follows from the results for empirical processes that

$$
\begin{aligned}
\sup_{a_F < x \le T} |I_1''| &\le \frac{\alpha}{a_n G^2(a)} \sup_{0 \le x \le T'} |F_n^*(x) - F^*(x)| \int_r^s |G(x) - G(x - a_n u)| |dk(u)| \\
&= \begin{cases} O_p(n^{-\frac{1}{2}}) \\ O((n^{-1} \log\log n)^{\frac{1}{2}}) \quad \text{a.s.} \end{cases}
\end{aligned}
\tag{2.7}
$$

Hence,

$$
\sup_{a_F < x \le T} \Big| I_1' - \frac{\alpha}{G(x)} [f^*(x) - E f_n^*(x)] \Big| = \begin{cases} O_p(n^{-\frac{1}{2}}), \\ O((n^{-1} \log\log n)^{\frac{1}{2}}) \quad \text{a.s.} \end{cases}
\tag{2.8}
$$

Theorem 1.1 follows from (2.3), (2.4), (2.5) and (2.8).

When $a_G = a_F$, the proofs are more difficult. We assume $a_G = a_F = 0$ without loss of generality. The following lemma is needed in the proof of Theorem 1.3.

Lemma 2.2. *Assume $a_G = a_F = 0$ and (1.6). Then for any $\varepsilon > 0$ and $p \ge 2$,*

$$
\sup_{0 < x \le T} \sup_{r \le u \le s} \Big| \int_x^{x - a_n u} \frac{C_n(z) - C(z)}{C^2(z)} dF^*(z) \Big| = o\Big(a_n^{\frac{2p-1}{2p}} \Big(\frac{\log^{1+\varepsilon} n}{n} \Big)^{\frac{1}{2}} \Big) \quad a.s.
\tag{2.9}
$$

If "$\sup_{0 < x \le T}$" in (2.9) is replaced by "$\sup_{c \le x \le T}$" for some given $c > 0$, then the right side of (2.9) is $o(n^{-1} a_n^2 \log^{1+\varepsilon} n)^{1/2}$ under the integral condition (1.5).

Proof. By the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
&\sup_{0 < x \le T} \sup_{r \le u \le s} \Big| \int_x^{x - a_n u} \frac{C_n(z) - C(z)}{C^2(z)} dF^*(z) \Big| \\
&\le \sup_{0 < x \le T'} \Big| \frac{C_n(x) - C(x)}{(C(x))^{\frac{1}{2}}} \Big| \sup_{0 < x \le T} \sup_{r \le u \le s} \Big| \int_x^{x - a_n u} \frac{dF^*(z)}{C^{\frac{3}{2}}(z)} \Big|
\end{aligned}
$$

$$\leq M \sup_{0 < x \leq T'} \left| \frac{C_n(x) - C(x)}{(C(x))^{\frac{1}{2}}} \right| \cdot \sup_{0 < x \leq T} \sup_{r \leq u \leq s} \left| \int_x^{x - a_n u} \frac{dF}{G^p} \right|^{\frac{1}{2p}} \left| \int_x^{x - a_n u} dF \right|^{\frac{2p-1}{2p}}$$

$$\leq M a_n^{\frac{2p-1}{2p}} \sup_{0 < x \leq T'} \left| \frac{C_n(x) - C(x)}{(C(x))^{\frac{1}{2}}} \right|, \tag{2.10}$$

where M is an absolute constant and $T \leq T' < T_F$.

On the other hand

$$\sup_{0 < x \leq T'} \left| \frac{C_n(x) - C(x)}{(C(x))^{\frac{1}{2}}} \right| \leq \alpha^{\frac{1}{2}} [1 - F(T')]^{-\frac{1}{2}} \left\{ \sup_{0 < x \leq T'} \left| \frac{G_n^*(x) - G^*(x)}{(G^*(x))^{\frac{1}{2}}} \right| \left(\frac{G^*(x)}{G(x)} \right)^{\frac{1}{2}} \right.$$

$$\left. + \sup_{0 < x \leq T'} \left| \frac{F_n^*(x) - F^*(x)}{(F^*(x))^{\frac{1}{2}}} \right| \left(\frac{F^*(x)}{G(x)} \right)^{\frac{1}{2}} \right\}.$$

On $0 < x \leq T'$ we have, by (1.6),

$$\infty > \int_0^{T'} \frac{1}{G^p(z)} dF(z) \geq \int_0^x \frac{1}{G^p(z)} dF(z) \geq \frac{\alpha F^*(x)}{G^{p+1}(x)}.$$

Then for some constant $c > 0$,

$$F^*(x) \leq c G^{p+1}(x). \tag{2.11}$$

Hence

$$\left(\frac{F^*(x)}{G(x)} \right)^{\frac{1}{2}} \leq c^{\frac{1}{2}} G^{\frac{p}{2}}(x) < \infty \tag{2.12}$$

uniformly in $0 < x \leq T'$, and

$$\left(\frac{G^*(x)}{G(x)} \right)^{\frac{1}{2}} \leq \left(\frac{1}{\alpha G(x)} \int_0^\infty G(x \wedge y) dF(y) \right)^{\frac{1}{2}}$$

$$\leq \left(\frac{1}{\alpha} \int_0^\infty dF(y) \right)^{\frac{1}{2}} < \infty \tag{2.13}$$

uniformly in $0 < x \leq T'$.

Furthermore, from Csáki (1975), we have for any $\varepsilon > 0$,

$$\sup_{x > 0} \left| \frac{L(x) - L(x)}{(L(x))^{\frac{1}{2}}} \right| = o\left(\left(\frac{\log^{1+\varepsilon} n}{n} \right)^{\frac{1}{2}} \right) \quad \text{a.s.} \tag{2.14}$$

for $L(x) = F^*(x)$ and $L(x) = G(x)$, respectively. Thus, from (2.11)-(2.14), we obtain

$$\sup_{0 < x \leq T'} \left| \frac{C_n(x) - C(x)}{(C(x))^{\frac{1}{2}}} \right| = o\left(\left(\frac{\log^{1+\varepsilon} n}{n} \right)^{\frac{1}{2}} \right) \quad \text{a.s.} \tag{2.15}$$

From (2.10) and (2.15), the lemma follows.

Now we proceed to prove Theorem 1.3. Assume that $a_G = a_F = 0$. Here the representations of I_1, I_2 and I_3 are the same as those in the proof of Theorem 1.1. Hence, it follows from Lemma 2.2 and (2.15) that

$$
\begin{aligned}
|I_2| &\le a_n^{-1} \sup_{0<x\le T'} \sup_{r\le u\le s} \Big| \int_x^{x-a_n u} \frac{C_n(z) - C(z)}{C^2(z)} dF^*(z) \Big| \int_r^s |dk(u)| \\
&\quad + \sup_{0<x\le T'} \Big| \frac{C_n(x) - C(x)}{(C(x))^{\frac12}} \Big| \int_0^{T'} \frac{dF^*(z)}{C^{\frac32}(z)} \sup_{0<x\le T'} f(x) \\
&= o\Big(a_n^{-\frac{1}{2p}} \Big(\frac{\log^{1+\varepsilon} n}{n} \Big)^{\frac12} \Big) \quad \text{a.s.}
\end{aligned}
$$

From (2.14), we have

$$
\begin{aligned}
&\sup_{0<x\le T} \sup_{r\le u\le s} \Big| \int_x^{x-a_n u} \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) \Big| \\
&\le \sup_{x>0} \Big| \frac{F_n^*(x) - F^*(x)}{(F^*(x))^{\frac12}} \Big| \sup_{0<x\le T} \sup_{r\le u\le s} \Big| \int_x^{x-a_n u} \frac{(F^*(z))^{\frac12}}{C^2(z)} dC(z) \Big| \\
&\le M a_n \sup_{x>0} \Big| \frac{F_n^*(x) - F^*(x)}{(F^*(x))^{\frac12}} \Big| \sup_{0<x\le T} |C'(x)| C^{\frac{p-3}{2}}(x),
\end{aligned}
$$

where M is a constant which has different values in each appearance.

By (2.14) and the boundaries of f and g, we have

$$
\begin{aligned}
&\sup_{0<x\le T} \sup_{r\le u\le s} \Big| \int_x^{x-a_n u} \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) \Big| \\
&= \begin{cases} o\Big(a_n \Big(\dfrac{\log^{1+\varepsilon} n}{n} \Big)^{\frac12} \Big) & \text{a.s. for } p \ge 3, \\[3mm] o\Big(\Big(\dfrac{\log^{1+\varepsilon} n}{n} \Big)^{\frac12} \Big) & \text{a.s.} \quad \text{for } 2 \le p < 3. \end{cases}
\end{aligned}
$$

For $|I_2'|$, we have that uniformly in $0 < x \le T$,

$$
\begin{aligned}
|I_2'| &\le a_n^{-1} \sup_{0<x\le T} \sup_{r\le u\le s} \Big| \int_x^{x-a_n u} \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) \Big| \cdot \int_r^s |dk(u)| \\
&\quad + M a_n^{-1} \sup_{0<x\le T} \Big| \int_0^x \frac{F_n^*(z) - F^*(z)}{C^2(z)} dC(z) \Big| \sup_{0<x\le T'} f(x) \\
&= \begin{cases} o\Big(\Big(\dfrac{\log^{1+\varepsilon} n}{n} \Big)^{\frac12} \Big) & \text{a.s.} \quad \text{for } p \ge 3, \\[3mm] o\Big(a_n^{-1} \Big(\dfrac{\log^{1+\varepsilon} n}{n} \Big)^{\frac12} \Big) & \text{a.s. for } 2 \le p < 3. \end{cases}
\end{aligned}
\tag{2.16}
$$

Finally, we get

$$G(x)I_1' = \alpha a_n^{-1} \int_0^\infty k\Big(\frac{x-t}{a_n}\Big) d[F_n^*(t) - F^*(t)]$$

$$+ \alpha a_n^{-1} \int_0^\infty \frac{1}{G(t)} [F_n^*(t) - F^*(t)][G(x) - G(t)] dk\Big(\frac{x-t}{a_n}\Big).$$

From (2.14), if $0 < x \leq T$,

$$\alpha a_n^{-1} \Big| \int_0^\infty \frac{1}{G(t)} [F_n^*(t) - F^*(t)][G(x) - G(t)] dk\Big(\frac{x-t}{a_n}\Big) \Big|$$

$$\leq \alpha a_n^{-1} \sup_{x>0} \frac{|F_n^*(x) - F^*(x)|}{(F^*(x))^{\frac{1}{2}}} \int_0^\infty G^{\frac{p-1}{2}}(t) |G(x) - G(x - a_n u)| |dk'(u)|$$

$$= o\Big(\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) \quad \text{a.s.}$$

Then

$$\sup_{0<x\leq T} \Big| G(x)I_1' - \alpha[f_n^*(x) - Ef^*(x)] \Big| = o\Big(\Big(\frac{\log^{1+\varepsilon} n}{n}\Big)^{\frac{1}{2}}\Big) \quad \text{a.s.} \qquad (2.17)$$

It follows from (2.5) that for any $\delta > 1/2$,

$$\sup_{0<x\leq T} G(x)|I_3| = O\Big(\frac{(\log n)^{1+\delta}}{na_n}\Big) \quad \text{a.s.} \qquad (2.18)$$

Theorem 1.3 follows from (2.16), (2.17) and (2.18).

Acknowledgement

References

Amemiya, J. (1985). *Advanced Econometrics*. Harvard Univ. Press, Cambridge, Mass.

Bhattacharya, P. K. (1983). Justification for a K-S type test for the slope of a truncated regression. *Ann. Statist.* **11**, 697-701.

Bhattacharya, P. K., Chernoff, H. and Yang, S. S. (1983). Nonparametric estimation of the slope of a truncated regression. *Ann. Statist.* **11**, 505-514.

Chao, M. T. and Lo, S. H. (1988). Some representations of the nonparametric maximum likelihood estimators with truncated data. *Ann. Statist.* **16**, 661-668.

Chen, K., Chao, M. T. and Lo, S. H. (1995). On strong uniform consistency of the Lynden-Bell estimator for truncation data. *Ann. Statist.* **23**, 440-449

Csáki, E. (1975). Some notes on the law of the iterated logarithm for empirical distribution function. In *Colloq. Math. Soc. János Bolyai II Limit Theorems of Probability Theory* (Edited by P. Révész), 47-57. North-Holland, Amsterdam.

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

Hall, P. (1981). Laws of the iterated logarithm for nonparametric density estimators. *Z. Wahrsch. Verw. Gebieta* **56**, 47-61.

Jackson, J. C. (1974). The analysis of quasar samples. *Monthly Notices Roy. Astronom. Soc.* **166**, 281-295.

Keiding, N. and Gill, R. D. (1990). Random truncated models and Markov process. *Ann. Statist.* **18**, 582-602.

Lagakos, S. W., Barraj, L. M. and De Gruttola, V. (1988). Nonparametric analysis of truncated survival data with application to AIDS. *Biometrika* **75**, 515-523.

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronom. Soc.* **155**, 95-118.

Nicoll, J. F. and Segal, I. E. (1980). Nonparametric elimination of the observational cutoff bias. *Astronom. and Astrophys.* **82**, L3-L6.

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

Stute, W.(1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21**, 146-156.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica* **26**, 24-36.

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.

Tsui, K. L., Jewell, N. P. and Wu, C. F. J. (1988). A nonparametric approach to the truncated regression problem. *J. Amer. Statist. Assoc.* **83**, 785-792.

Wang, M. C., Jewell, N. P. and Tsai, W. Y. (1986). Asymptotic properties of the product-limit estimate under random truncation. *Ann. Statist.* **4**, 1597-1605.

Wellek, S. (1990). A nonparametric model for product-limit estimation under right censoring and left truncation. *Comm. Statist. Models* **6**, 561-592.

Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13**, 163-177.

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

Institute of Applied Mathematics, Academia Sinica, Beijing 100080.

E-mail: yzhou@amath4.amt.ac.cn