A Strong Representation of the Product-Limit Estimator for Left Truncated and Right Censored Data

Yong Zhou*

Institute of Applied Mathematics, Academia Sinica, Beijing 100080, China and
Department of Statistics, The University of Hong Kong, Hong Kong

and

Paul S. F. Yip

Department of Statistics, The University of Hong Kong, Hong Kong

Received December 20, 1994; revised September 28, 1997

In this paper we consider the TJW product-limit estimator \( \hat{F}_n(x) \) of an unknown distribution function \( F \) when the data are subject to random left truncation and right censorship. An almost sure representation of PL-estimator \( \hat{F}_n(x) \) is derived with an improved error bound under some weaker assumptions. We obtain the strong approximation of \( \hat{F}_n(x) - F(x) \) by Gaussian processes and the functional law of the iterated logarithm is proved for maximal derivation of the product-limit estimator to \( F \). A sharp rate of convergence theorem concerning the smoothed TJW product-limit estimator is obtained. Asymptotic properties of kernel estimators of density function based on TJW product-limit estimator is given.


Key words and phrases: truncated data; censored data; product-limit estimator; almost sure representation.

1. INTRODUCTION

Let \((X, T, S)\) denote a random vector where \(X\) is the variable of interest with continuous distribution function (d.f.) \(F\); \(T\) is a random left truncation time with unknown d.f. \(G\) and \(S\) is a random right censoring time with arbitrary d.f. \(L\). It is assumed that \(X\) and \((T, S)\) are mutually independent, but \(T\) and \(S\) may be dependent. In this model, one observes \((Y, T, \delta)\) if \(Y \geq T\) where \(Y = X \wedge S = \min(X, S)\) and \(\delta = I(X \leq S)\) indicates the cause

* Correspondence to should be addressed to Young Zhou, Institute of Applied Mathematics, Academia Sinica, Beijing 100080, China. The work was supported by a Post-doctoral fellowship of the University of Hong Kong.
of failure. When $X < T$ nothing is observed. Let $\alpha = P(Y \geq T) > 0$, and let $W$ denote the d.f. of $Y$. By the independence assumption we have $1 - W = (1 - F)(1 - L)$. Let $(X_i, T_i, S_i, \delta_i), 1 \leq i \leq N$, be a sequence of independent and identically distributed (i.i.d.) random vectors. As a consequence of truncation, $n$, the size of the actually observed sample, is random. Obviously $n \leq N$ and $N$ is unknown. But we can regard the observed sample, $(X_i, T_i, S_i, \delta_i), i = 1, 2, ..., n$, as being generated by independent random variables $X_i, T_i, S_i, i = 1, 2, ..., N$. From the SLLN,

$$n/N \to \alpha = P(Y \geq T) > 0 \ a.s.$$ 

Now given the value of $n$, the observed data $(X_i, T_i, S_i, \delta_i)$ are still i.i.d. That is, $(Y_i, T_i, \delta_i), i = 1, 2, ..., n$ are i.i.d. random samples of $(Y, T, \delta)$, but the joint distribution of $Y$ and $T$ becomes $W^*(y, t) = W(Y \leq y, T \leq t \mid T \leq Y)$. Let $C(z) = P(T \leq z \leq Y \mid T \leq Y) = \alpha^{-1}P(T \leq z \leq S) \ (1 - F(z - ))$, and $C(z)$ can be can be consistently estimated by its empirical function $\hat{C}_n(z) = n^{-1} \sum_{i=1}^n I(T_i \leq z \leq Y_i)$. The nonparametric maximum likelihood estimator of $F$ based on the data $(Y_i, T_i, \delta_i), i = 1, 2, ..., n$ is the product-limit estimator (PL) estimator $\hat{F}_n$ introduced by Tsai, Jewell and Wang (1987) and defined by

$$1 - \hat{F}_n(x) = \prod_{y_i \leq x} [1 - (nC_n(y_i))^{-1}]^{d_i} \quad (1.1)$$

which is called the TJW product-limit estimator. Note that $\hat{F}_n$ reduces to the Kaplan-Meier (1958) PL-estimator when there is no left truncation ($T = 0$) and to the Lynden-Bell (1971) PL-estimator when there is no right censoring. For left truncated and right censored (LTRC) data the properties of the PL-estimator $\hat{F}_n$ were studied by Tsai et al. (1987), Gijbels and Wang (1993) and Zhou (1996). Gu and Lai (1990) and Lai and Ying (1991) slightly modified the PL estimator $\hat{F}_n$ for LTRC data. They obtained a functional law of the iterated logarithm and strong approximation results for the modified estimator using martingale theory.

Woodroofe (1985) pointed out that $F$ is identifiable only if some conditions on the support of $F$ and $G$ are satisfied. For any d.f. $K$, let $a_K = \inf \{ t : K(t) > 0 \}$ and $b_K = \sup \{ t : K(t) < 1 \}$ denote the left and right endpoints of its support. For LTRC data, compared with Woodroofe’s results, $F$ is identifiable if $a_G \leq a_W$ and $b_G \leq b_W$. In this paper we also assume that $a_G \leq a_W$ and $b_G \leq b_W$ hold. For left truncated data, Chao and Lo (1988) obtained an almost sure representation of the Lynden-Bell PL-estimator in term of a sum of i.i.d. random variables with a remainder $o(n^{-1/3})$. Such results are useful, for example, in estimating density and hazard function and in investigating the oscillation modulus of PL-estimators and properties of quantile estimators. Similarly, Stute (1993) derived a strong approximation for this PL-estimator with an improved
rate of approximation of a negligible term, \( \norm{R_n(x)} = O(n^{-1}(\log n)^3) \), where \( \norm{R_n(x)} = \sup_{a < x < b} \| R(x) \|, b < b_W \). But in \( a_G = a_W \) Stute (1993) required an integral condition
\[
\int_0^\infty \frac{dF(z)}{G(z)} < \infty, \tag{1.2}
\]
where \( a_W = a_F \) for only left-truncated data. Arcones and Giné (1995) obtained a correct almost sure bound \( \norm{R_n(x)} = O(n^{-1} \log \log n) \) for the remainder term under some more restrictive integrability hypotheses than that of Stute (1993), i.e.,
\[
\int_0^\infty \frac{dF(z)}{G(z)} < \infty. \tag{1.3}
\]
For right censored data, Lo and Singh (1986) decomposed the Kaplan–Meier PL-estimator as a mean of i.i.d. random variables plus a negligible remainder term of the order \( O(n^{-3/4}(\log n)^{3/4}) \) almost surely, uniformly over compact intervals. Major and Rejtő (1988) also obtained the same results as that of Lo and Singh (1986) with an improved order \( O(n^{-1} \log n) \) almost surely, see also Lo et al. (1989) and Gu (1991). Burke et al. (1988) considered different approximations by Wiener processes and proved that the rate of approximation to a Wiener processes is \( O(n^{-1} \log n) \) almost surely.

Based on left truncated and right censored data, Gijbels and Wang (1993) similarly proved a strong approximation for TJW product-limit estimator at the rate \( O(n^{-1} \log n) \) almost surely when \( a_G < a_W \). Zhou (1996) also established a strong approximation for TJW product-limit estimator at the rate \( O(n^{-1} \log 1 + \epsilon) \) for \( \epsilon > 1/2 \) when \( a_G = a_W \) under (1.2).

In this paper we further show that the rates of remainder terms are of order \( O(n^{-1} \log \log n) \) almost surely, in the case of \( a_G < a_W \) and the critical case of \( a_G = a_W \), with (2.1) below for LTRC data. Section 2 gives the main results and some remarks. Some applications of the main results are provided in Section 3. The proofs of lemmas are given in the Appendix.

2. MAIN THEOREMS

We introduce some notation. Assume that \( F, G \) and \( L \) are continuous.

Let \( \beta_n(y) = W_n(y) - W_1(y) \), where
\[
W_1(y) = P(Y \leq y, T = 1 | T \leq Y)
\]
and

\[ W_n(y) = n^{-1} \sum_{i=1}^{n} I(Y_i \leq y, \delta_i = 1). \]

It can easily be shown for the truncated and censored data that the cumulative hazard function is

\[ A(x) = \int_0^x \frac{dW_1(z)}{C(z)}, \quad 0 \leq x < \infty, \]

where \( dW_1(y) = \alpha^{-1} P(T \leq y \leq S) \, dF(y) \). Hence \( A(x) \) is consistently estimated by

\[ \tilde{A}_n(x) = \int_0^x \frac{dW_n(z)}{C_n(z)} = \sum_{i=1}^{n} \frac{I(Y_i \leq y, \delta_i = 1)}{nC_n(Y_i)}, \]

with the convention, put \( 0/0 = 0 \). Consider the following integrability condition, for \( a_a < b < b_w \)

\[ \int_{a_a}^{b} \frac{dW_1(z)}{C^3(z)} < \infty. \quad (2.1) \]

Obviously, this reduces to (1.2) for truncated data when the random variables \( T \) and \( S \) are independent. Certainly (1.2), (1.3) and (2.1) are satisfied when \( a_a < a_w \).

Write

\[ L_n(x) = \int_{a_a}^{x} \frac{d\tilde{A}_n(z)}{C(z)} - \int_{a_a}^{x} \frac{C_n(z) - C(z)}{C^2(z)} \, dW_1(z). \]

Our main results are the following theorems.

**Theorem 2.1.** Suppose that \( a_a \leq a_w \) and (2.1) are satisfied for some \( b < b_w \). Then we have uniformly in \( a_a \leq x \leq b < b_w \),

\[ \tilde{A}_n(x) - A(x) = L_n(x) + R_{n1}(x) \]

with

\[ \sup_{a_a \leq x \leq b} |R_{n1}(x)| = O(n^{-1} \log \log n) \quad a.s. \quad (2.2) \]

For the TJW PL-estimator we have a strong representation as follows:
Theorem 2.2. Under the assumptions of Theorem 1, then uniformly in $a_w \leq x \leq b < b_w$, we have

$$
\hat{F}_n(x) - F(x) = (1 - F(x))(A_n(x) - A(x)) + R_{n2}(x)
$$

$$
\hat{F}_n(x) - F(x) = (1 - F(x)) L_n(x) + R_{n2}(x)
$$

with

$$
\sup_{a_w \leq x \leq b} |R_{n2}(x)| = \sup_{a_w \leq x \leq b} |R'_{n2}(x)| = O(n^{-1} \log \log n) \text{ a.s. (2.9)}
$$

Remark 2.1. Note that integrability condition (2.1) reduced to (1.2) when there is no right censoring (i.e. $S = \infty$) and it always holds in the case of no left truncation (i.e. $T = 0$). Of course, (2.1) is weaker than the integrability condition in Arcones and Giné (1995 Eq 5.2). Hence our results are extensions of Arcones and Giné (1995) to both left-truncated and right censored data under weaker integrability condition.

Remark 2.2. It can easily be shown that the order of error term, $O(n^{-1} \log \log n)$ almost surely which, is sharper than those of Stute (1993), Chao and Lo (1988) for truncated data and of Gijbels and Wang (1993) and Zhou (1996) for left-truncated and right-censored data. In the case of no left truncation, i.e. $T = 0$, the integrability condition (2.1) always holds, our results include and improve those of Lo and Singh (1986), Major and Rejtő (1988).

Using the main theorems we can easily show the following strong approximation results for the product-limit estimator. Write

$$
d(t) = \int_{a_w}^{t} \frac{dW_i(u)}{C^2(z)} \text{ for } a_w < t < b_w.
$$

Corollary 2.1. Assume that $a_G \leq a_F$ and (2.1) are satisfied. If the probability space is rich enough, then (i) there exists a Gaussian process $B(z)$, $0 \leq z < \infty$, $EB(z) = 0$, with covariance function

$$
EB(s) B(t) = d(t \min(s, t)) \text{ for } a_w < s, t < b_w \text{ (2.4)}
$$

such that

$$
\sup_{a_w \leq z \leq b} |\sqrt{n}(\hat{F}_n(z) - F(z)) - (1 - F(z)) B(z)| = O(n^{-1/2} \log n) \text{ a.s.}
$$

for $b < b_w$. 
(ii) there exists a sequence of independent and identically distributed Gaussian processes $B_1(z)$, $B_2(z)$, ..., $EB_i(z) = 0$, $i = 1, 2, ...$ with covariance function (2.4) such that

$$
sup_{a_W < z \leq b} \left| \sqrt{n} \left( \hat{F}_n(z) - F(z) \right) - n^{-1/2} (1 - F(z)) \sum_{j=1}^{n} B_j(z) \right| = O(n^{-1/2} \log^2 n) \text{ a.s.}\n$$

(iii) there exists a two-parameter Gaussian process $\{G(z, u), 0 \leq z < \infty, u \geq 0\}$ with mean zero and covariance function

$$
EG(s, u) G(t, v) = u^{-1/2} v^{-1/2} \min(u, v) (1 - F(z)) (1 - F(t)) \delta(\min(s, t)),
$$

such that

$$
sup_{a_W < z \leq b} \left| \sqrt{n} \left( \hat{F}_n(z) - F(z) \right) - G(z, n) \right| = O(n^{-1/2} \log^2 n) \text{ a.s. (2.5)}\n$$

There are some similar strong approximation results for the cumulative hazard function by Theorem 2.1.

Proof. It is a relatively simple consequence of Komlós et al. (1975) and Theorem 2.2.

**Corollary 2.2.** Assume that $a_G \leq a_W$ and (2.1) are satisfied. Then the stochastic sequence $\{(n / (2 \log n))^{1/2} \left( \hat{F}_n(z) - F(z) \right)\}$ is almost surely relatively compact in the supremum norm of functions over $(a_W, b]$, and its set of limit point is

$$
\{(d(b))^{1/2} (1 - F(\cdot)) g(d(\cdot)/d(b)) : g \in \mathcal{G}\},
$$

where $\mathcal{G}$ is Strassen's set of absolutely continuous functions:

$$
\mathcal{G} = \left\{ g : [0, 1] \rightarrow \mathcal{R}, g(0) = 0, \int_0^1 \left( \frac{dg(x)}{dx} \right)^2 dx \leq 1 \right\}.
$$

Consequently,

$$
\limsup_{n \to \infty} \left( \frac{n}{2 \log \log n} \right)^{1/2} \sup_{a_W < z \leq b} |\hat{F}_n(z) - F(z)| = \sup_{a_W < z \leq b} v(t) \text{ a.s. (2.6)}
$$
and
\[ \liminf_{n \to \infty} (n \log \log n)^{1/2} \sup_{a < t \leq b} \left| \frac{\tilde{F}_n(z) - F(z)}{1 - F(z)} \right| = \frac{\pi}{8^{1/2}} (d(b))^{1/2} \quad \text{a.s.} \quad (2.7) \]
where \( v^2(t) = (1 - F(t)) d(t) \).

**Proof.** Observe that the process \( \{G(t, u), a < t \leq b, u \geq 0\} \) equals in distribution to the process \( \{(d(b))^{1/2} (1 - F(t)) a^{-1/2} W(d(t)/d(b)), t), a < t \leq b, u \geq 0\} \)
where \( W(t, u) \) is a standard two-parameter Wiener process. Hence the main parts of Corollary 2.2 follow from the standard functional laws of the iterated logarithm for a two-parameter Wiener process (cf. Csórgó and Révész (1981), Theorem 1.14.1). In fact, with probability one, the left hand side of (2.6) equals to
\[ \sup_{g \in \mathcal{G}} \max_{a < t \leq b} \| (d(b))^{1/2} (1 - F(t)) g(d(t)/d(b)) \| \leq \sup_{a < t \leq b} v(t) \]
where the inequality is obtained by Riesz's Lemma (Lemma 1.3.1 in Csórgó and Révész (1981)), according to which \( |g(t)| \leq t^{1/2} \) on \((0, 1)\) for any \( g \in \mathcal{G} \). The opposite inequality is trivial.

The “liminf” part also follows from (b) of Corollary 2.1 and Chung’s law of the iterated logarithm for partial sums of independent Wiener processes (Csórgó and Révész (1981) p. 122).

We can also obtain similar results of Corollary 2.1 and 2.2 for the cumulative hazard estimator \( \bar{A}_n(z) \) by Theorem 2.1. For example
\[ \limsup_{n \to \infty} \sqrt{\frac{n}{2 \log \log n}} \sup_{a < t \leq b} |\bar{A}_n(t) - A(t)| = (d(b))^{1/2} \quad \text{a.s.} \]
\[ \liminf_{n \to \infty} \sqrt{n \log \log n} \sup_{a < t \leq b} |\bar{A}_n(t) - A(t)| = \frac{\pi}{8^{1/2}} (d(b))^{1/2} \quad \text{a.s.} \]

Without loss of generality, we assume \( w = 0 \) and put \( 0/0 = 0 \). Csáki (1975) showed a law of the iterated logarithm for weighted empirical process, that is
\[ \sup_{x > 0} \left| \frac{U_n(x) - x}{\sqrt{x}} \right| = o(n^{-1/2} (\log n)^{(1+\varepsilon)/2}) \quad \text{a.s.} \]
for each \( \varepsilon > 0 \). The question concerning here is whether the weighted empirical process \( C_n(x) - C(x) \) has a similar law of the iterated logarithm.
when $C(x)$ has no monotony at $x$. The law of the iterated logarithm play a key role to prove Theorems 2.1 and 2.2.

Note that $C(z) = P(T \leq z \mid T \leq Y) - P(Y \leq z \mid T \leq Y)$. Write $G^*(z) = P(T \leq z \mid T \leq Y)$ and $F^*(z) = P(Y \leq z \mid T \leq Y)$. Define $W_2(y) = P(Y \leq y, \delta = 0 \mid T \leq Y)$, we have $F^*(z) = W_1(z) + W_2(z)$. We consider the following transformation

$$V_i = \begin{cases} W_1(Y_i), & \text{if } \delta_i = 1, \\ 1 - W_2(Y_i), & \text{if } \delta_i = 0. \end{cases} \quad (2.8)$$

Since $W_1$ and $W_2$ are continuous, then $V_1, ..., V_n$ are i.i.d. uniform [0, 1] random variables. Let $U_n(x) = n^{-1} \sum_{i=1}^n I(V_i \leq x)$ be the corresponding empirical distribution function. Then $W_{n1}(x) = U_n(W_1(x))$ for $W_1(x) \leq W_1(\infty)$ and $W_{n2}(x) = U_n(W_2(x))$ for $W_2(x) \leq W_2(\infty)$. It follows from Csaki (1975) that

$$\sup_{0 < x < b} \frac{H_{n1}(x) - H_1(x)}{\sqrt{H_1(x)}} = o(n^{-1/2} (\log n)^{(1+\delta)/2}) \quad a.s. \quad (2.9)$$

for $H_i = G^*$ or $W_i$ and $H_{n1} = G^*_n$ or $W_{n1}(i = 1, 2)$. It follows immediately that

$$G^*(z) = \pi^{-1} \int_{-\infty}^z P(T \leq x \land T \leq S) \, dF(x)$$

$$\leq \pi^{-1} P(T \leq z, T \leq S),$$

Similarly, we have $W_{2}(z) \leq \pi^{-1} P(T \leq S \leq z)$, where $a \land b = \min(a, b)$. We can easily choose a positive $\theta \in (0, b)$, $b < b_M$ such that for $0 < z \leq \theta$, $P(T \leq S \leq z) \leq P(T \leq z \leq S)$. Hence

$$\sup_{0 < z < \theta} \frac{G^*(z)}{C(z)} \leq \sup_{0 < z < \theta} \frac{2P(T \leq z \leq S)}{P(T \leq z \leq S)(1 - F(z))} \leq M, \quad (2.10)$$

where $M$ is some positive constant. Similarly, we have

$$\sup_{0 < z < \theta} \frac{W_i(z)}{C(z)} \leq M, \quad \text{for } i = 1, 2. \quad (2.11)$$

Thus by LIL of the empirical process, it follows from (2.9), (2.10) and (2.11) that
for each \( \varepsilon > 0 \).

**Lemma 2.1.** Under (2.1), we have for \( b < b_w \)

\[
\sup_{0 \leq x \leq b} \frac{|C_n(x) - C(x)|}{\sqrt{C(x)}} \leq \left( \sup_{0 \leq x \leq \theta} + \sup_{\theta \leq x \leq b} \right) \frac{|C_n(x) - C(x)|}{\sqrt{C(x)}} \\
\leq \sup_{0 < x \leq \theta} \frac{|G(x) - G_n(x)|}{\sqrt{G^*(x)}} \times \sup_{0 \leq x \leq \theta} \frac{|G_n(x)|}{\sqrt{C(x)}} \\
+ \sum_{j=1}^2 \sup_{0 < x \leq \theta} \frac{|W_j(x) - W_n(x)|}{\sqrt{W_j(x)}} \times \sup_{0 \leq x \leq \theta} \frac{|W_n(x)|}{\sqrt{C(x)}} \\
+ O(n^{-1/2}(\log \log n)^{1/2}) \quad a.s. \\
= o(n^{-1/2}(\log n)^{(1+\varepsilon)/2}) \quad a.s. \quad (2.12)
\]

The lemma is crucial to the following lemmas.

For the proof of Theorem 2.2 a modification of \( \hat{F}_n \) is needed to safeguard against log 0 when taking logarithm of \( 1 - \hat{F}_n(x) \). Define \( \tilde{F}_n \) by

\[
1 - \tilde{F}_n(x) = \prod_{i: X_i \leq x} \left( \frac{nc_n(X_i)}{nc_n(Y_i)} \right)^{\varepsilon_i}.
\]

The integrability condition (2.1) is assumed. The following lemmas are similar to those of Stute (1993) and Zhou (1996), but we have obtained a sharper order of approximation against those of Stute (1993) and have improved the order of convergence of Lemmas 3 and 4 of Zhou (1996) by using Lemma 2.1.

**Lemma 2.2.** For \( b < b_w \), we have

\[
\sup_{0 \leq x \leq b} |\tilde{F}_n(x) - \hat{F}_n(x)| = O(n^{-1}) \quad a.s.
\]

**Lemma 2.3.** For \( b < b_w \), we have

\[
\sup_{0 \leq x \leq b} |\log(1 - \tilde{F}_n(x)) + A_n(x)| = O(n^{-1}) \quad a.s. \quad (2.13)
\]

In the following lemma we consider the asymptotic bound of the error term in Theorem 1.
Lemma 2.4. (i) Assume that \( a_G \leq a_W \) and (2.1). Then
\[
\sup_{0 < x < b} |R_{n3}(x)| = \sup_{0 < x < b} |R_{n4}(x)| = O(n^{-1} \log \log n) \quad a.s. \tag{2.14}
\]

where
\[
R_{n3} = \int_0^x \frac{C(z) - C_n(z)}{C^2(z)} \, dF_n(z) \quad \text{and} \quad R_{n4} = \int_0^x \left( \frac{C(z) - C_n(z)}{C^2(z)} \right)^2 \, dW_n(z).
\]

The law of the iterated logarithm for the cumulative hazard estimator are proved in the next lemma.

Lemma 2.5. Under (2.1) we have
\[
\sup_{0 < x < b} |A_n(x) - A(x)| = \sup_{0 < x < b} \left| \int_0^x \frac{dW_n(z)}{C_n(z)} - \int_0^x \frac{dW(z)}{C(z)} \right| = O(n^{-1/2} (\log \log n)^{1/2}) \quad a.s.
\]

Proof of Theorem 2.1. It is easy to check from Lemma 2.4 that
\[
\sup_{0 < x < b} |R_{n3}(x)| \leq \sup_{0 < x < b} |R_{n3}(x)| + \sup_{0 < x < b} |R_{n4}(x)| = O(n^{-1} \log \log n) \quad a.s.
\]

Note that a representation for the estimator of cumulative hazard function can easily be proved by a similar argument of Lemma 1.9 of Stute (1993). This completes the proof of Theorem 2.1.

Proof of Theorem 2.2. The Taylor’s expansion for \( F(x) - \hat{F}_n(x) \) can readily be obtained as in Lemma 1.8 of Stute (1993). Therefore, Theorem 2.2 can be proved from Lemmas 2.2, 2.3, 2.5 and a straightforward application of the result of Theorem 2.1.

3. APPLICATIONS

1. Smoothed PL-estimator. We consider a smoothed PL-estimator by kernel distribution function. Reiss (1989) gave a particularly lucid exposition of the mathematical attractions of smoothing an empirical distribution function, which has many good properties. It is intuitively appealing and easy to compute. It provides an increased second order efficiency and improves the rate of convergence in bootstrapping. The kernel smoothed estimator of \( F \) based on PL-estimator can be defined as
\[
\hat{F}_n(t) = \int_{-\infty}^{\infty} K \left( \frac{t-x}{a_n} \right) d\hat{F}_n(x), \tag{3.1}
\]
where \( \{ a_n, n \geq 1 \} \) is a sequence of bandwidths tending to zero and \( K(x) = \int_{-\infty}^{\infty} k(t) \, dt \), in which \( k(t) \) is a smooth probability density function and \( F_n \) is the PL-estimator of (1.1). Corresponding to \( F_n(t) \), we can define a kernel smooth estimator of cumulative hazard function based on \( A_n(t) \) replacing \( F_n(x) \) by \( A_n(x) \) in (3.1). Write \( f = F' \).

Using Theorems 2.1 and 2.2 we can easily obtain the law of the iterated logarithm for the smoothed PL-estimator.

**Theorem 3.1.** Suppose that \( F \) is twice differentiable on \((a_W, A] \) with \( f(x) \leq M < \infty \) for all \( x \in (a_W, A] \), \( A < b_F \) and \( G \) is a \( \frac{1}{2} \)-Hölder continuous function and (2.5) are satisfied. Let \( k \) be a symmetric density function of bounded variation with \( k(x) = 0 \) outside some finite interval \([-r, r]\). Let \( \{ a_n, n \geq 1 \} \) satisfy \( n a_n^{-1} \log \log n \to \infty \). Then for any interval \([a, b] \subset (a_W, A] \), if \( n a_n^{-1} \log \log n \to 0 \), we have

\[
\limsup_{n \to \infty} \sqrt{\frac{n}{2 \log n}} \sup_{a \leq t \leq b} |F_n(t) - F(t)| = \sup_{a \leq t \leq b} (1 - F(t))(d(t))^{1/2} \quad a.s.
\]

if \( n a_n^{-1} \log \log n \to 0 \), we have

\[
\liminf_{n \to \infty} \sqrt{n \log \log n} \sup_{a \leq t \leq b} \frac{|F_n(t) - F(t)|}{1 - F(t)} = \frac{\pi}{8^{1/2}} (d(b))^{1/2} \quad a.s.
\]

**Theorem 3.2.** Suppose that the conditions of Theorem 3.1 are satisfied. Then for any interval \([a, b] \subset (a_W, A] \), if \( n a_n^{-1} \log \log n \to 0 \), we have

\[
\limsup_{n \to \infty} \sqrt{\frac{n}{2 \log n}} \sup_{a \leq t \leq b} |A_n(t) - A(t)| = (d(b))^{1/2} \quad a.s.
\]

if \( n a_n^{-1} \log \log n \to 0 \), we have

\[
\liminf_{n \to \infty} \sqrt{n \log \log n} \sup_{a \leq t \leq b} |A_n(t) - A(t)| = \frac{\pi}{8^{1/2}} (d(b))^{1/2} \quad a.s.
\]

Similarly, using Theorems 2.1 and 2.2 we can easily prove the asymptotic normality of \( F_n(x) \) and \( A_n(x) \), respectively. Define \( \gamma_n(t) = \sqrt{n} F_n(t) - F(t) \) and if \( t \leq a_W \), put \( \gamma_n(t) = \gamma_n(a_W) \).
Lemma 3.1. Assume that $a_0 \leq a_W$ and (2.1) are satisfied. Let $f$ be bounded on $(a_W, A)$, where $A < b_W$ and $a_n^{-1} \log \log n \to \infty$. Then for any $[a, b] \subset (a_W, A)$

$$\sup_{a \leq t \leq b} \sup_{|u| \leq M} |\gamma_n(t + a_n u) - \gamma_n(t)| = O(n^{-1/2} \log^2 n) \lor (a_n \log a_n^{-1})^{1/2} \quad a.s.$$ 

where $M$ is some positive constant and $a_1 \lor a_2 = \max(a_1, a_2)$.

Proof. Note that the process $\{G(t, u), a_W < t \leq b, u \geq 0\}$ equals in distribution to the process $\{(d(b))_{1/2} (1 - F(t)) n^{-1/2} W(d(t)/d(b), u), a_W < t \leq b, u \geq 0\}$ where $W(t, u)$ is a standard two-parameter Wiener process. Obviously,

$$\sup_{a \leq t \leq b} \sup_{|u| \leq M} |d(t + a_n u) - d(t)| \leq \sup_{a \leq t \leq b} \int_t^{t + a_n M} \frac{dW_1(x)}{C^2(x)} \leq M_1 a_n,$$

where $M_1$ is some constant. Hence Lemma 3.1 follows from Theorem 1.14.2 of Csörgö and Révész (1981).

Proof of Theorem 3.1. We only prove the second part of Theorem 3.1 here. We can obtain the following decomposition from integration by parts

$$n^{1/2} (\tilde{F}_n(t) - \hat{F}_n(t)) = \int_{-\infty}^\infty (\gamma_n(t - a_n x) - \gamma_n(t)) k(x) \, dx$$

$$+ n^{1/2} \int_{-\infty}^\infty (F(t - a_n x) - F(t)) k(x) \, dx$$

$$= A_n(t) + R_n(t).$$

By Lemma 3.1 and the assumption of $k$, we have

$$\sup_{a \leq t \leq b} |A_n(t)| = O((a_n \log a_n^{-1})^{1/2} \lor n^{-1/2} \log^2 n) \quad a.s.$$ 

Obviously, it follows that $\sup_{a \leq t \leq b} |R_n(t)| = O(n^{1/2} a_n^2)$. Therefore

$$n^{1/2} \sup_{a \leq t \leq b} |(\tilde{F}(t) - F(t)) - (\hat{F}_n(t) - F(t))|$$

$$= O((a_n \log a_n^{-1})^{1/2} \lor n^{1/2} a_n^2 \lor n^{-1/2} \log^2 n) \quad a.s.$$ 

Thus by Corollary 2.2, we can easily show that Theorem 3.1 holds.

Proof of Theorem 3.2. It follows immediately from Theorems 3.1 and 2.2.
2. Kernel estimators of density function. The strong i.i.d representations of TJW PL-estimator can be applied to density and hazard function estimation such as those of Gijbels and Wang (1993) and Arcones and Giné (1995).

Assume that $f$ and $\lambda$ are the density function and hazard rate function of distribution function $F$, respectively. Assume that $a_G \leq a_W$ as in theorems above and $f$ is $p \geq 1$ times continuous differentiable at $z$ with $f(z) > 0$ for $a_W < z < b_W$. A natural estimator of $f$ is defined by

$$
\hat{f}_n(z) = h_n^{-1} \int_{-\infty}^{\infty} k \left( \frac{t-z}{h_n} \right) d\hat{F}_n(t)
$$

where $k$ is a $p$th-order kernel function in $L^2[-1, 1]$ of bounded variation with support in $[-1, 1]$. The bandwidth sequence $\{h_n, n \geq 1\}$ satisfies the usual condition $h_n \downarrow 0$ and $nh_n \uparrow \infty$. Similarly, we can consider the kernel estimator for $\lambda$.

Consider the following decomposition of $\hat{f}_n(z) - f(z)$ using notation of Theorems 2.1 and 2.2, we can obtain that

$$
\hat{f}_n(z) - f(z) = \int_{-\infty}^{\infty} F(z - h_nu) \, dk(u) - f(z)
$$

$$
+ h_n^{-1} \int_{-\infty}^{\infty} [(1 - F) L_n](z - h_nu) \, dk(u)
$$

$$
+ h_n^{-1} \int_{-\infty}^{\infty} R_{n2}(z - h_nu) \, dk(u)
$$

$$
= b_n(z) + \sigma_n(z) + e_n(z),
$$

where $b_n(z)$ is essentially the bias, $\sigma_n(z)$ the main component and $e_n(z)$ the error term coming from that of Theorem 2.2. By Theorem 2.2, we have

$$
\sup_{a_W < z < b} |e_n(z)| = O(\log \log n(nh_n)) \quad a.s.
$$

(3.2)

Remark 3.1. It can easily be shown that the order of error term in (3.2) is better than that of Gijbels and Wang (1993), and is an improvement of Lo et al. (1989) for right censored data and of Arcones and Giné (1995) for left truncated data. Furthermore, the critical case $a_W = a_G$ has also been considered.

Define

$$
B_p = \left[ (-1)^p/p \right] \int_{-\infty}^{\infty} x^p k(x) \, dx \quad \text{and} \quad V = \int_{-1}^{1} k^2(x) \, dx.
$$
Asymptotic normality of \( \hat{f}_n(z) \) follows by the standard result of density function in the absence of both left truncation and right censoring.

**Theorem 3.3.** Assume that the conditions of Theorem 1.1 are satisfied. Let \( \sigma^2 = f(z)[1 - F(z)]C(z)^{-1} V \) and \( d = \lim_{n \to \infty} nh_n^{2p + 1} \). Then

\[
(nh_n)^{1/2} (\hat{f}_n(z) - f(z)) \xrightarrow{d} N(0, \sigma^2), \quad \text{if} \quad \frac{(\log \log n)^2}{nh_n} \to 0,
\]

\[
(nh_n)^{1/2} (\hat{f}_n(z) - f(z)) \xrightarrow{d} N(d^{1/2} f^{(p)}(z) B_p, \sigma^2), \quad \text{for} \quad d > 0,
\]

where \( \hat{f}_n(z) = h_n^{-1} \int k((t - z)/h_n) dF(t) \).

**Theorem 3.4.** Let \( a < a_W \). Suppose that \( F \), \( G \) and \( L \) be continuous distribution functions and (2.1) are satisfied and that \( f \) is twice continuous differentiable and non-negative on \((a, b) \cap [a_W, b_W]\). Assume that \( P(T \leq x \leq S) \) is \( \gamma \)-Hölder continuous on \((a', b') \) for some \( \gamma > 1/2 \) and \( P(T \leq x \leq S) \) and \( f \) are uniformly continuous on \((a', \infty) \). Let \( k \) be a kernel and \( \{h_n\} \) satisfy that \( \log h_n \to \log \log n \to \infty \) and \( nh_n / (\log h_n)^{-1} \to \infty \). Then for any \([a, b] \subset (a', b')\) we have

\[
\lim_{n \to \infty} \left( \frac{n h_n}{2 \log h_n} \right)^{1/2} \sup_{z \in [a, b]} \left( \frac{P(T \leq z \leq S)}{\sigma(f(z))} \right)^{1/2} |\sigma_u(z)| = V^{1/2}. \quad (3.3)
\]

**Proof:** Integration by parts and using a similar arguments in Arcones and Giné (1995), we obtain that uniformly in \([a, b]\)

\[
\sigma_u(z) = \alpha \int k \frac{(x - z)}{h_n} dF(x) + I_n^u(z), \quad (3.4)
\]

where \( I_n^u(z) \) is a sum of four empirical processes. By the laws of the iterated logarithm for empirical processes, we obtain that

\[
\lim_{n \to \infty} \left( \frac{n h_n}{2 \log h_n} \right)^{1/2} \sup_{a \leq z \leq b} |I_n^u(z)| = 0.
\]

The first term of (3.4) is the dominant part. It follows from the transformation (2.8) that the first term of right hand side of (3.4) equals to

\[
\frac{\alpha}{h_n P(T \leq z \leq S)} \int k \frac{(x - z)}{h_n} dU_n(W_1(x)) - W_1(x). \]
By strong approximations of uniform empirical process (cf. Csörgő and Révész (1981) p. 133) and Lemma 2.1 of Xiang (1994) we can easily obtain that
\[
\lim_{n \to \infty} \left( \frac{nh_n}{2 \log h_n} \right)^{1/2} \times \int k \left( \frac{x-z}{h_n} \right) d[ U_n(W_1(x)) - W_1(x)] / \sqrt{W'_n(z)} = V^{1/2} \quad a.s.
\]
This completes the proof of Theorem 3.4.

**Theorem 3.5.** Let \(-\infty < a_W \leq a_w \) and (2.1) holds. Suppose that \(F, G\) and \(L\) are continuous distribution functions, \(f\) is continuous and \(f > 0\) at neighborhood of \( z \in (a_W, b_W)\). Assume that \(P(T \leq x \leq S)\) is \(\gamma\)-Hölder continuous in neighborhood of \(z\) and \(f(A) \to 0\) as \(A \to \infty\). Let \(k\) be a kernel function and \(h_n\) satisfies \(nh_n \log n \to \infty\). Then
\[
\limsup_{n \to \infty} \left( \frac{nh_n}{2 \log n} \right)^{1/2} \pm \sigma_n(z) = \left( \frac{af(z) \ V}{P(T \leq z \leq S)} \right)^{1/2}.
\]

**Proof.** Integration by parts we have a similar expansion of \(\sigma_n(z)\) by the same argument of Diehl and Stute (1988) for \( z \in (a_W, b_W)\). Let \( k\) be a kernel function and \(h_n\) satisfies \(nh_n \log n \to \infty\). Then
\[
\limsup_{n \to \infty} \left( \frac{nh_n}{2 \log n} \right)^{1/2} \pm \sigma_n(z) = \left( \frac{af(z) \ V}{P(T \leq z \leq S)} \right)^{1/2}.
\]

Theorem 3.5 can be reduced to the untruncated and uncensored case. Define \(Z_i = Y_i\) if \(T_i \leq Y_i\), and \(\delta_i = 1\), and \(Z_i = X_i - A\) otherwise, with \(A = 2 + (b' - a_W)\), where \(a_W > -\infty\), then for fixed \(n\) on and \(z \in (a_W, b_W)\)
\[
k \left( \frac{z - Z_i}{h_n} \right) = k \left( \frac{z - Y_i}{h_n} \right) I(T_i \leq Y_i, \delta_i = 1),
\]
since \( z-x-A \leq 2 \), \( K((z-x-A)/h_n) = 0 \). Note that the density of \( Z_i \) at \( z \in (a_W, b) \) is 
\[
\phi(z, A) = x^{-1} \left\{ P(T \leq z \leq S) f(z) + (1 - P(T \leq z + A \leq S)) f(z + A) \right\}
\]
Theorem 1.1 of Deheuvels and Einmahl (1996) applies to the uncensored case and we easily obtain that

\[
\limsup_{n \to \infty} \left( \frac{nh_n}{2 \log \log n} \right)^{1/2} \left\{ \int k \left( \frac{x-z}{h_n} \right) d[ W_{n1}(x) - W_1(x) ] \right\} = V^{1/2} \text{ a.s.}
\]

Since \( A \) is arbitrarily large and \( f(z + A) \to 0 \) as \( A \to \infty \), it then follows that

\[
\limsup_{n \to \infty} \left( \frac{nh_n}{2 \log \log n} \right)^{1/2} \left\{ \int k \left( \frac{x-z}{h_n} \right) d[ W_{n1}(x) - W_1(x) ] \right\} = \frac{\pi f(z) V}{P(T \leq z \leq S)} \left( \frac{\log \log n}{nh_n} \right)^{1/2} \left( \frac{\log \log n}{h_n} \right) \left( \frac{\log \log n}{nh_n} \right)^{1/2} \text{ a.s.}
\]

**Remark 3.2.** (i) The size of error term in (3.2) is negligible as long as the sequence \( \{h_n, n \geq 1\} \) satisfies \( nh_n/\log \log n \to \infty \). The condition is often required to prove the uniform rates of convergence and the laws of the iterated logarithm for kernel density estimator, see Stute (1982 a, b) in the absence of both truncation and censorship, Arcones and Giné (1995) for truncated data and Diehl and Stute (1988) for censored data. Note that if we were to use the corresponding estimate of the error term in Gijbels and Wang (1993), i.e. (1.16), which does not give the exact rate, a smaller range of \( \{h_n, n \geq 1\} \) would be possible. Theorem 2.2 asserts that asymptotically the only effect on estimation of density function \( f \) when the data are truncated and censored is the factor \( P(T \leq z \leq S) \) in Theorem 3.4 and 3.5.

(ii) Consider the difference between \( \sigma_n(z) \) and the first term of the right hand size of (3.5) which is similar to Diehl and Stute (1988) for right censored data, we obtain that the pointwise difference of \( z \in (a_W, b_W) \)

\[
nh_n^{1/2} \sigma_n(z) = \frac{n^{1/2}x}{h_n^{1/2} P(T \leq z \leq S)} \int k \left( \frac{x-z}{h_n} \right) d[ W_{n1}(x) - W_1(x) ] + O \left( \frac{\log \log n}{(nh_n)^{1/2}} + (h_n \log \log n)^{1/2} \right) \text{ a.s.}
\]
providing that \( P(T \leq z \leq S) \) if \( r \)-Hölder continuous in a neighborhood of \( z \) for \( r > 1/2 \). Note that \( P(T \leq z \leq S) = G(z)(1 - L(z - )) \) as the random variables \( T \) and \( S \) are independent.

4. APPENDIX

Proofs of Lemmas 2.1–2.5

The following Lemma gives a bound for the ratio of \( C \) and \( C_n \). Its proof is a straight-forward adaptation of Stute’s (1993) Corollary 1.3.

**Lemma A.1.** For \( b < b_W \), as \( N \to \infty \), we have

\[
\sup_{i, Y_i \leq b} \frac{C(Y_i)}{C_n(Y_i)} = O(\log N) \quad \text{a.s.}
\]

**Proof.** See Corollary 1.3 of Stute (1993). □

Note that for \( 0 < b < b_W \) and \( 0 < p \leq 2 \), it follows from the SLLN that

\[
\int_0^b \frac{dW_{n1}(z)}{C^p(z)(C_n(z) + n^{-1})} < \infty.
\]

**Proof of Lemma 2.1.** It follows from Lemma A.1 that

\[
\sup_{0 \leq s \leq b} \sum_{i, X_i \leq s} \frac{1}{nC_n(X_i)[nC_n(X_i) + 1]}
\]

\[
\leq \frac{1}{n} \int_0^b \frac{dW_{n1}(z)}{C(z)(C_n(z) + n^{-1})}
\]

\[
+ \sup_{0 \leq s \leq b} \frac{C(x) - C_n(x)}{\sqrt{C(x)}} \sup_{i, X_i \leq b} \frac{C(X_i)}{C_n(X_i)} \int_0^b \frac{dW_{n1}(x)}{\sqrt{C(x)}(C_n(z) + n^{-1})}
\]

\[
= O(n^{-1}) \quad \text{a.s.} \quad □
\]

**Proof of Lemma 2.3.** Obviously

\[
\sup_{0 \leq s \leq b} \left| \prod_{i, Y_i \leq s} \left( \frac{nC_n(Y_i) - 1}{nC_n(Y_i)} \right)^{d_i} - \prod_{i, Y_i \leq s} \left( \frac{nC_n(Y_i) + 1}{nC_n(Y_i)} \right)^{d_i} \right|
\]

\[
\leq \sup_{0 \leq s \leq b} \sum_{i, Y_i \leq s} \frac{I(\delta_i = 1)}{nC_n(Y_i)[nC_n(Y_i) + 1]}
\]

\[
= O(n^{-1}) \quad \text{a.s.} \quad □
\]
Proof of Lemma 2.3. By expansion of the function \( \log(1-x) \) for \( x < 1 \) and Lemma 2.1, the right hand side of (2.13) equals to

\[
\sup_{0 < x \leq b} \left| \sum_{i: Y_i \leq x} \frac{\delta_i}{nC_n(Y_i)(nC_n(Y_i) + 1)} - \sum_{i: Y_i \leq x} \sum_{m=2}^{\infty} \frac{\delta_i}{m[nC_n(Y_i) + 1]^m} \right|
\]

\[
= \sup_{0 < x \leq b} \left| \sum_{i: Y_i \leq x} \frac{1}{nC_n(Y_i)(nC_n(Y_i) + 1)} \right|
\]

\[
= O(n^{-1}) \text{ a.s.}
\]

Proof of Lemma 2.4. Using a similar argument in (I) of Theorem 5.1 of Arcones and Giné (1995), we have

\[
\sup_{0 < x \leq b} |R_{a3}(x)| = O(n^{-1} \log \log n) \text{ a.s.}
\]

Note that

\[
\sup_{0 < x \leq b} |R_{a4}| \leq \sup_{0 < x \leq b} \left| \frac{C(x) - C_n(x)}{\sqrt{C(x)}} \right|^2 \int_0^x \frac{|C(z) - C_n(z)|}{C^2(z) C_n(z)} dW_{a1}(z)
\]

\[
+ \sup_{0 < x \leq b} |C(x) - C_n(x)|^2 \int_0^x \frac{dW_{a1}(z)}{C^3(z)}
\]

\[
= I_1 + I_2.
\]

The laws of large number and LIL imply \( I_1 = O(n^{-1} \log \log n) \) a.s. Again using the law of the iterated logarithm for empirical process, (2.12) and Lemma A.1, we have for each \( \varepsilon > 0 \)

\[
I_1 = o(n^{-1} (\log n)^{1+\varepsilon}) \sup_{0 < x \leq b} |C(x) - C_n(x)| \sup_{i: x_i \leq b} \frac{C(X_i)}{C_n(X_i)} \int_0^x \frac{dW_{a1}(z)}{C^3(z)}
\]

\[
= o(n^{-3/2} (\log n)^{2+\varepsilon} (\log \log n)^{1/2}) \text{ a.s.}
\]

\[
= O(n^{-1} \log \log n) \text{ a.s.}
\]

Hence this completes the proof of (2.14).

Proof of Lemma 2.5. We first prove that

\[
\sup_{0 < x \leq b} \left| \int_0^x \frac{dW_{a1}(z)}{C_n(z)} - \int_0^x \frac{dW_{a1}(z)}{C(z)} \right| = O(n^{-1/2} (\log \log n)^{1/2}) \text{ a.s.}
\]
for $b < b_w$. In fact, by Lemma A.1 and LIL for empirical processes we have

$$A_n := \sup_{0 < x \leq b} \left| \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) - \frac{1}{n} \sum_{i=1}^{n} I(C(X_i, X_i, |x|^{1/2})) \right|$$

$$= \sup_{c, 0 < x \leq b} \left( \frac{C(X,c)}{C_n(X,c)} \right) \left| \int_0^b \frac{(C_n(z) - C(z))^2}{C^3(z)} dW_{\alpha_1}(z) \right|$$

$$+ \left| \int_0^b \frac{C_n(z) - C(z)}{C^2(z)} dW_{\alpha_1}(z) \right|$$

$$= O(n^{-1/2}(\log \log n)^{1/2}) \ a.s.$$

where we have used (2.1). Write

$$I_n(x) = \int_0^x \frac{dW_{\alpha_1}(z)}{C(z)} - \int_0^x \frac{dW_1(z)}{C(z)}$$

The process $I_n(x)$ is an empirical process over VC classes of function with square integral envelope, so it satisfies the law of the iterated logarithm (e.g. Alexander and Talagrand (1989)), i.e. its sup over $0 < x \leq b$ is a.s. of the order $(\log \log n)^{1/2}$.

ACKNOWLEDGMENTS

The authors thank the referee and the Editor for their many valuable suggestions and comments.

REFERENCES


