Strong Convergence Rates of the Product-limit Estimator for Left Truncated and Right Censored Data under Association

A.H. Shabani, H. Jabbari* and V. Fakoor

Department of Statistics, Faculty of Mathematics, Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran

Received: 15 October 2018 / Revised: 1 May 2019 / Accepted: 21 May 2019

Abstract

Non-parametric estimation of a survival function from left truncated data subject to right censoring has been extensively studied in the literature. It is commonly assumed in such studies that the lifetime variables are a sample of independent and identically distributed random variables from the target population. This assumption is often prone to failure in practical studies. For instance, when recruited subjects are all from the same institute or the same geographical region. To the best of our knowledge, there is no study in the past literature addressing such situations. In this article, we study large and small sample behavior of Tsai-Jewell-Wang estimator under positive and negative association.

Keywords: Negative and positive association; Random left truncation and right censorship model; Strong consistency; Nelson-Allen estimator; Product-limit estimator.

Introduction

Failure time data typically comprise an initiating event, say the onset of a disease, and a terminating event, and say death due to the disease. The time between these two events is essentially what one hopes to observe in a follow-up study. Two commonly encountered issues that complicate the analysis of failure time data are left truncation and right censoring. The former occurs when the recruited subjects to the study have already experienced the initiating events, while the latter happens when the follow-up on some subjects is lost or the event time is not observed before the end of the study. Failure time data collected from follow-up studies on prevalent subjects, i.e. patients who have experienced the initiating event before being recruited to the study, are left truncated and can be subject to right censoring.

The classical setting in survival analysis, i.e. right censored failure time data, has been the subject of an intensive research over the past almost five decades. Properties of the estimators of the survivor and the cumulative hazard function, Kaplan-Meier and Nelson-Aalen estimator respectively, have been extensively studied under this classical setting. For example, asymptotic properties like uniform consistency and weak convergence were obtained by Gill [6, 8], Stute and Wang [20] and Stute [19]. Zhou [28] studied the asymptotic behavior of an estimator of distribution function in independent observation under left truncation and right censoring.

There are some results available for the case that these observations exhibit some kind of dependence. For instance, for positively associated (PA) r.v.’s, Bagai

* Corresponding author: Tel: +9838805691; Fax: +9838805691; Email: Jabbarinh@um.ac.ir

PA data can occur in many applications, like, systems reliability and statistical mechanics (see, for instance, Newman [13]). Brindley and Thompson [3] investigated positive dependence in Harris’ multivariate extension of increasing failure rate. The concept of negatively associated (NA) r.v.’s was introduced by Alam and Saxena [1]. Joag-Dev and Proschan [10] studied NA property, carefully and introduced some applications of NA r.v.’s. Also, the asymptotic normality of the random fields under PA (NA) case has been established by Roussas [15]. Moreover, Ying and Wei [26] studied the survival function estimator for censored negatively superadditive dependent data.

Our focus in this paper is on the study of asymptotic properties of TJW product-limit estimator (introduced by Tsai et al., [22]) for right censored and left truncated PA (NA) failure times (see Definition 1). Gijbels and Wang [5] obtained a strong iid representation for survival function and a kernel estimator of density function under left truncation and right censoring. Although, our line of attack is similar to that of Cai and Roussas [4], who considered the same problem, only under right censoring, many steps in the paper should be modified or adjusted to accurate for left truncation in addition to right censoring.

The layout of this paper is as follows: Section 2 includes preliminary materials and methods required to establish the limit theorems. In Section 3, the strong uniform consistency of TJW product-limit estimator is discussed, under either PA or NA and a number of simulation studies are carried out in this section. The proofs of main theorems are documented in the last subsection of Section 3. Section 4 consists of our discussion on the results of this paper.

Materials and Methods

Let \( \{X_n, N \geq 1\} \) be a sequence of lifetime variables with a common continuous marginal distribution function (df) \( F \), which may not be mutually independent. Suppose that \( \{Y_n, N \geq 1\} \) is a sequence of iid r.v.’s with continuous df \( H \) and also is independent of \( X_i \)’s. Let the r.v.’s \( X_i \) be right censored by the r.v.’s \( Y_i \), so that one observes only pairs \( (Z_i, \delta_i) \) such that

\[
Z_i = X_i \wedge Y_i \quad \text{and} \quad \delta_i = I(\{X_i \leq Y_i\}),
\]

where \( \wedge \) and \( I(\cdot) \) stand for minimum and indicator function of the event specified in the parenthesis, respectively. Let \( \{T_N, N \geq 1\} \) be a sequence of iid r.v.’s with continuous df \( G \). The major concern is drawing nonparametric inference about \( F \), based on the right censored and left truncated observations \( (Z_i, T_i, \delta_i), i = 1, \ldots, N \), where sample size \( N \) is fixed but unknown. In the left truncated model, \( (Z_i, T_i) \) is observed only when \( Z_i \geq T_i \) for \( i = 1, \ldots, n \), where \( n < N \). In fact, we observe a subsequence of original data, but we show both sequences by same letters.

Let \( \gamma = P(T_i \leq Z_i) > 0 \) and without loss of generality \( X_i, T_i \) and \( Y_i \) are non-negative r.v.’s and they are independent of each other. Suppose that the cumulative hazard function of \( F \) is \( \Lambda(x) \equiv \int_0^1 (1-F(u))^{-1} dF(u) \cdot \)

we define

\[
C(x) = P(T_i \leq x, T_i \leq Z_i) = \gamma^{-1} P(T_i \leq x \leq Y_i) \times (1-F(x)),
\]

and

\[
W(x) = P(Z_i \leq x, \delta_i = 1 | T_i \leq Z_i) = \gamma^{-1} \int_0^x P(T_i \leq u \leq Y_i) dF(u).
\]

It can be shown that

\[
\Lambda(x) = \int_0^x \frac{dW(u)}{C(u)}
\]

Let \( C_n(x) \) and \( W_n(x) \) be the empirical estimators of \( C(x) \) and \( F(x) \), respectively, i.e.
\[
C_n(x) = n^{-1} \sum_{i=1}^{n} I(T_i \leq x \leq Z_i)
\]
and
\[
W_{1n}(x) = n^{-1} \sum_{i=1}^{n} I(Z_i \leq x, \delta_i = 1),
\]
where \( F_i(x) = P(Z_i \leq x, \delta_i = 1) \) is a sub-distribution of uncensored observations. Then, TJW product-limit estimator \( \hat{F}_n \) of \( F \), which is proposed by Tsai et al. [22] is given by
\[
\hat{F}_n(x) = \begin{cases} 
1 - \prod_{z < x} \left( 1 - \frac{1}{\hat{C}_n(z)} \right)^{0.5} & ; x < Z_n, \\
1 & ; x \geq Z_n.
\end{cases}
\]
The estimator of \( \Lambda(x) \) which is comparable with Nelson-Aalen estimator of the cumulative hazard function for right censored data is
\[
\hat{\Lambda}_n(x) = \int_0^x \frac{dW_{1n}(u)}{C_1(u)} = \sum_{i=1}^{n} \frac{I(Z_i \leq x, \delta_i = 1)}{nC_n(Z_i)}.
\]
The definition of the underlying dependence considered here is as follows.

**Definition 1** A finite family of random variables \( \{ X_i, 1 \leq i \leq n \} \) is said to be PA if, for every coordinate-wise non-decreasing functions \( f_1, f_2 : R^d \rightarrow R, E[f_2^2(X_i, 1 \leq i \leq n)] < \infty \) and \( E[f_2^2(X_j, 1 \leq j \leq n)] < \infty \), we have
\[
\text{Cov}(f_1(X_i, 1 \leq i \leq n), f_2(X_j, 1 \leq j \leq n)) \geq 0.
\]
The above r.v.’s are said to be NA, if for every non-empty proper subset \( A \) of \( \{ 1, 2, \ldots, n \} \) and for every coordinate-wise non-decreasing functions, \( f_1 : R^{\text{Card}(A)} \rightarrow R, f_2 : R^{\text{Card}(A')} \rightarrow R \), \( E[f_2^2(X_i, i \in A)] < \infty \) and \( E[f_2^2(X_j, j \in A')] < \infty \), we have
\[
\text{Cov}(f_1(X_i, i \in A), f_2(X_j, j \in A')) \leq 0.
\]
An infinite family of random variables is PA (NA), if every finite subfamily is PA (NA).

Now, we introduce some general assumptions to be used throughout the article.

(A1). \( \{ X_n, n \geq 1 \} \) is a stationary sequence of PA (NA) r.v.’s with marginal df \( F \), having bounded density and finite second moment.

(A2). The censoring sequence \( \{ Y_n, n \geq 1 \} \) is iid r.v.’s with bounded density and independent of \( \{ X_n, n \geq 1 \} \). The truncated sequence \( \{ T_n, n \geq 1 \} \) is also iid r.v.’s and independent of \( \{ X_n, n \geq 1 \} \) and \( \{ Y_n, n \geq 1 \} \).

(A3). \[ \sum_{j=2}^{\infty} j^{-2} \sum_{i=1}^{j-1} |\text{Cov}(X_i, X_j)|^{1/3} < \infty. \]

(A4). \[ \sum_{i=1}^{\infty} |\text{Cov}(X_i, X_j)|^{1} = O(n^{\alpha}) , \text{ for some } 0 < \alpha \leq 1. \]

(A5). \[ \sum_{j=2}^{\infty} j^{-2} \sum_{i=1}^{j-1} |\text{Cov}(X_i, X_j)|^{1} = O(n^{-2r-2}) , \text{ for some } r > 2. \]

For every df \( L \) denotes the left and right endpoints of its support by \( a_L = \inf \{ x; L(x) > 0 \} \) and \( \tau_L = \sup \{ x; L(x) < 1 \} \), respectively. For the dfs \( F, G \) and \( H \), the values \( \tau_F, \tau_G, \tau_H \) (possibly infinite), \( a_F, a_G \) and \( a_H \) (non-negative) are defined by \( \tau_F = \sup \{ x; F(x) < 1 \} \) and \( a_F = \inf \{ x; F(x) > 0 \} \), and \( \tau_G, \tau_H, a_G \) and \( a_H \) are defined in the same way.

We can find df of \( Z \) by \( W(\cdot) = 1 - F(\cdot)H(\cdot) \). Then according to Stute and Wang [20], left and right endpoints of df \( W \) are
\[
\tau_W = \tau_F \wedge \tau_H \text{ and } a_W = a_F \wedge a_H.
\]
Then under the current model, we assume that \( a_G \leq a_W \) and \( \tau_G \leq \tau_W \).

**Results**

In this section, some theorems including strong uniform consistency of \( \hat{F}_n \) and \( \hat{\Lambda}_n \) with rates of convergence are introduced. In order to prove all limit theorems, we need the following proposition.

**Proposition 1** Suppose that (A1) and (A2) hold. Then
(i) \( \{ X_n, i \geq 1 \} \) has PA property and (A3) is satisfied, it holds
\[
\sup_{a_F \leq x \leq a_G} |C_n(x) - G(x)[1 - W(x)]| \rightarrow 0 \text{ as } n \rightarrow \infty,
\]
and
\[
\sum_{j=2}^{\infty} j^{-2} \sum_{i=1}^{j-1} |\text{Cov}(X_i, X_j)|^{1} < \infty. \]

and
\[
\sup_{x \in \mathbb{R}} \left| W_{I_n}(x) - F_n(x) \right| \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty.
\]

\begin{align}
(9)
\end{align}

(ii) If \( \{X_n, n \geq 1\} \) is NA, it follows that (8) and (9) hold true.

**Proof.** See the Appendix.

The proof of the following theorem has been omitted due to the similarity to the proof of Theorem 1.1 of Cai and Roussas [4]. The proof is yielded by using Lemma 2 in Gill [7] and an application of Proposition 1.

**Theorem 1** Suppose that (A1) and (A2) hold. Then, for all \( a \) \( < a < \tau < \tau_W \), we have

(i) If \( \{X_i, i \geq 1\} \) is PA and (A3) is satisfied, then

\[
\sup_{x \in \mathbb{R}} \left| \hat{\Lambda}_n(x) - \Lambda(x) \right| \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty.
\]

\begin{align}
(10)
\end{align}

(ii) If \( \{X_i, i \geq 1\} \) is NA, then (10) holds.

The uniform strong consistency of TJW product-limit estimator is proved in the following theorem.

**Theorem 2** Suppose that (A1), (A2) and assumptions either in part (i) or part (ii) of Theorem hold. Then

\[
\sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty.
\]

\begin{align}
(11)
\end{align}

and

\[
\sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty.
\]

\begin{align}
(12)
\end{align}

where

\[
Z_{\min} = \max_{i \leq n} Z_i \quad \text{and} \quad Z_{\max} = \min_{i \leq n} Z_i.
\]

**Proof.** See the Appendix.

**Theorem 3** Suppose that (A1), (A2) and (A5) hold. Then for every \( a \) \( < a < \tau < \tau_W \), we have

\[
\sup_{x \in \mathbb{R}} \left| \hat{\Lambda}_n(x) - \Lambda(x) \right| = O(n^{-\delta}) \quad \text{a.s.},
\]

where \( 0 < \delta < \frac{1}{2(r^2+2+r^3)} \), for any \( \delta > 0 \) and \( r \) is given by (A5).

**Proof.** For PA case, (A5) implies (26). Then by applying Lemma (i) and (ii), Remark 1.3 and Corollary 2.1 in Roussas [14], (13) holds true for both PA and NA cases (repeat the same way of the proof of Theorem 1.3 in Cai and Roussas, [4]).

Due to the similarity of the proof of the following theorem to that of Theorem 1.4 in Cai and Roussas [4], we omit it.

**Theorem 4** Suppose that (A1), (A2) and (A4) hold. Then for every \( a \) \( < a < \tau < \tau_W \), we have

\[
\sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right| = O(n^{-\theta}) \quad \text{a.s.},
\]

where \( \theta \) is defined in Theorem 3.

**Simulation study**

In the first subsection, a Monte Carlo simulation has been carried out to check the goodness of convergence rate of the estimators. For this purpose, MATLAB software version R2012a is used. We generate PA (NA) data and calculate \( \hat{F}_n(\cdot), \Lambda_n(\cdot) \) and \( \Lambda_n^{(\rho)}(\cdot) \), then, the convergence rate for these estimators are obtained by iterating this process.

Using Ghosh [9] method, to generate NA data, we could use \( n \)-variate normal distribution with \( \mu = (\mu_1, \mu_2, \ldots, \mu_n) \) and the following covariance matrix:

\[
\Sigma = \frac{1}{1-\rho^2} \begin{pmatrix}
1 & -\rho & -\rho^2 & \cdots & -\rho^{n-1} \\
-\rho & 1 & -\rho & \cdots & -\rho^{n-2} \\
-\rho^2 & -\rho & 1 & \cdots & -\rho^{n-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\rho^{n-1} & -\rho^{n-2} & -\rho^{n-3} & \cdots & 1
\end{pmatrix}.
\]

\begin{align}
(20)
\end{align}

We have a vector with NA property when \( \rho > 0 \). In order to generate PA data, it is enough to follow the same structure which used in NA case, except removing the negative sign from every non-diagonal elements of \( \Sigma \) in (20).

Accordingly, a vector with PA (NA) property will be obtained. We set \( \rho = 0.2, \mu_i = 12; \forall i = 10, \ldots, 1000 \) and the censored and truncation samples are generated from \( N(13,1) \) and \( N(11,1) \), respectively. In addition, \( \hat{F}_n(\cdot) \) and \( \hat{\Lambda}_n(\cdot) \) are calculated. Then, (10) and (11) are obtained as a function of \( n \) in NA (PA) case and we show them by \( d_{\hat{\Lambda}_n}(n) \) and \( d_{\hat{F}_n}(n) \), respectively.

Moreover, we calculate the following function for checking almost sure convergence of \( \hat{\Lambda}_n^{(\rho)}(\cdot) \).

\[
d_{\hat{\Lambda}_n^{(\rho)}}(n) = \sup_{x \in \mathbb{R}} \left| \hat{\Lambda}_n^{(\rho)}(x) - \Lambda_n^{(\rho)}(x) \right|.
\]

\begin{align}
(21)
\end{align}

For smoothing the results, this process is iterated 1000 times, independently. So, the average value of \( d_{\hat{\Lambda}_n}(n) \), \( d_{\hat{\Lambda}_n^{(\rho)}}(n) \) and \( d_{\hat{F}_n}(n) \) are indicated in Figure 1 which shows the plot of \( d_{\hat{F}_n}(n) \) (Graph 1), \( d_{\hat{\Lambda}_n}(n) \) and \( d_{\hat{\Lambda}_n^{(\rho)}}(n) \) (Graph 3) versus \( n \) for NA case, moreover, Graph 2 and Graph 4 show the same functions for PA case. The red lines are the convergence rates (14) for \( \hat{F}_n(\cdot) \) using \( \theta = 0.27 \) and the yellow lines are the convergence rates (13) for \( \hat{\Lambda}_n(\cdot) \) using...
In order to show the convergence rate of our estimators numerically, we summarized the simulations of PA and NA cases in the following two tables for \( n = \{20, 30, 50, 100, 200, 500\} \), by iterating 1000 times for \( \rho = \{0.1, 0.2, 0.3\} \).

According to the estimator of \( \Lambda(\cdot) \) in Figure 1, we see that the convergence rates of both estimators are really closed but the convergence rate of \( \hat{\Lambda}_n(\cdot) \) is mostly located at the bottom of \( \hat{\Lambda}_n(\cdot) \). Therefore, it is preferred to use \( \hat{\Lambda}_n(\cdot) \) due to easier computation and better convergence rate.

In these figures, we can see that the convergence rates are good in NA (PA) case i.e.:

(a) In Graph 1 and Graph 2 of Figure 1, we can see the convergence rate could get sharper and these graphs show that the convergence of \( \hat{F}_n(\cdot) \) is good.

(b) In Graph 3 and Graph 4 of this Figure, the convergence rate is not reasonable as well as Graph 1 and Graph 2, but it is good enough to present. Since the acceptable range of \( \Lambda(\cdot) \) is \([0, +\infty)\), so the differences less than 0.5 could be reasonable.

**Remark 1** In these simulation studies, it is easy to show that the Assumptions (A3), (A4) and (A5) are held true.

**Proofs**

The following lemmas are necessary for proving Proposition 1 and other limit theorems.

**Lemma 1** (Cai and Roussas, [4]) Let \( \{X_n, n \geq 1\} \) be a stationary sequence of r.v.’s. Then

(i) If the r.v.’s are PA, having finite variance and (A3) holds, it follows that

\[
\frac{1}{n} \sum_{i=1}^{n} (X_i - EX_i) \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty \quad (25)
\]

(ii) If the r.v.’s are NA with finite first moment, the convergence in (25) holds true.

**Lemma 2** (Cai and Roussas, [4]) Let \( U = (U_1, \ldots, U_m)' \) and \( V = (V_1, \ldots, V_m)' \) be two independent random vectors. Then

(i) If \( X_i = g_i(U_i, V) \), \( i = 1, \ldots, n \) where \( g_i(U, V) \) is non-decreasing in each \( V_j \), \( j = 1, \ldots, m \) for fixed \( U \) and \( V \) is PA, then so are \( X = (X_1, \ldots, X_n)' \).

(ii) For \( m = n \), if \( X_i = g_i(U_i, V) \), \( i = 1, \ldots, n \) and \( g_i(U_i, V) \) is non-decreasing in \( U_i \), for fixed \( U_i \) and \( V \) is NA, then \( X = (X_1, \ldots, X_n)' \) is also NA.

**Lemma 3** (Cai and Roussas, [4]) Let \( \{X_n, n \geq 1\} \) be a sequence of r.v.’s where \( EX_j = 0 \) and
for $i, j$ are iid and $\rho^2 = 2$. By corollary of Theorem 1 in Sadikova [16] in order to do this, by independence of $\text{PA}$, Lemma (i) implies that

$$\text{Vol. 30 No. 2 Spring 2019}$$

then exists a constant $C < \infty$ for $j \geq 1$. Then

(i) If the r.v.’s are PA and satisfy the condition

$$\sup_{k \leq 1} \sum_{j \neq k} \left| \text{Cov}(X_j, X_i) \right| = O(n^{-(r-2)/2}),$$

for some $r > 2$, it follows that, for all $n \geq 1$, there exists a constant $B > 0$, not depending on $n$, such that

$$\sup_{m \geq 2} E \left[ \sum_{j=m+1}^{m+n} X_j \right] \leq Bn^{r/2}. \tag{27}$$

(ii) If the r.v.’s are NA, the inequality in (27) holds true for every $r > 2$.

**Proof of Proposition 1.** (i) Since $\{X_i, 1 \leq i \leq n\}$ is PA, Lemma (i) implies that $\{I(T_i \leq x \leq Z_i), 1 \leq i \leq n\}$ is also PA on interval $[d_{\min}, \tau_{\max}]$. By corollary of Theorem 1 in Sadikova [16] and (21) in Newman [13] and because $T_j$’s are iid and also, independent of $Z_i$’s, there exists $M > 0$, such that, for all $i \neq j$ we have

$$\text{Cov}\left( I(T_i \leq x \leq Z_i), I(T_j \leq x \leq Z_j) \right) = E[I(T_i \leq x)]E[I(T_j \leq x)]E[I(Z_i \geq x)]E[I(Z_j \geq x)] - E[I(T_i \leq x)]E[I(T_j \leq x)]E[I(Z_i \geq x)]E[I(Z_j \geq x)] \leq \text{Cov}(I(Z_i \geq x), I(Z_j \geq x))$$

$$= \text{Cov}(I(Z_i \geq x), I(Z_j < x)) \leq M \text{Cov}^3(Z_i, Z_j). \tag{28}$$

Next, we will find an upper bound for $\text{Cov}(Z_i, Z_j)$ for all $i \neq j$. In order to do this, by independence of $\{X_i, i \geq 1\}$ and $\{Y_i, i \geq 1\}$, we obtain

$$\text{Cov}(Z_i, Z_j) = \int E[(X_i, X_j, X_i, X_j)] - E(X_i, X_j)E(X_i, X_j)dG(y_1)dG(y_2)$$

$$= \int E((X_i, X_j, X_i, X_j)) dG(y_1)dG(y_2) \tag{29}$$

Let $f_y(x) = \min(x, y)$ for fixed $y > 0$. So, $f_y(x)$ is a non-decreasing function of $x$ on interval $[d_{\min}, \tau_{\max}]$. Then, by Hoeffding’s inequality (see, for example, Theorem 2.2.6 in Vershynin [23]) for fixed $y_1 > 0$ and $y_2 > 0$, we conclude that $\text{Error! Bookmark not defined.}$

### Table 1. The values of $d_{\text{PA}}(n), d_{\text{NA}}(n)$ and $d_{\text{IIT}}(n)$ for different values of $\rho$ and $n$ in PA.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Estimator</th>
<th>$n=20$</th>
<th>$n=30$</th>
<th>$n=50$</th>
<th>$n=100$</th>
<th>$n=200$</th>
<th>$n=500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>Survival function</td>
<td>0.2477</td>
<td>0.1538</td>
<td>0.1272</td>
<td>0.1018</td>
<td>0.0898</td>
<td>0.0885</td>
</tr>
<tr>
<td>0.2</td>
<td>Survival function</td>
<td>0.2525</td>
<td>0.1605</td>
<td>0.1317</td>
<td>0.1041</td>
<td>0.0907</td>
<td>0.0889</td>
</tr>
<tr>
<td>0.3</td>
<td>Survival function</td>
<td>0.2625</td>
<td>0.1661</td>
<td>0.1403</td>
<td>0.1108</td>
<td>0.0994</td>
<td>0.0947</td>
</tr>
<tr>
<td>0.1</td>
<td>Cumulative hazard function</td>
<td>0.8109</td>
<td>0.5689</td>
<td>0.4852</td>
<td>0.4160</td>
<td>0.3937</td>
<td>0.3914</td>
</tr>
<tr>
<td>0.2</td>
<td>Cumulative hazard function</td>
<td>0.8105</td>
<td>0.5534</td>
<td>0.4699</td>
<td>0.4013</td>
<td>0.3728</td>
<td>0.3671</td>
</tr>
<tr>
<td>0.3</td>
<td>Cumulative hazard function</td>
<td>0.7995</td>
<td>0.5447</td>
<td>0.4444</td>
<td>0.3798</td>
<td>0.3494</td>
<td>0.3362</td>
</tr>
<tr>
<td>0.1</td>
<td>Cumulative hazard function</td>
<td>1.2254</td>
<td>0.5628</td>
<td>0.4752</td>
<td>0.4036</td>
<td>0.3844</td>
<td>0.3883</td>
</tr>
<tr>
<td>0.2</td>
<td>Cumulative hazard function</td>
<td>1.3412</td>
<td>0.5618</td>
<td>0.4659</td>
<td>0.3953</td>
<td>0.3667</td>
<td>0.3646</td>
</tr>
<tr>
<td>0.3</td>
<td>Cumulative hazard function</td>
<td>0.9115</td>
<td>0.5422</td>
<td>0.4479</td>
<td>0.3766</td>
<td>0.3467</td>
<td>0.3350</td>
</tr>
</tbody>
</table>

### Table 2. The values of $d_{\text{PA}}(n), d_{\text{NA}}(n)$ and $d_{\text{IIT}}(n)$ for different values of $\rho$ and $n$ in NA.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Estimator</th>
<th>$n=20$</th>
<th>$n=30$</th>
<th>$n=50$</th>
<th>$n=100$</th>
<th>$n=200$</th>
<th>$n=500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>Survival function</td>
<td>0.1398</td>
<td>0.1187</td>
<td>0.1004</td>
<td>0.0655</td>
<td>0.0575</td>
<td>0.0452</td>
</tr>
<tr>
<td>0.2</td>
<td>Survival function</td>
<td>0.1552</td>
<td>0.1314</td>
<td>0.1302</td>
<td>0.1117</td>
<td>0.0543</td>
<td>0.0340</td>
</tr>
<tr>
<td>0.3</td>
<td>Survival function</td>
<td>0.1795</td>
<td>0.1795</td>
<td>0.1244</td>
<td>0.0743</td>
<td>0.0586</td>
<td>0.0402</td>
</tr>
<tr>
<td>0.1</td>
<td>Cumulative hazard function</td>
<td>0.5629</td>
<td>0.4753</td>
<td>0.4343</td>
<td>0.2597</td>
<td>0.2421</td>
<td>0.2327</td>
</tr>
<tr>
<td>0.2</td>
<td>Cumulative hazard function</td>
<td>0.6456</td>
<td>0.6106</td>
<td>0.4685</td>
<td>0.3288</td>
<td>0.2802</td>
<td>0.2047</td>
</tr>
<tr>
<td>0.3</td>
<td>Cumulative hazard function</td>
<td>0.3608</td>
<td>0.2793</td>
<td>0.2119</td>
<td>0.1988</td>
<td>0.1723</td>
<td>0.1324</td>
</tr>
<tr>
<td>0.1</td>
<td>Cumulative hazard function</td>
<td>0.5017</td>
<td>0.4921</td>
<td>0.3638</td>
<td>0.2578</td>
<td>0.2338</td>
<td>0.2117</td>
</tr>
<tr>
<td>0.2</td>
<td>Cumulative hazard function</td>
<td>0.6737</td>
<td>0.6136</td>
<td>0.4666</td>
<td>0.3293</td>
<td>0.2872</td>
<td>0.2317</td>
</tr>
<tr>
<td>0.3</td>
<td>Cumulative hazard function</td>
<td>0.6917</td>
<td>0.4121</td>
<td>0.3807</td>
<td>0.2847</td>
<td>0.2123</td>
<td>0.1928</td>
</tr>
</tbody>
</table>
\[
\text{Cov}(X, Y) = \text{Cov}(f_a(X), f_b(X))
\]
\[
= \int \{[P(X \leq r, X \leq s) - P(X \leq r)P(X \leq s)]drds
\]
\[
\leq \int \{[P(X \leq r, X \leq s) - P(X \leq r)P(X \leq s)]drds
\]
\[
= \text{Cov}(X, X).
\]

Hence by substituting of (30) into (29), we have
\[
\text{Cov}(Z_i, Z_j) \leq \text{Cov}(X_i, X_j).
\]

So,
\[
\text{Cov}(I(T_i \leq s \leq Z_i), I(T_i \leq s \leq Z_i)) \leq M \text{Cov}^1(X_i, X_i),
\]

for some constant \( M > 0 \), such that it is not depending on \( i, j \) and \( x \). Thus, an application of Lemma 1 (i) yields, for all \( a \),
\[
\text{Cov}(x) \xrightarrow{a.s.} P(T \leq x \leq Z) = G(x)[1 - W(x)] = G(x)\bar{W}(x),
\]

but both \( 1 - C_n(x) \) and \( 1 - G(x)\bar{W}(x) \) are not d.f.'s on observation space \([a_\nu, \tau_\nu]\) and so we cannot use Glivenko-Cantelli theorem. In order to prove (8) and we use
\[
\text{G}_n(x) \xrightarrow{a.s.} G(x), \quad \forall x \in [a_\nu, \tau_\nu]
\]

and
\[
\text{G}_n(a) \xrightarrow{a.s.} G(a).
\]

Now by computing \( \frac{C_n(x)}{\text{G}_n(a)} \) and \( \frac{G(x)\bar{W}(x)}{G(a)} \), it is easy to show that \( 1 - \frac{C_n(x)}{\text{G}_n(a)} \) and \( 1 - \frac{G(x)\bar{W}(x)}{G(a)} \) are two d.f.'s on \([a_\nu, \tau_\nu] \) such that
\[
1 - \frac{C_n(x)}{\text{G}_n(a)} \xrightarrow{a.s.} 1 - \frac{G(x)\bar{W}(x)}{G(a)}, \quad \text{as } n \to \infty.
\]

So according to Glivenko-Cantelli theorem, we obtain
\[
\sup_{a_\nu \leq x \leq \tau_\nu} \left| \frac{C_n(x)}{\text{G}_n(a)} - \frac{G(x)\bar{W}(x)}{G(a)} \right| \xrightarrow{a.s.} 0, \quad \text{as } n \to \infty.
\]

By using
\[
(a_\nu + a)(b_\nu(x) - b(x)) = (a_\nu b_\nu(x) - ab(x)) + (ab(x) - a_\nu b(x)),
\]

and setting
\[
a_\nu = \frac{1}{G(a_\nu)}, \quad a = \frac{1}{G(a)}, \quad b_\nu(x) = C_n(x), \quad b(x) = G(x)\bar{W}(x),
\]

and with paying attention to this fact that two first terms above are greater than zero, we conclude that
\[
\int_{a_\nu}^a \frac{dA_n(s)}{\bar{B}(s)} \exp \left(- \int_{a_\nu}^s \frac{dA_n(s)}{\bar{B}(s)} \right),
\]

Then, (8) is achieved.

Next, Lemma 2 (i) implies that \( \{I(X_j > x \wedge Y_j), 1 \leq j \leq n\} \) is PA, for each fixed \( x \) on interval \([a_\nu, \tau_\nu]\). Therefore, an application of corollary of Theorem 1 in Sadikova [16] and (21) in Newman [13] again gives
\[
\text{Cov}(I(T_i \leq s \leq Z_i), I(T_j \leq s \leq Z_j)) = \text{Cov}(I(T_i \leq s \leq Z_i), I(T_j \leq s \leq Z_j))
\]

Hence, for all \( a_\nu \leq x \leq \tau_\nu \), using Lemma 1 (i),
\[
W_{in}(x) \xrightarrow{a.s.} P(X_i \leq x \wedge Y_i) = F_i(x), \quad \text{as } n \to \infty.
\]

Then based on similar idea for establishing (8), we have
\[
\sup_{a_\nu \leq x \leq \tau_\nu} \left| W_{in}(x) - F_i(x) \right| \xrightarrow{a.s.} 0, \quad \text{as } n \to \infty.
\]

This completes the proof of part (i).

Proof of part (ii) follows by an argument similar to one used in the proof of part (i) and utilizing Lemma 2 (ii).

**Lemma 4.** Let \((A, B), (A', B') \in \mathcal{H} \times \mathcal{B}\) and for any interval \([a_\nu, k]\) such that \(B(k) > 0\) and \(B'(k) > 0\), we show the supremum metric by \( \rho_\kappa \). If \( \max(\rho_\kappa(A, A'), \rho_\kappa(B, B')) \to 0 \) then
\[
\rho_\kappa(\Phi(A, B), \Phi(A', B')) \to 0.
\]

**Proof.** In order to prove, we must repeat the same lines of proof Lemma 2 of Gill [7] by using interval \([a_\nu, \kappa]\) for some \( \kappa < \tau_\nu \).

**Proof of Theorem 2.** To establish the asymptotic behavior of \( \hat{F}_i(\cdot) \) which is mentioned in this theorem, let us introduce some notations as follows.

Let \( A \) be a bounded, non-decreasing and right-continuous function on \([a_\nu, \tau_\nu]\) such that \( A(a_\nu) = 0 \) and \( \mathcal{A} \) be the set of all such functions. Also, \( B \) be a bounded, non-increasing and positive function on \([a_\nu, \tau_\nu]\) and \( \mathcal{B} \) be the set of all such functions.

For every pair \((A, B) \in \mathcal{A} \times \mathcal{B}\), we define some version of (6) in Gill [7]:
\[
\Phi(A, B)(x) = \prod_{i \in \mathbb{Z}} \left(1 - \frac{dA_i(x)}{B(x)} \right) \exp \left(- \int_{a_\nu}^x \frac{dA_i(s)}{B(s)} \right),
\]
where $A_c$ is the continuous part of $A$. We prove $\Phi(A,B)(\cdot)$ is a right-continuous, non-negative and non-increasing function on $(a_W, \tau_W)$ with $\Phi(A,B)(a_W) = 1$. Also, by setting $A(x) = F_1(x)$ and $B(x) = G(x)W(x)$.

Clearly, $(F_1, G W) \in \mathcal{A} \times \mathcal{B}$ and

$$\Phi(F_1, G W)(x) = \prod_{i \leq x}(1 - \frac{dF_i(s)}{G(s)W(s)}\exp(-\int_{a_W}^{x} \frac{dF_i(s)}{a(s)W(s)})).$$

Because $F_1$ is continuous, then $\Phi(F_1, G W)(x) = 1 - F(x)$. Moreover, by setting $A(x) = W_{1n}(x)$ and $B(x) = C_n(x)$, and since $(W_{1n}, C_n) \in \mathcal{A} \times \mathcal{B}$, it is easy to see $\Phi(W_{1n}, C_n)(x) = 1 - F_n(x)$. By Proposition 1, for every $\tau \leq \tau_W$ and $a_W \leq a$, we obtain (8) and (9). According to Lemma 4, we have

$$\sup_{a \in [0, \tau_W]} \left| \tilde{F}_n(x) - F_n(x) \right| \leq \frac{a}{n \rightarrow \infty},$$

for every $\tau < \tau_W$ and $a_W < a$ for which $F(\tau) < 1$ and $F(a) > 0$. We may now proceed for the proof of (12), in view of (11), it suffices to consider the case $F(\tau_W) = 0$ and $F(\tau_W) = 1$. Also, since $W(\tau_W) = 1$, then $Z_{n,m} < \tau_W$ and $a_W < Z_{1n}$. Hence, (12) is a consequence of (11). So, the proof is completed (see proof of Theorem 1.2 in Cai and Roussas [4], for more details).

**Discussion**

In the current manuscript, we prove the strong uniform convergence of TJW product-limit estimator for PA (NA) data under random left truncation and right censorship model. Most of the applied papers in the field of survival analysis would not assume the left truncation model, subjected to right censorship and dependency of the lifetime observations. Therefore, we try to solve this open problem in view of Cai and Roussas [4], which proved strong and weak convergence of product-limit estimator for PA (NA) data under right random censorship model. One application of the desired model used in this paper can be found in Shabani et al. [17], which are estimated point-wise confidence bound of survival and hazard functions for lung cancer patients under PA. For further research, we propose a new model in which the lifetimes, censorship and truncation sequences are dependent.

**Acknowledgement**

The authors would like to thank the editor and referees for their careful reading and constructive comments that improved presentation of the paper.

**References**


