The strong representation for the nonparametric estimation of length-biased and right-censored data

Jianhua Shi\textsuperscript{1,3}, Xiaoping Chen\textsuperscript{2,3*}, and Yong Zhou\textsuperscript{3}

\textsuperscript{1}School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, China
\textsuperscript{2}School of Mathematics and Computer Science, Fujian Normal University, Fuzhou, China
\textsuperscript{3}School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China

September 17, 2014

Abstract

In this article, we consider a useful product-limit estimator of distribution function proposed by Huang & Qin(2011) when the observations are subject to length-biased and right-censored data. The estimator retains the simple closed-form expression of the truncation product-limit estimator with some good properties. An almost sure representation for the estimator is obtained which can be used to derive many properties of functional statistics based on this product-limit estimator. The rate for the remainder in the representation is of order $O(n^{-3/4} \log n^{3/4})$ a.s.

Keywords: Random truncation; Strong representation; Length-biased; Right-censored.

1 Introduction

Length-biased data are frequently appear in observational studies, when the observed samples are not randomly selected from the population of interest but with probability proportional

Acknowledgements: Zhou’s work was supported by National Natural Science Foundation of China (NSFC) (71271128), the State Key Program of National Natural Science Foundation of China (71331006), NCMIS, Shanghai First-class Discipline A and IRTSHUFE. Shi and Chen’s work was supported by National Science Foundation of Fujian Province, China (2012J01028) and National Natural Science Foundation of China (NSFC) (11301473).

*Corresponding author: Xiaoping Chen, Fujian Normal University, No.1 Keji Road, Shangjie University Town, Minhou County, Fuzhou, Fujian, China 350117. (E-mail: xpcchen@fjnu.edu.cn).
to their length (Yu et al. 2009), such as in the prevalent sampling design which only considers subjects with disease. In the prevalent sampling, the occurrence of disease onset follows a stationary Poisson process (Winter & Foldes, 1988), and the truncation time has a uniform distribution. As a result, the survival time in the prevalent cohort has a length-biased distribution, where the probability of a survival time being sampled is proportional to its length. In addition to length-biased, survival sampling data are subject to right censoring due to loss of follow-up. Hence, the length-biased and right-censored (LBRC) data is associate with left-truncation and right-censored (LTCR) data, as those who fail before sampling time are not observable. When the distribution of the truncation time is unspecified, various estimation methods to estimate survival distribution have been developed, and many works are devoted to the strong representation for estimator of the survival distribution under LTCR data but few under LBRC setting. A brief review of the theoretic development for LTCR data in recent years is worth mentioning firstly.

Assume that \((T^0, A, C^0)\) is a random vector where \(T^0\) is the survival time of interest, with unknown cumulative distribution function (c.d.f) \(F(\cdot)\), \(A\) is the random left truncation time with unknown c.d.f. \(F_A(\cdot)\) and \(C^0\) is a random right censored time with arbitrary c.d.f. \(F_C(\cdot)\). Denote \(Y = \min(T^0, C^0)\), \(\Delta = I(T^0 \leq C^0)\) is the usual indicator of censoring status and \(\bar{R}(t) = n^{-1} \sum_{i=1}^{n} I(a_i \leq t \leq y_i)\) is the empirical estimator for \(R(t) = P(A \leq t \leq Y)\), where and throughout the paper the lowercase letters for the random variables indicate the sampling value from a population. The asymptotic behavior for the estimator of \(F(\cdot)\) has been studied in many articles under some suitable conditions, such as Tsai et al. (1987) proposed the nonparametric estimator \(\hat{F}_n(x)\) for \(F(\cdot)\),

\[
1 - \hat{F}_n(x) = \prod_{y_i \leq x} [1 - (n\bar{R}(y_i))^{-1}]^{\delta_i}
\]

which is the well-known TJW product-limit (PL) estimator. Obviously, the estimator reduces to the Lynden-Bell (1971) PL-estimator if there is only left truncation and the Kaplan-Meier (1958) PL-estimator if there is only right censored \((T = 0)\).

Extensive literature has focused on the strong representation for the TJW PL estimator. As for the truncated data without censoring, Chao & Lo (1988) considered this type nonparametric maximum likelihood estimators for the survival distribution function, the remainder terms in their results are of order \(O(n^{-3/4}(\log n)^{3/4})\) and of order \(o(n^{-1/2})\) a.s. under different condition on the support of d.f., respectively. Stute (1993) utilized several new technical tools, establishing almost sure representations of the estimator \(\hat{F}_n(\cdot)\) and its cumulative hazard function estimator \(\Lambda_n(\cdot)\) with the negligible remainder term being order \(O(n^{-1}(\log n)^3)\).
almost surely. Under more restrictive integrability conditions, Arcones & Giné (1995) show that the remainder term has more strong approximation rate, i.e. \( O(n^{-1} \log \log n) \) a.s.

As to random censorship model, Lo & Singh (1986) considered the PL estimator and its quantile process, by express them as i.i.d. means of bounded random variables plus remainders with order \( O(n^{-3/4} (\log n)^{3/4}) \) a.s. uniformly over compact intervals. The results in Lo & Singh (1986) not only present a justification of the bootstrap approach for estimating the standard error of the PL estimator and its quantiles, but also provide a way of constructing confidence intervals for the unknown parameters based on the simulated data. Furthermore, the order of the remainder term is later improved to be \( O(n^{-1} \log n) \) a.s. by Burke et al. (1988), Lo et al. (1989), and Gu (1991). To characterize the rate of uniform behavior of the PL estimator, Chen & Lo (1997) give some sufficient and necessary conditions on the distribution of survival time, which somehow fill a longstanding gap in the asymptotic theory of lifetime analysis.

Under LTRC sampling, Lai & Ying (1991) using martingale theory, obtained the functional law of the iterated logarithm for modified PL estimator. Furthermore, as an extension of Theorem 1 of Chao & Lo (1988), Gijbels & Wang (1993) studied a strong representation for TJW PL estimator, and obtain the order \( O(n^{-1} \log n) \) a.s. of the remainder term under suitable condition. Zhou (1996) considered more complicated situation when the support of unknown d.f. \( G \) of the random left-truncation time is equal to the support of d.f. \( W \) of the minimum variable between interest variable and the random-censoring time, obtained some almost sure representations for the TJW PL estimator of a distribution function. The same representations are also considered by Zhou & Yip (1999), the remainder terms in Zhou Zhou & Yip (1999) are of order \( O(n^{-1} \log \log n) \) a.s. under suitable integrability assumption, which currently are the best result about the convergent rate. For references to some other related strong approximation of PL estimator see Tse (2003), Sellero et al. (2005), Liang et al. (2009), etc.

Therefore the representation of the PL estimator is interesting problem in the field of probability and statistic, which attracts many attention. However, the strong behavior for LBRC design are relatively few studied. Recently, Huang & Qin (2011) used sufficiently the potential information in the LBRC design, obtained the weak representation for the nonparametric estimator of the unknown survival distribution, and correspondingly the cumulative hazard estimator.

In this article, we consider the almost sure representation for the cumulative hazard
estimator and the estimator of the distribution function, the remainder terms in the representation are of order $O(n^{-3/4}(\log n)^{3/4})$ a.s. that extends the weak representation of Huang & Qin (2011) to the strong one. The rest of the paper is organized as follows. Some notations and the main theoretical results concerning the nonparametric estimators are presented in Section 2. Section 3 is devoted to several prepared lemmas and their proofs. Proofs for main results are deferred to Appendix finally.

2 Notations and Main Results

We now introduce related random variable notations for LBRC sampling, Huang & Qin’s (2011) notation is followed whenever possible. Let $(T^0, A, C^0)$ denote random variables where $T^0$ is the interest survival time from the disease incidence to the failure event with marginal density function $f(t)$ and survival function $S(t)$; $A$ is the random left truncation time from the disease incidence to sampling time, $\xi$, and $C^0$ is the total censoring time from the disease onset. Meanwhile, suppose $W^0$ be the onset time for the disease incidence, $C$ be the time from sampling time to censoring, i.e. it is the residual censored time.

Several basic assumptions for the general population needed in the paper are presented in the following.

Assumption 1. The distribution of $T^0$ is independent of $W^0$.

Assumption 2. The incidence of disease onset occurs over calendar time at a constant rate, that is, $W^0$ has a constant density function.

Let $F^u(t) = P(\Delta = 1, Y \leq t)$ denote the subdistribution function, and define $a_G = \inf\{t : G(t) > 0\}, b_G = \sup\{t : G(t) < 1\}$ for any d.f. $G(\cdot)$. Assume $Y \sim H(\cdot)$, compared with Woodrofe’s (1985) results, $F(\cdot)$ can be reconstructed if $a_{F_A} \leq a_H$ and $b_{F_A} \leq b_H$. Therefore we shall assume $a_{F_A} \leq a_H$ and $b_{F_A} \leq b_H$ throughout this paper. Meanwhile, put $0/0 = 0$ for convenience.

Assumption 3. For $a_H < b < b_H$,

$$\int_{a_H}^{b} \frac{dF^u(u)}{R^3(u)} < \infty.$$ 

And for illustrating the left truncated sampling, we drop the superscript 0 in the notation of $W^0$ and $T^0$, and thus

$$(W, T) \sim_d (W^0, T^0) \mid T^0 \geq \xi - W^0 > 0,$$
where \( \sim_d \) denote identical distribution.

In our setting, \( C \) is assumed to be independent of \( (W, \xi, T) \) and \( \xi \) is independent with \( (W, T) \). Since the total censoring time \( C^0 \) and the survival time \( T \) share the same \( A \), therefore, they are dependent random variables, which indicates the informative censoring for survival time \( T \). Just as the notation defined before, in LBRC model, when \( Y < A \) nothing is observed, then the only observed data here is \( (Y, A, \Delta) \) if \( Y \geq A \). Define \( \alpha =: P(Y \geq A) \), naturally one needs to assume \( \alpha > 0 \). Let \( \mu = E(T^0) = \int_0^{\infty} uf(u)du \), under Assumptions 1 and 2, the joint density function for \( (A, T) \) (Lancaster, 1990, Ch.3) is

\[
f_{A,T}(a, t) = \frac{f(t)}{\mu} I(t > a > 0),
\]

If we set \( \tilde{V} = \min(V, C) \), where \( V \) is the residual survival time from the sampling time, then the observed data is i.i.d. copies of \( (W, A, \tilde{V}, \Delta) \).

Denote the survival functions of the random variables \( A, T, C \) and \( V \) defined in the prevalent population as \( S_A(t), S_T(t), S_C(t), \) and \( S_V(t) \), which are always assumed continuous in the paper, and the corresponding marginal density functions as \( f_A(t), f_T(t), f_C(t) \) as well as \( f_V(t) \), respectively. Let \( \Lambda(\cdot) \) be the cumulative hazard function of \( F(\cdot) \), note that \( E[dF^u(t)] = \mu^{-1} f(t) \int_0^t G(s)dsdt \) under LBRC mechanism, it is easy to see that

\[
\Lambda(t) = \int_0^t \frac{dF^u(u)}{R(u)}.
\]

Therefore, the usual estimated cumulative hazard function \( \hat{\Lambda}(\cdot) \) for \( \Lambda(\cdot) \) is

\[
\hat{\Lambda}(t) = \int_0^t \frac{d\tilde{N}(u)}{R(u)},
\]

with \( \tilde{N}(t) = n^{-1} \sum_{j=1}^n \delta_j I(y_j \leq t) \). Wang (1991) and others proposed and studied the PL estimator for LTRC survival data:

\[
\hat{S}(t) = \prod_{u \in [0, t]} \{1 - d\hat{\Lambda}(u)\},
\]

Under LBRC sampling, Huang & Qin (2011) introduce Assumption 1 and Assumption 2 for the general population, and find an important relation that the truncation time, \( A \), and the residual survival time, \( V \), share the same marginal density function, i.e.

\[
f_A(t) = f_V(t) = \frac{S(t)}{\mu} I(t > 0).
\]
Basing on the key property, Huang & Qin (2011) proposed to replace the empirical estimate

\[ \bar{R}(t) = n^{-1} \sum_{j=1}^{n} I(y_j \geq t \geq a_j), \]

as the estimator

\[ \tilde{R}(t) = n^{-1} \sum_{j=1}^{n} I(y_j \geq t) - \tilde{S}_A(t), \]

where \( \tilde{S}_A(t) = \prod_{u \in [0,t]} \{1 - \frac{d\tilde{Q}(u)}{K(u)}\} \) is Kaplan–Meier estimator for \( A \), and

\[ \tilde{Q}(t) =: \frac{1}{n} \sum_{i=1}^{n} [I(a_i \leq t) + \delta_i I(\tilde{v}_i \leq t)] =: \tilde{Q}_1(t) + \tilde{Q}_2(t), \]

as well as

\[ \tilde{K}(t) =: \frac{1}{n} \sum_{i=1}^{n} [I(a_i \geq t) + I(\tilde{v}_i \geq t)] =: \tilde{K}_1(t) + \tilde{K}_2(t). \]

Here \( \tilde{Q}(t) \) and \( \tilde{K}(t) \) are the estimators for

\[ Q(t) = E\{I(A \leq t) + \Delta I(\tilde{V} \leq t)\} =: Q_1(t) + Q_2(t), \]

and

\[ K(t) = E\{I(A \geq t) + I(\tilde{V} \geq t)\} =: K_1(t) + K_2(t), \]

respectively.

Furthermore, Huang & Qin (2011) based on the estimator \( \tilde{R}(t) \), construct an alternative nonparametric estimator for \( \Lambda(\cdot) \) by combining information from both \( A \) and \( V \) under length-biased sampling:

\[ \tilde{\Lambda}(t) = \int_0^t \frac{d\tilde{N}(u)}{R(u)}, \]

The new estimator retains the simplicity of the truncation product-limit estimator with a closed-form expression. Huang & Qin (2011) further present the asymptotic large sample property, expressing the survival estimator as i.i.d. means of random variables, which is different from the previous literature. But the result in Huang & Qin (2011) is a weak asymptotic type, which may be insufficient, such as in a representation of the quantile function of \( \tilde{F}_n(t) \) with \( 1 - \tilde{F}_n(t) = \prod_{u \in [0,t]} \{1 - d\tilde{\Lambda}(u)\} \), or in studying the oscillation modulus of PL estimator, or for estimating the density and hazard function, etc. for example those of Gijbels & Wang (1993) and Zhou & Yip (1999).

Define the function \( \psi_i(t) = \psi_{1i}(t) + \psi_{2i}(t) \), where

\[ \psi_{1i}(t) = \int_0^t R^{-2}(u)I(y_i \geq u \geq a_i)dF^u(u) - \frac{\delta_i I(y_i \leq t)}{R(y_i)}, \]
and

$$\psi_{2i}(t) = \int_0^t R^{-2}(u)\{I(a_i > u) - S_A(u) - S_A(u)\phi_i(u)\}dF^u(u).$$

Then one may get

$$-n^{-1} \sum_{i=1}^n \psi_1(t) = \int_0^t \frac{d\hat{N}(u)}{R(u)} - \int_0^t \frac{\bar{R}(u)}{R^2(u)} dF^u(u),$$

and by Lemma 3.4 in Section 3,

$$-n^{-1} \sum_{i=1}^n \psi_2(t) = -\int_0^t n^{-1} \sum_{i=1}^n \left[ I(a_i > u) - S_A(u) - S_A(u)\phi_i(u) \right] \frac{dF^u(u)}{R^2(u)}$$

$$= -\int_0^t \frac{\bar{R}(u) - \check{R}(u) + O(n^{-3/4}n^{3/4})}{R^2(u)} dF^u(u) \text{ a.s.}$$

$$= -\int_0^t \frac{\bar{R}(u) - \check{R}(u)}{R^2(u)} dF^u(u) + O(n^{-3/4}n^{3/4}) \text{ a.s.}$$

The strong asymptotic properties of the proposed estimator $\tilde{\Lambda}(\cdot)$ and $\tilde{F}_n(\cdot)$ are summarized in Theorem 2.1 and Theorem 2.2, respectively, whose proofs are given in Appendix.

**Theorem 2.1.** Suppose that Assumption 3 holds for some $b < b_H$. Then uniformly in $a_H \leq x \leq b < b_H$, the random process $\tilde{\Lambda}(t) - \Lambda(t)$ has an asymptotic representation

$$\tilde{\Lambda}(t) - \Lambda(t) = -n^{-1} \sum_{i=1}^n [\psi_1(t) + \psi_2(t)] + R_{n1}(t)$$

with $\sup_{a_H \leq t \leq b} |R_{n1}(t)| = O(n^{-3/4}(\log n)^{3/4})$ a.s.

**Theorem 2.2.** Suppose that Assumption 3 holds for some $b < b_H$, then uniformly in $a_H \leq x \leq b < b_H$, we have

$$\tilde{F}_n(t) - F(t) = n^{-1} \sum_{i=1}^n (1 - F(t))[\psi_1(t) + \psi_2(t)] + R_{n2}(t)$$

with $\sup_{a_H \leq t \leq b} |R_{n2}(t)| = O(n^{-3/4}(\log n)^{3/4})$ a.s.

As an application of Theorem 2.2, one can obtain the LIL asymptotic result for the PL estimator in the following. In fact, some other similar results in Zhou & Yip(1999) can also be obtained by Theorem 2.1 or Theorem 2.2, we will consider these topics in future.

**Corollary 2.1** Assume that $a_{FA} \leq b_H$ and Assumption 3 are satisfied. Then the stochastic sequence

$$\{(n/(2 \log \log n)^{1/2})(\tilde{F}_n(x) - F(x))\}$$
is almost surely relatively compact in the supremum norm of functions over \((a_H, b]\), and its set of limit point is
\[
\{(d(b))^{1/2}(1 - F(\cdot))g(d(\cdot)/d(b)) : g \in G\}
\]
where \(d(t) = \int_{a_H}^t \frac{dN(u)}{R(u)}\) and \(G\) is Strassen’s set of absolutely continuous functions:
\[
G = \{g | g : [0, 1] \to \mathbb{R}, \; g(0) = 0, \; \int_0^1 \left(\frac{dg(x)}{dx}\right)^2 dx \leq 1\}.
\]
Consequently, write \(v^2(t) = (1 - F(t))d(t)\), then
\[
\limsup_{n \to \infty} \left(\frac{n}{2 \log \log n}\right)^{1/2} \sup_{a_H < x \leq b} \left|\tilde{F}_n(x) - F(x)\right| = \sup_{a_H < x \leq b} v(x) \text{ a.s.,}
\]
and
\[
\liminf_{n \to \infty} (n \log \log n)^{1/2} \sup_{a_H < x \leq b} \left|\tilde{F}_n(x) - F(x)\right| \leq \frac{\pi}{2\sqrt{2}} (d(b))^{1/2} \text{ a.s.}
\]

### 3 Some Lemmas and their proofs

**Lemma 3.1** (Lo & Singh, 1985) If \(\eta_1, \ldots, \eta_n\) are i.i.d with mean zero, \(|\eta_i| \leq c\) and \(\text{Var}(\eta_1) = \sigma^2\), then for any positive \(z\) and \(d\) satisfying \(cz \leq d\) and \(nz\sigma^2 \leq d^2\) one has
\[
P\left(\left|\sum_{i=1}^n \eta_i\right| \geq 3d\right) \leq 2e^{-z}.
\]

**Lemma 3.2** Under the d.f. continuity of random variable assumed above,
\[
\sup_{0 \leq x \leq t} \left|\int_0^x (K^{-1} - K^{-1})d(\tilde{Q}_1 - Q_1)\right| = O(n^{-3/4}(\log n)^{3/4}) \text{ a.s.,}
\]
and
\[
\sup_{0 \leq x \leq t} \left|\int_0^x (K^{-1} - K^{-1})d(\tilde{Q}_2 - Q_2)\right| = O(n^{-3/4}(\log n)^{3/4}) \text{ a.s.}
\]

**Proof of Lemma 3.2** Partitioning firstly the interval \([0, x]\) into subintervals \([x_i, x_{i+1}], i = 1, 2, \ldots, k_n\), with \(k_n = O(\sqrt{n}/(\log n)^{1/2})\), and \(0 = x_1 < x_2 < \cdots < x_{k_n+1} = x\) are such that
\[
K_1(x_i) - K_1(x_{i+1}) \leq M \cdot O((\log n)^{1/2}/\sqrt{n}),
\]
\[
K_2(x_i) - K_2(x_{i+1}) \leq M \cdot O((\log n)^{1/2}/\sqrt{n}),
\]
and then
\[
K(x_i) - K(x_{i+1}) \leq M \cdot O((\log n)^{1/2}/\sqrt{n}).
\]
Note that \( \tilde{Q}_1(x), Q_1(x) \) are monotone increasing function, we have, as in the proof of Lemma 2 of Lo & Singh (1986), that the left hand side in (3.1) is bounded above by

\[
\left| \int_0^x (\tilde{K}^{-1} - K^{-1}) d(\tilde{Q}_1 - Q_1) \right|
\]

\[
\leq \sup_{y \in [0,x]} \left| \tilde{K}^{-1}(y) - K^{-1}(y) \right| \sum_{i=1}^{k_n} \left| [\tilde{Q}_1(x_{i+1}) - Q_1(x_{i+1})] - [\tilde{Q}_1(x_i) - Q_1(x_i)] \right|
\]

\[
+ \sum_{i=1}^{k_n} \int_{x_i}^{x_{i+1}} \left| [\tilde{K}^{-1}(x) - K^{-1}(x)] - [\tilde{K}^{-1}(x_i) - K^{-1}(x_i)] \right| d\tilde{Q}_1(x)
\]

\[
+ \sum_{i=1}^{k_n} \int_{x_i}^{x_{i+1}} \left| [\tilde{K}^{-1}(x) - K^{-1}(x)] - [\tilde{K}^{-1}(x_i) - K^{-1}(x_i)] \right| dQ_1(x)
\]

\[
\leq 2 \max_{1 \leq i \leq k_n} \sup_{y \in [x_i,x_{i+1}]} \left| \tilde{K}^{-1}(x_i) - K^{-1}(x_i) \right| \quad \max_{1 \leq i \leq k_n} \left| \tilde{Q}_1(x_{i+1}) - Q_1(x_{i+1}) - (\tilde{Q}_1(x_i) - Q_1(x_i)) \right|
\]

\[
=: \quad A + B.
\]

For estimating \( A \), we further subdivide every \( [x_i, x_{i+1}] \) into subintervals \( [x_{ij}, x_{i(j+1)}], j = 1, \cdots, a_n \), with \( a_n = O(n^{-3/4} \log n)^{3/4} \), such that

\[
K_1(x_{ij}) - K_1(x_{i(j+1)}) = O(n^{-3/4} \log n)^{3/4}; \quad K_2(x_{ij}) - K_2(x_{i(j+1)}) = O(n^{-3/4} \log n)^{3/4},
\]

uniformly in \( i, j \). Now, since \( \left| \tilde{K} - K \right|^2 = O(n^{-1} \log n) \) a.s. by LIL, and \( K(\cdot) \) and \( \tilde{K}(\cdot) \) are bound in the intervals \( [x_i, x_{i+1}], i = 1, 2, \cdots, k_n \), it follows that

\[
\sup_{y \in [x_i,x_{i+1}]} \left| [\tilde{K}^{-1}(x_i) - K^{-1}(x_i)] - [\tilde{K}^{-1}(y) - K^{-1}(y)] \right|
\]

\[
\leq \sup_{y \in [x_i,x_{i+1}]} \left\{ \frac{\tilde{K}(y) - K(y)}{K^2(y)} - \frac{\tilde{K}(x_i) - K(x_i)}{K^2(x_i)} \right\} + O(n^{-1} \log n) \text{ a.s.}
\]

\[
\leq \sup_{y \in [x_i,x_{i+1}]} \left( K^{-2}(x_{i+1}) \left| [\tilde{K}(y) - K(y)] - [\tilde{K}(x_i) - K(x_i)] \right| \right.
\]

\[
\left. + \frac{K_i^2(x_{i+1}) - K_i^2(y)}{K_i^2(y) K_i^2(x_{i+1})} \left| \tilde{K}(y) - K(y) \right| \right.
\]

\[
\left. + \frac{K_i^2(x_{i+1}) - K_i^2(x_i)}{K_i^2(x_i) K_i^2(x_{i+1})} \left| \tilde{K}(x_i) - K(x_i) \right| \right) + O(n^{-1} \log n) \text{ a.s.}
\]

\[
\leq K_i^{-2}(x_{i+1}) \sup_{y \in [x_i,x_{i+1}]} \left| \tilde{K}(y) - \tilde{K}(x_i) - K(y) + K(x_i) \right| + O(n^{-1} \log n) \text{ a.s.}
\]

\[
\leq \text{const.} \max_{1 \leq j \leq a_n} \left| \tilde{K}(x_{ij}) - \tilde{K}(x_i) - K(x_{ij}) + K(x_i) \right| + O(n^{-3/4} (\log n)^{3/4}) \text{ a.s.}
\]

Set \( \eta_k = I(x_i < a_k \leq x_{ij}) - P(x_i < A \leq x_{ij}) \), the following probability bound can be verified from the exponential inequality of Lemma 3.1 under proper choices of parameter.
(e.g. \( c = 1, \sigma^2 = Mn^{-1/2}(\log n)^{1/2}, z = M \log n \)).

\[
\max_{i \leq k_n} \max_{1 \leq j \leq n} P\left( |\tilde{K}(x_{ij}) - \tilde{K}(x_i) - K(x_{ij}) + K(x_i)| > const. n^{-3/4}(\log n)^{3/4} \right) = O(n^{-3}).
\]

Utilizing the bound, Bonferroni inequality together with the Borel Cantelli Lemma, it follows that \( A = O(n^{-3/4}(\log n)^{3/4}) \) a.s. The estimation of \( B \) is treated similarly and leads to the same order. The proof of (3.2) is similar, we omit the details here. This completes the proof.

Assume \( a_H = 0 \) w.l.o.g. throughout to avoid trivialities. Define \( \tilde{\Lambda}_A(t) = \int_0^t \frac{d\tilde{Q}(u)}{K(u)} \) as the estimate for \( \Lambda_A(\cdot) \), the cumulative hazard function of \( A \), then.

**Lemma 3.3** Under Assumption 3, we have

\[
\sup_{0 \leq x \leq b} \left| \tilde{\Lambda}_A(x) - \Lambda_A(x) \right| = \sup_{0 \leq x \leq b} \left| \int_0^x \frac{d\tilde{Q}(u)}{K(u)} - \int_0^x \frac{dQ(u)}{K(u)} \right| = O(n^{-1/2}(\log \log n)^{1/2}) \text{ a.s.}
\]

**Proof of Lemma 3.3** Obviously,

\[
\sup_{0 \leq x \leq b} \left| \int_0^x \frac{d\tilde{Q}(u)}{K(u)} - \int_0^x \frac{dQ(u)}{K(u)} \right| \\
\leq \sup_{0 \leq x \leq b} \left| \int_0^x \frac{d\tilde{Q}(u)}{K(u)} - \int_0^x \frac{dQ(u)}{K(u)} \right| + \sup_{0 \leq x \leq b} \left| \int_0^x \frac{d\tilde{Q}(u)}{K(u)} - \int_0^x \frac{dQ(u)}{K(u)} \right| \\
=: J_{n1} + J_{n2}
\]

As to \( J_{n1} \), under Assumption 3, combining LIL for empirical processes, there is

\[
J_{n1} = \sup_{0 \leq x \leq b} \left| \int_0^x \frac{\tilde{K}(u) - K(u)}{K(u)\tilde{K}(u)} d\tilde{Q}(u) \right| \\
\leq \sup_{0 \leq x \leq b} \left( \sup_{0 \leq u \leq x} [\tilde{K}(u) - K(u)]^2 \right) \int_0^x \frac{1}{K^2(u)\tilde{K}(u)} d\tilde{Q}(u) \\
+ \sup_{0 \leq x \leq b} \int_0^x \left| \frac{\tilde{K}(u) - K(u)}{K^2(u)} \right| d\tilde{Q}(u) \\
= O((n^{-1} \log \log n)^{1/2}) \text{ a.s.}
\]

Next, put

\[
I_n(x) = \int_0^x \frac{1}{K(u)} d[\tilde{Q}(u) - Q(u)],
\]

then \( J_{n2} = \sup_{0 \leq x \leq b} I_n(x) \) is of the order \( (n^{-1} \log \log n)^{1/2} \) almost sure. In fact, the process \( I_n(x) \) satisfies LIL (e.g. Alexander & Talagrand\((1989))\), since it is an empirical process over VC
classes of function with square integral envelope, and thus \( \sup_{0 \leq x \leq b} |\tilde{\Lambda}_A(x) - \Lambda_A(x)| \) is also of the same order. This ends the proof.

We now establish the the strong representation for \( \tilde{S}_A(\cdot) \), which is constructed by pooling data from the truncation time \( A \) and the observed residual survival time \( \tilde{V} \). Lemma 3.4 in the following indicates that \( \tilde{S}_A(\cdot) \) is a strong consistent estimator of \( S_A(\cdot) \), and obviously it implies the asymptotic representation for \( \tilde{R}(\cdot) \):

\[
\tilde{R}(t) = \bar{R}(t) + \tilde{S}_A(t) - S_A(t)
\]

\[
= \bar{R}(t) + \frac{1}{n} \sum_{i=1}^{n} \{I(a_i > t) - S_A(t) - S_A(t)\phi_i(t)\} + O(n^{-3/4}(\log n)^{3/4}) \text{ a.s.}
\]

**Lemma 3.4.** Suppose that Assumption 3 holds for some \( b < b_H \). Then, the stochastic process \( \tilde{S}_A(t) - S_A(t) \) has an asymptotic representation

\[
\tilde{S}_A(t) - S_A(t) = n^{-1} \sum_{i=1}^{n} S_A(t)\phi_i(t) + R_{n3}(t),
\]

where

\[
\phi_i(t) = \int_{0}^{t} K^{-2}(u)\{I(a_i \geq u) + I(\tilde{v}_i \geq u)\}dQ(u) - \frac{I(a_i \leq t)}{K(a_i)} - \frac{\delta_i I(\tilde{v}_i \leq t)}{K(\tilde{v}_i)},
\]

and \( \sup_{a_H \leq t \leq b} |R_{n3}(t)| = O(n^{-3/4}(\log n)^{3/4}) \text{ a.s.} \)

**Proof of Lemma 3.4** By the definition of \( \phi_i(t) \), there is

\[
-n^{-1} \sum_{i=1}^{n} \phi_i(t) = -\int_{0}^{t} \frac{\tilde{K}(u)}{K^2(u)}dQ(u) + \int_{0}^{t} \frac{1}{K(u)}d\tilde{Q}(u).
\]

Since again

\[
\left| \int_{0}^{t} \left( \frac{1}{K(u)} - \frac{1}{K(u)} \right)d(\tilde{Q}(u) - Q(u)) \right| \leq \left| \int_{0}^{t} \left( \frac{1}{K(u)} - \frac{1}{K(u)} \right)d(\tilde{Q}_1(u) - Q_1(u)) \right| + \left| \int_{0}^{t} \left( \frac{1}{K(u)} - \frac{1}{K(u)} \right)d(\tilde{Q}_2(u) - Q_2(u)) \right| =: S_1 + S_2.
\]

Thus applying Lemma 3.2 to \( S_1 \) and \( S_2 \), one can derive the following asymptotic repre-
sentation under Assumption 3:

\[
\tilde{\Lambda}_A(t) - \Lambda_A(t) = \int_0^t \frac{d\tilde{Q}(u)}{K(u)} - \int_0^t \frac{dQ(u)}{K(u)}
\]

\[
= \int_0^t \left( \frac{1}{K(u)} - \frac{1}{\tilde{K}(u)} \right) d(\tilde{Q}(u) - Q(u)) - \int_0^t \frac{\tilde{K}(u)}{K^2(u)} dQ(u)
\]

\[
+ \int_0^t \frac{1}{K(u)} d\tilde{Q}(u) + \int_0^t \frac{1}{K(u)} dQ(u) \leq \frac{1}{K(u)} d(\tilde{Q}(u) - Q(u))
\]

\[
= -\frac{1}{n} \sum_{i=1}^n \phi_i(t) + \int_0^t \frac{(K(u) - \tilde{K}(u))^2}{K(u)K^2(u)} dQ(u) + \int_0^t \frac{1}{K(u)K^2(u)} dQ(u) + O(n^{-3/4}(\log n)^{3/4}) \ a.s.
\]

\[
= -\frac{1}{n} \sum_{i=1}^n \phi_i(t) + O(n^{-3/4}(\log n)^{3/4}) \ a.s.
\]

where

\[
\sup_{0 \leq u \leq b} (K(u) - \tilde{K}(u))^2 \left| \int_0^t \frac{1}{K(u)K^2(u)} dQ(u) \right| 
\]

\[
= O(n^{-1}\log n) \left[ \int_0^t \frac{1}{K^3(u)} dQ(u) \right] \ a.s.
\]

According to Lemma 3.3, by expansion of the function \( \exp \{-x\} \) in neighborhood of zero.

\[
\tilde{S}_A(t) - S_A(t) = \exp\{-\tilde{\Lambda}_A(t)\} - \exp\{-\Lambda_A(t)\}
\]

\[
= -\exp\{-\Lambda_A(t)\} (\tilde{\Lambda}_A(t) - \Lambda_A(t) + O(n^{-1}\log \log n))
\]

\[
\leq -\exp\{-\Lambda_A(t)\} \left[ -\frac{1}{n} \sum_{i=1}^n \phi_i(t) + O(n^{-3/4}(\log n)^{3/4}) + O(n^{-1}\log \log n) \right] \ a.s.
\]

That is

\[
\tilde{S}_A(t) - S_A(t) = \frac{1}{n} \sum_{i=1}^n S_A(t) \phi_i(t) + O(n^{-3/4}(\log n)^{3/4}) \ a.s.
\]

This ends the proof.

Next, similar to the discussion of Lemmas in Zhou & Yip(1999), we may derive several Lemmas. Note that for \( 0 < b < b_H \), it follows from the SLLN that

\[
\int_0^b \frac{d\tilde{N}(u)}{R(u)\left[R(u) + n^{-1}\right]} < \infty. \tag{3.3}
\]
In fact, by the LIL for the empirical process, we have
\[
\left| \int_0^b \frac{d\tilde{N}(u)}{R(u)[R(u) + n^{-1}]} - \int_0^b \frac{dN(u)}{R^2(u)} \right| \\
\leq \sup_{0 < u < \infty} \left( |R(u) - \bar{R}(u)| + n^{-1} \right) \int_0^b \frac{1}{R^2(u)\bar{R}(u)} d\tilde{N}(u).
\]

Similarly,
\[
\int_0^b \frac{1}{R^2(u)\bar{R}(u)} d\tilde{N}(u) \leq \left| \int_0^b \frac{1}{R^3(u)} d\tilde{N}(u) \right| (1 + \sup_{0 < u \leq b} \left| \frac{R(u) - \bar{R}(u)}{\bar{R}(u)} \right|),
\]

The LIL for the empirical process and SLLN yield
\[
\int_0^b \frac{1}{R^3(u)} d\tilde{N}(u) \to \int_0^b \frac{1}{R^3(u)} dN(u) < \infty \text{ a.s.}
\]

Hence the result of (3.3) holds.

For proof of Theorem 2.2, we need a slight modification of \(\bar{F}_n(\cdot)\), define a new estimator \(\bar{F}_n(\cdot)\) as
\[
1 - \bar{F}_n(x) = \prod_{y_i \leq x} \left[ 1 - \frac{1}{n\bar{R}(y_i) + 1} \right] \delta_i,
\]
which is only to safeguard against log0 when taking logarithms of \(1 - \bar{F}_n(x)\).

**Lemma 3.5** Under Assumption 3, when \(b < b_H\), there is
\[
\sup_{0 \leq x \leq b} \left| \bar{F}_n(x) - \bar{F}_n(x) \right| = O(n^{-1}) \text{ a.s.}
\]

**Proof of Lemma 3.5** Obviously, with (3.3),
\[
\sup_{0 \leq x \leq b} \left| \bar{F}_n(x) - \bar{F}_n(x) \right| \\
\leq \sup_{0 \leq x \leq b} \left| \sum_{z : y_i \leq x} \frac{\delta_i}{n\bar{R}(y_i)[n\bar{R}(y_i) + 1]} \right| \\
\leq \frac{1}{n} \sup_{0 \leq x \leq b} \int_0^x \frac{d\tilde{N}(u)}{R(u)[R(u) + n^{-1}]} + \frac{1}{n} \sup_{0 \leq x \leq b} \int_0^x \left| \frac{R(u) - \bar{R}(u)}{R(u)[R(u) + n^{-1}] + 1} \right| d\tilde{N}(u) \\
\leq \frac{1}{n} \int_0^b \frac{d\tilde{N}(u)}{R(u)[R(u) + n^{-1}]} + \frac{1}{n} \sup_{0 \leq x \leq b} \left| \frac{R(u) - \bar{R}(u)}{R(u) + n^{-1}} \right| \int_0^b \frac{d\tilde{N}(u)}{R(u)[R(u) + n^{-1}]} \\
= O(n^{-1}) \text{ a.s.}
\]

**Lemma 3.6** When \(b < b_H\), there is
\[
\sup_{0 \leq x \leq b} \left| \log(1 - \bar{F}_n(x)) + \Lambda(x) \right| = O(n^{-1}) \text{ a.s.}
\]
**Proof of Lemma 3.6** Similar to the discussion in Lemma 3.5, and using the Taylor’s expansion for the function \( \log(1 - x) \) when \( x < 1 \),

\[
\sup_{0 \leq x \leq b} \left| \log(1 - \tilde{F}_n(x)) + \tilde{\Lambda}(x) \right| = \\
= \sup_{0 \leq x \leq b} \left| \sum_{i : y_i \leq x} \delta_i \log\left(1 - \frac{1}{n\tilde{R}(y_i) + 1}\right) + \sum_{i=1}^{n} \frac{\delta_i I(y_i \leq x)}{n\tilde{R}(y_i)} \right| = \\
= \sup_{0 \leq x \leq b} \left| - \sum_{i : y_i \leq x} \frac{\delta_i}{n\tilde{R}(y_i) + 1} - \sum_{i : y_i \leq x} \delta_i \sum_{k=2}^{\infty} \frac{1}{k[n\tilde{R}(y_i) + 1]^k} + \sum_{i : y_i \leq x} \frac{\delta_i}{n\tilde{R}(y_i)} \right| = \\
= \sup_{0 \leq x \leq b} \left| \sum_{i : y_i \leq x} \frac{\delta_i}{n\tilde{R}(y_i)[n\tilde{R}(y_i) + 1]} - \sum_{i : y_i \leq x} \sum_{k=2}^{\infty} \frac{\delta_i}{k[n\tilde{R}(y_i) + 1]^k} \right| \leq \\
\leq \sup_{0 \leq x \leq b, i : y_i \leq x} \frac{I(\delta_i = 1)}{n\tilde{R}(y_i)[n\tilde{R}(y_i) + 1]} = O(n^{-1}) \text{ a.s.}
\]

**Lemma 3.7** Under Assumption 3, we have

\[
\sup_{0 \leq x \leq b} \left| \tilde{\Lambda}(x) - \Lambda(x) \right| = \sup_{0 \leq x \leq b} \left| \int_0^x \frac{d\tilde{N}(u)}{R(u)} - \int_0^x \frac{dF^u(u)}{R(u)} \right| = O(n^{-1/2} (\log \log n)^{1/2}) \text{ a.s.}
\]

**Proof of Lemma 3.7** The proof is similar to Lemma 3.3, we omit the details here.

### 4 Appendix

**Proof of Theorem 2.1.** Utilizing Lemma 3.4, we have by following the discussion of Lemma 3.2 that

\[
\int_0^t \left( \frac{1}{R(u)} - \frac{1}{\tilde{R}(u)} \right) d(\tilde{N}(u) - F^u(u)) = O(n^{-3/4} (\log n)^{3/4}) \text{ a.s.}
\]

Meanwhile,

\[
\int_0^t \frac{(R(u) - \tilde{R}(u))^2}{R(u)R^2(u)} dF^u(u) \leq \sup_{0 \leq u \leq t} (R(u) - \tilde{R}(u))^2 \int_0^t \frac{1}{\tilde{R}(u)R^2(u)} dF^u(u) \leq \sup_{0 \leq u \leq t} (R(u) - \tilde{R}(u))^2 \left\{ \int_0^t \frac{1}{R^3(u)} dF^u(u) \right\} \leq \left( \sup_{0 \leq u \leq t} |R(u) - \tilde{R}(u)| \right)^2 \left\{ \sup_{0 \leq u \leq t} \frac{|R(u) - \tilde{R}(u)|}{R(u)} \right\} \left\{ \int_0^t \frac{1}{R^3(u)} dF^u(u) \right\} = O(n^{-1} \log n) \text{ a.s.}
\]
Then, one can easily decompose the term $\tilde{\Lambda}(t) - \Lambda(t)$.

$$
\tilde{\Lambda}(t) - \Lambda(t) = \int_0^t \frac{d(\bar{N}(u) - F^u(u))}{R(u)} + \int_0^t (\frac{1}{R(u)} - \frac{1}{\bar{R}(u)}) dF^u(u)
$$

$$
= \left[ \int_0^t \frac{1}{R(u)} d\bar{N}(u) - \int_0^t \frac{\bar{R}(u)}{R^2(u)} dF^u(u) \right] - \int_0^t \frac{\bar{R}(u) - \bar{R}(u)}{R^2(u)} dF^u(u)
$$

$$
+ \int_0^t \left( \frac{1}{R(u)} - \frac{1}{\bar{R}(u)} \right) d(\bar{N}(u) - F^u(u)) + \int_0^t \frac{(\bar{R}(u) - \bar{R}(u))^2}{R(u)R^2(u)} dF^u(u)
$$

$$
= -\frac{1}{n} \sum_{i=1}^n [\psi_1(t) + \psi_2(t)] + O(n^{-3/4}(\log n)^{3/4}) \text{ a.s.}
$$

Completing the proof of Theorem 2.1.

**Proof of Theorem 2.2.** Note the fact from Lemma 1.8 of Stute (1993) that:

$$
F(t) - \bar{F}_n(t) = -(1 - F(t))[\tilde{\Lambda}_n(t) - \Lambda(t)] + R'_{n1}(t) + R'_{n2}(t),
$$

where

$$
R'_{n1}(t) = 2^{-1} \exp\{-\tilde{\Lambda}_n(t)\}[\tilde{\Lambda}_n(t) - \Lambda(t)]^2
$$

$$
R'_{n2}(t) = \exp\{-\tilde{\Lambda}^{**}_n(t)\}[\ln(1 - \bar{F}_n(t)) + \tilde{\Lambda}_n(t)]
$$

with $\tilde{\Lambda}_n(t)$ between $\tilde{\Lambda}_n(t)$ and $\Lambda(t)$ as well as $\tilde{\Lambda}^{**}_n(t)$ between $\tilde{\Lambda}_n(t)$ and $-\ln(1 - \bar{F}_n(t))$, respectively. Hence, Lemmas 3.5, 3.6, 3.7 together with the result of Theorem 2.1 yield Theorem 2.2.

**Proof of Corollary 2.1.** The rate of the strong convergence in Theorem 2.2 enough support the result of Corollary 2.1, and the proof is similar to the procedure of Corollary 2.2 in Zhou & Yip (1999), the details are omitted here.

**References**

[1] Y. Shen, J. Ning, J. Qin. Analyzing length-biased data with semiparametric transformation and accelerated failure time models. *J. Amer. Statist. Assoc.* 104 (2009) 1192-1202.

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

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

[4] E. L. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, *J. Amer. Statist. Assoc.* 53 (1958) 457-481.

[5] D. Lynden-Bell, A method of allowing for known observational selection in small samples applied to 3CR quasars, *Mon. Not. Roy. Astronom. Soc.* 155 (1971) 95-118.

[6] M. A. Arcones, E. Giné, On the law of the iterated logarithm for canonical U-statistics and processes, *Stochastic Proc. Appl.* 58 (1995) 217-245.
[7] M. T. Chao, S. H. Lo, Some representations of the nonparametric maximum likelihood estimators with truncated data, *Ann. Statist.* 16 (1988) 661-668.

[8] S. H. Lo, K. Singh, The product-limit estimator and the Bootstrap: some asymptotic representations, *Probab. Theory Related Fields*. 71 (1986) 455-465.

[9] M. D. Burke, S. Csörgö, L. Horváth, A correction to improvement of strong approximations of some biometric estimates under random censorship, *Probab. Theory Related Fields* 79 (1988) 51-57.

[10] S. H. Lo, Y. P. Mack, J. T. Wang, Density and hazard rate estimation for censored data via strong representations of the Kaplan Meier estimator, *Probab. Theory Related Fields* 80 (1989) 461-473.

[11] M. G. Gu, The Chung-Smirnov law for the Product-limit estimator under random censorship, *Chinese Ann. Math. Ser. B* 12 (1991) 96-105.

[12] K. Chen, S. H. Lo, On the rate of uniform convergence of the product-limit estimator: strong and weak laws, *Ann. Statist.* 25 (1997) 1050-1087.

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

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

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

[16] Y. Zhou, P. S. F. Yip, A strong representation of the product-limit estimator for left truncated and right censored data, *J. Multivariate Analysis*, 69 (1999) 261-280.

[17] SzeMan Tse, Strong Gaussian approximations in the left truncated and right censored model. *Statistica Sinica* 13 (2003) 275-282.

[18] C.S. Sellero, W.G. Manteiga, I.V. Keilegom, Uniform Representation of Product-Limit Integrals with Applications. *Scandinavian Journal of Statistics* 32 (2005) 563-581.

[19] M.C. Wang, Nonparametric estimation from cross-sectional survival data, *J. Amer. Statist. Assoc.* 86 (1991) 130-43.

[20] T. Lan caster, The Econometric Analysis of Transition Data, *Cambridge: Cambridge University Press* (1990).

[21] C.Y. Huang, J.Qin, Nonparametric estimation for length-biased and right-censored data. *Biometrika* 98(1) (2011) 177-186.

[22] H.Y. Liang, D.L. Li, Y.C. Qi, Strong convergence in nonparametric regression with truncated dependent data, *J. Multivariate Analysis* 100(1) (2009) 162-174.

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

[24] K. Alexander, M. Talagrand, The law of the iterated logarithm for empirical process on Vanik-Červonenkis classes, *J. Multivariate Anal.* 30 (1989) 155-166.

[25] B. Winter, A. Földes, A product-limit estimator for use with length-biased data. *Canadian Journal of Statistics* 16 (1988) 337–355.