Journal of Sciences, Islamic Republic of Iran 30(2): 177 - 185 (2019) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

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 Mathematic, 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

and Prakasa Rao [2] discussed the strong consistency and asymptotic normality of the empirical survival function and Yu [27] obtained weak convergence of the empirical process. Large sample properties of the product-limit estimator for the left truncated data in α-mixing observations were studied by Sun and Zhou [21]. Wang et al. [25] and Lee and Tsai [11] investigated the estimating of survival function when censoring variables are dependent. Wang [24] proved some properties of marginal survival functions for dependent censored data under an assumed Archimedean copula. Li and Patilea [12] proposed a dimension reduction approach for conditional product-limit estimators. Shen and Wang [18] investigated asymptotic properties of the product-limit estimator and Nelson-Allen estimator for censored negatively superadditive dependent data.

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 dependent lifetime.

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, δ_i) such that

$$Z_i = X_i \wedge Y_i \quad \textit{and} \quad \delta_i = I(X_i \leq Y_i),$$

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 sub-sequence of original data, but we show both sequences by same letters.

Let $\gamma \equiv P(T_1 \leq Z_1) > 0$ 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 $\Lambda(x) = \int_0^x (1 - F(u))^{-1} dF(u)$, we define

$$C(x) = P(T_1 \leq x \leq Z_1 \mid T_1 \leq Z_1) = \gamma^{-1} P(T_1 \leq x \leq Y_1) \times (1 - F(x)), \tag{1}$$

and

$$W(x) = P(Z_1 \leq x, \delta_1 = 1 \mid T_1 \leq Z_1) = \gamma^{-1} \int_0^x P(T_1 \leq u \leq Y_1) dF(u). \tag{2}$$

It can be shown that

$$\Lambda(x) = \int_0^x \frac{dW_1(u)}{C(u)}.$$

Let $C_n(x)$ and $W_{1n}(x)$ be the empirical estimators of $C(x)$ and $F_1(x)$, respectively, i.e.

$$C_n(x) = n^{-1}\sum_{i=1}^{n} I(T_i \le x \le Z_i) \tag{3}$$

and

$$W_{1n}(x) = n^{-1}\sum_{i=1}^{n} I(Z_i \le x, \delta_i = 1), \tag{4}$$

where $F_1(x) = P(Z_1 \le x, \delta_1 = 1)$ 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

$$\hat{F}_n(x) = \begin{cases} 1 - \prod_{Z_i < x}\left(1 - \frac{1}{nC_n(Z_i)}\right)^{\delta_i} & ; x < Z_{(n)} \\ 1 & ; x \ge Z_{(n)} \end{cases}. \tag{5}$$

The estimator of $\Lambda(x)$ which is comparable with Nelson-Aalen estimator of the cumulative hazard function for right censored data is

$$\hat{\Lambda}_n(x) = \int_0^x \frac{dW_{1n}(u)}{C_n(u)} = \sum_{i=1}^{n} \frac{I(Z_i \le x, \delta_i = 1)}{nC_n(Z_i)}. \tag{6}$$

The definition of the underlying dependence considered here is as follows.

Definition 1 *A finite family of random variables* $\{X_i, 1 \le i \le n\}$ *is said to be PA if, for every coordinate-wise non-decreasing functions* $f_1, f_2 : R^n \to R$, $E[f_1^2(X_i, 1 \le i \le n)] < \infty$ *and* $E[f_2^2(X_i, 1 \le i \le n)] < \infty$, *we have*

$$Cov(f_1(X_i, 1 \le i \le n), f_2(X_j, 1 \le j \le n)) \ge 0.$$

The above r.v.'s are said to be NA, if for every non-empty proper subset A *of* $\{1, 2, \dots, n\}$ *and for every coordinate-wise non-decreasing functions,* $f_1 : R^{Card(A)} \to R,\ f_2 : R^{Card(A^C)} \to R,$ $E[f_1^2(X_i, i \in A)] < \infty$ *and* $E[f_2^2(X_i, i \in A^c)] < \infty$, *we have*

$$Cov(f_1(X_i, i \in A), f_2(X_j, j \in A^c)) \le 0.$$

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). $\{X_n, n \ge 1\}$ 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 $\{Y_n, n \ge 1\}$ is iid r.v.'s with bounded density and independent of $\{X_n, n \ge 1\}$. The truncated sequence $\{T_n, n \ge 1\}$ is also iid r.v.'s and independent of $\{X_n, n \ge 1\}$ and $\{Y_n, n \ge 1\}$.

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

(A4). $\sum_{i=2}^{n} |Cov(X_1, X_i)|^{\frac{1}{3}} = O(n^{\alpha})$, for some $0 < \alpha \le 1$.

(A5). $\sum_{j=n+1}^{\infty} |Cov(X_1, X_j)|^{1/3} = O(n^{-(r-2)/2})$, 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 τ_F, τ_G, τ_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 τ_G, τ_H, a_G and a_H are defined in the same way.

We can find df of Z by $W(\cdot) = 1 - \overline{F}(\cdot)\overline{H}(\cdot)$. Then according to Stute and Wang [20], left and right endpoints of df W are

$$\tau_W = \tau_F \wedge \tau_H \quad and \quad a_W = a_F \wedge a_H. \tag{7}$$

Then under the current model, we assume that $a_G \le a_W$ and $\tau_G \le \tau_W$.

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) If $\{X_i, i \ge 1\}$ *has PA property and **(A3)** is satisfied, it holds*

$$\sup_{a_W \le x \le \tau_W} \left| C_n(x) - G(x)[1 - W(x)] \right| \xrightarrow{a.s.} 0 \quad as \quad n \to \infty, \tag{8}$$

and

$$\sup_{a_W \le x \le \tau_W} \left| W_{1n}(x) - F_1(x) \right| \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty. \tag{9}$$

(ii) If $\{X_n, n \ge 1\}$ *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* $a_W < \tau < \tau_W$ *, we have*

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

$$\sup_{a_W \le x \le \tau_W} \left| \hat{\Lambda}_n(x) - \Lambda(x) \right| \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty. \tag{10}$$

(ii) If $\{X_i, i \ge 1\}$ *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*

$$\sup_{a_W \le x \le \tau_W} \left| \hat{F}_n(x) - F(x) \right| \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty, \tag{11}$$

and

$$\sup_{Z_{1:n} \le x \le Z_{n:n}} \left| \hat{F}_n(x) - F(x) \right| \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty, \tag{12}$$

where

$$Z_{n:n} = \max_{1 \le i \le n} Z_i \quad and \quad Z_{1:n} = \min_{1 \le i \le n} Z_i .$$

Proof. See the Appendix.

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

$$\sup_{a \le x \le \tau} \left| \hat{\Lambda}_n(x) - \Lambda(x) \right| = O(n^{-\theta}) \quad a.s., \tag{13}$$

Where $0 < \theta < {}^{(r-2)}\!/_{(2r+2+\delta)}$*, 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* $a_W < a < \tau < \tau_W$ *, we have*

$$\sup_{a \le x \le \tau} \left| \hat{F}_n(x) - F(x) \right| = O(n^{-\theta}) \quad a.s., \tag{14}$$

where θ *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 $\hat{F}_n(\cdot)$, $\hat{\Lambda}_n(\cdot)$ and $\hat{\Lambda}_n^{(p)}(\cdot)$, 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:

$$\Sigma = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho & -\rho^2 \ldots & -\rho^{n-1} \\ -\rho & 1 & -\rho \ldots & -\rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ -\rho^{n-1} & -\rho^{n-2} & -\rho \ldots & 1 \end{pmatrix}. \tag{20}$$

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 Σ 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 $\hat{\Lambda}_n^{(p)}(\cdot)$.

$$d_{\hat{\Lambda}_n^{(p)}}(n) = \sup_{a_W \le x \le \tau_W} | \hat{\Lambda}_n^{(p)}(x) - \Lambda(x) |. \tag{21}$$

For smoothing the results, this process is iterated 1000 times, independently. So, the average value of $d_{\hat{F}_n}(n)$, $d_{\hat{\Lambda}_n}(n)$ and $d_{\hat{\Lambda}_n^{(p)}}(n)$ are indicated in Figure 1 which shows the plot of $d_{\hat{F}_n}(n)$ (Graph 1), $d_{\hat{\Lambda}_n}(n)$ and $d_{\hat{\Lambda}_n^{(p)}}(n)$ (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

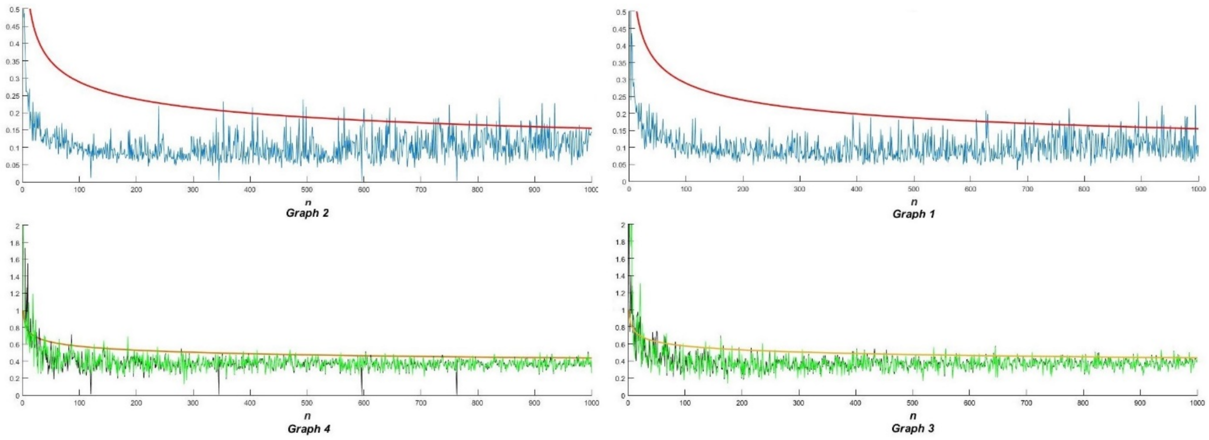

Figure 1. $d_{\hat{F}_n}(n)$ (blue line) and its convergence rate (red line) for $\theta = 0.27$ in NA (Graph 1) and PA (Graph 2) cases, $d_{\hat{\Lambda}_n}(n)$ (black line), $d_{\hat{\Lambda}_n^{(p)}}(n)$ (green line) and their convergence rates (yellow line) for $\theta = 0.12$ in NA (Graph 3) and PA (Graph 4) cases.

$\theta = 0.12$.

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 $\hat{\Lambda}_n^{(p)}(\cdot)$ is mostly located at the bottom of $\hat{\Lambda}_n(\cdot)$. Therefore, it is preferred to use $\hat{\Lambda}_n^{(p)}(\cdot)$ 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* $\{X_n, n \ge 1\}$ *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*

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

(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* $\underline{U} = (U_1, \dots, U_n)'$ *and* $\underline{V} = (V_1, \dots, V_m)'$ *be two independent random vectors. Then*

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

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

Lemma 3 *(Cai and Roussas, [4]) Let* $\{X_n, n \ge 1\}$ *be a sequence of r.v.'s where* $E(X_j) = 0$ *and*

Table 1. The values of $d_{\hat{F}_n}(n)$, $d_{\hat{\Lambda}_n}(n)$ and $d_{\hat{\Lambda}_n^{(p)}}(n)$ for different values of ρ and n in PA.

ρ	Estimator	n=20	n=30	n=50	n=100	n=200	n=500
0.1		0.2477	0.1538	0.1272	0.1018	0.0898	0.0885
0.2	Survival function	0.2525	0.1605	0.1317	0.1041	0.0907	0.0889
0.3		0.2625	0.1661	0.1403	0.1108	0.0994	.09457
0.1		0.8109	0.5689	0.4852	0.4160	0.3937	0.3914
0.2	Cumulative hazard	0.8105	0.5534	0.4699	0.4013	0.3728	0.3671
0.3	function	0.7995	0.5447	0.4444	0.3798	0.3494	0.3362
0.1		1.2254	0.5628	0.4752	0.4036	0.3844	0.3883
0.2	Cumulative hazard	1.3412	0.5618	0.4659	0.3953	0.3667	0.3646
0.3	function (plug-in)	0.9115	0.5422	0.4479	0.3766	0.3467	0.3350

Table 2. The values of $d_{\hat{F}_n}(n)$, $d_{\hat{\Lambda}_n}(n)$ and $d_{\hat{\Lambda}_n^{(p)}}(n)$ for different values of ρ and n in NA.

ρ	Estimator	n=20	n=30	n=50	n=100	n=200	n=500
0.1		0.1398	0.1187	0.1004	0.0655	0.05751	0.04526
0.2	Survival function	0.1552	0.1314	0.1302	0.1117	0.05435	0.03405
0.3		0.1795	0.1795	0.1244	0.07433	0.05867	0.04023
0.1		0.5629	0.4753	0.4343	0.2597	0.2421	0.2327
0.2	Cumulative hazard	0.6456	0.6106	0.4685	0.3288	0.2802	0.2047
0.3	function	0.36083	0.27931	0.21195	0.19888	0.17238	0.13241
0.1		0.5017	0.4921	0.3638	0.2578	0.2338	0.2117
0.2	Cumulative hazard	0.6737	0.6136	0.4666	0.3293	0.2872	0.2317
0.3	function (plug-in)	0.6917	0.4121	0.3807	0.2847	0.2123	0.1928

$|X_j| \le C < \infty$ for $j \ge 1$. Then

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

$$\sup_{k\ge1} \sum_{j:|j-k|\ge n} \left|Cov(X_j, X_k)\right| = O(n^{-(r-2)/2}), \qquad (26)$$

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

$$\sup_{m\ge0} E\left|\sum_{j=m+1}^{m+n} X_j\right|^r \le Bn^{r/2}. \qquad (27)$$

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

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

$$\begin{aligned}
Cov\left[I(T_i \le x \le Z_i), I(T_j \le x \le Z_j)\right] &= Cov\left[I(T_i \le x)I(x \le Z_i), I(T_j \le x)I(x \le Z_j)\right]\\
&= E\left[I(T_i \le x)\right]E\left[I(T_j \le x)\right]E\left[I(Z_i \ge x)I(Z_j \ge x)\right]\\
&\quad - E\left[I(T_i \le x)\right]E\left[I(T_j \le x)\right]E\left[I(Z_i \ge x)\right]E\left[I(Z_j \ge x)\right]\\
&\le Cov\left(I(Z_i \ge x), I(Z_j \ge x)\right)\\
&= Cov\left(I(Z_i < x), I(Z_j < x)\right)\\
&\le M\, Cov^{\frac{1}{3}}\left(Z_i, Z_j\right)
\end{aligned} \qquad (28)$$

Next, we will find an upper bound for $Cov(Z_i, Z_j)$ for all $i \ne j$. In order to do this, by independence of $\{X_i, i\ge1\}$ and $\{Y_i, i\ge1\}$, we obtain

$$\begin{aligned}
Cov\left(Z_i, Z_j\right) &= \int_{R^2} E\left((X_i \wedge y_1)(X_j \wedge y_2)\right) - E\left(X_i \wedge y_1\right)E\left(X_j \wedge y_2\right) dG(y_1)dG(y_2)\\
&= \int_{R^2} Cov\left((X_i \wedge y_1)(X_j \wedge y_2)\right) dG(y_1)dG(y_2)
\end{aligned} \qquad (29)$$

Let $f_y(x) = \min(x, y)$ for fixed y>0. So, $f_y(x)$ is a non-decreasing function of x on interval $[a_W, \tau_W]$. 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

Error! Bookmark not defined.

$$\begin{aligned}Cov(X_i \wedge y_1, X_j \wedge y_2) &= Cov\left[f_{y_1}(X_i), f_{y_2}(X_j)\right]\\ &= \int_0^{y_1}\int_0^{y_2}[P(X_i \le r, X_j \le s) - P(X_i \le r)P(X_j \le s)]drds\\ &\le \int_0^{\infty}\int_0^{\infty}\left[P(X_i \le r, X_j \le s) - P(X_i \le r)P(X_j \le s)\right]drds\\ &= \mathrm{Cov}(X_i, X_j). \end{aligned} \tag{30}$$

Hence by substituting of (30) into (29), we have

$$Cov(Z_i, Z_j) \le \mathrm{Cov}(X_i, X_j). \tag{31}$$

So,

$$Cov(I(T_i \le x \le Z_i), I(T_j \le x \le Z_j)) \le M\ \mathrm{Cov}^{\frac{1}{3}}(X_i, X_j), \tag{32}$$

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 $a_W \le x \le \tau_W$,

$$C_n(x) \xrightarrow{a.s.} P(T_1 \le x \le Z_1) = G(x)[1 - W(x)] = G(x)\overline{W}(x),$$

but both $1 - C_n(x)$ and $1 - G(x)\overline{W}(x)$ are not df's on observation space $[a_W, \tau_W]$ and so we cannot use Glivenko-Cantelli theorem. In order to prove (8) and we use

$$G_n(x) \xrightarrow{a.s.} G(x), \quad \forall\ x \in [a_W, \tau_W]$$

and

$$G_n(a_W) \xrightarrow{a.s.} G(a_W).$$

Now by computing $\frac{C_n(x)}{G_n(a_W)}$ and $\frac{G(x)\overline{W}(x)}{G(a_W)}$, it is easy to show that $1 - \frac{C_n(x)}{G_n(a_W)}$ and $1 - \frac{G(x)\overline{W}(x)}{G(a_W)}$ are two df's on $[a_W, \tau_W]$ such that

$$1 - \frac{C_n(x)}{G_n(a_W)} \xrightarrow{a.s.} 1 - \frac{G(x)\overline{W}(x)}{G(a_W)}, \qquad as \quad n \to \infty.$$

So according to Glivenko-Cantelli theorem, we obtain

$$\sup_{a_W \le x \le \tau_W}\left|\frac{C_n(x)}{G_n(a_W)} - \frac{G(x)\overline{W}(x)}{G(a_W)}\right| \xrightarrow{a.s.} 0, \qquad as \quad n \to \infty.$$

By using

$$(a_n + a)\left(b_n(x) - b(x)\right) = \left(a_n b_n(x) - ab(x)\right) + \left(ab_n(x) - a_n b(x)\right),$$

and setting

$$a_n = \frac{1}{G_n(a_W)}, \quad a = \frac{1}{G(a_W)}, \quad b_n(x) = C_n(x), \quad b(x) = G(x)\overline{W}(x),$$

and with paying attention to this fact that two first terms above are greater than zero, we conclude that

$$\left|C_n(x) - G(x)\overline{W}(x)\right| \le G_n(a_W)\left|\frac{C_n(x)}{G_n(a_W)} - \frac{G(x)\overline{W}(x)}{G(a_W)}\right| + G(x)\overline{W}(x)G_n(a_W)\left|\frac{1}{G(a_W)} - \frac{1}{G_n(a_W)}\right|.$$

Then, (8) is achieved.

Next, Lemma 2 (i) implies that $\{I(X_j > x \wedge Y_j), 1 \le j \le n\}$ is PA, for each fixed x on interval $[a_W, \tau_W]$. Therefore, an application of corollary of Theorem 1 in Sadikova [16] and (21) in Newman [13] again gives

$$\begin{aligned}Cov\left(I(X_i > x \wedge Y_i), I(X_j > x \wedge Y_j)\right) &= Cov\left(I(X_i \le x \wedge Y_i), I(X_j \le x \wedge Y_j)\right)\\ &= \int_0^{\infty}\int_0^{\infty} Cov\left(I(X_i \le x \wedge y_i), I(X_j \le x \wedge y_j)\right)dG(y_1)dG(y_2)\\ &\le M\ Cov^{\frac{1}{3}}(X_i, X_j). \end{aligned} \tag{33}$$

Hence, for all $a_W \le x \le \tau_W$, using Lemma 1 (i),

$$W_{1n}(x) \xrightarrow{a.s.} P(X_1 \le x \wedge Y_1) = F_1(x), \qquad as \quad n \to \infty.$$

Then based on similar idea for establishing (8), we have

$$\sup_{a_W \le x \le \tau_W}\left|W_{1n}(x) - F_1(x)\right| \xrightarrow{a.s.} 0, \qquad as \quad n \to \infty.$$

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 $[a_W, \kappa]$ such that $B(\kappa) > 0$ and $B'(\kappa) > 0$, we show the supremum metric by ρ_κ. If $\max\left(\rho_\kappa(A, A'), \rho_\kappa(B, B')\right) \longrightarrow 0$ then

$$\rho_k(\Phi(A, B), \Phi(A', B')) \longrightarrow 0.$$

Proof. In order to prove, we must repeat the same lines of proof Lemma 2 of Gill [7] by using interval $[a_W, \kappa]$ for some $\kappa < \tau_W$.

Proof of Theorem 2. To establish the asymptotic behavior of $\hat{F}_n(\ \cdot\ )$ 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 $[a_W,\ \tau_W)$ such that $A(a_W) = 0$ and $\mathcal{A}$ be the set of all such functions. Also, B be a bounded, non-increasing and positive function on $[a_W,\ \tau_W)$ 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]:

$$\Phi(A, B)(x) = \prod_{\{s \le x\}}\left(1 - \frac{dA(s)}{B(s)}\right)\ exp\left(-\int_{a_W}^{x}\frac{dA_C(s)}{B(s)}\right),$$

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

$$A(x) = F_1(x) \qquad and \qquad B(x) = G(x)\overline{W}(x).$$

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

$$\Phi(F_1, G\,\overline{W}\,)(x) = \prod_{\{s \le x\}}\left(1 - \frac{dF_1(s)}{G(s)\,\overline{W}(s)}\right) exp\left(-\int_{a\_W}^{x} \frac{dF_{1c}(s)}{G(s)\,\overline{W}(s)}\right).$$

Because F_1 is continuous, then $\Phi(F_1, G\,\overline{W}\,)(x) = 1 - F(x)$. Moreover, by setting

$$A(x) = W_{1n}(x) \qquad and \qquad B(x) = C_n(x),$$

and since $(W_{1n}, C_n) \in \mathcal{A} \times \mathcal{B}$, it is easy to see $\Phi(W_{1n}, C_n)(x) = 1 - \hat{F}_n(x)$. By Proposition 1, for every $\tau \le \tau_W$ and $a_W \le a$, we obtain (8) and (9). According to Lemma 4, we have

$$\sup_{a \le x \le \tau}\left|\hat{F}_n(x) - F(x)\right| \xrightarrow{a.s.} 0, \qquad as \quad n \to \infty,$$

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 $F(a_W) = 0$ and $F(\tau_W) = 1$. Also, since $W(\tau_W) = 1$, then $Z_{n:n} < \tau_W$ and $a_W < Z_{1:n}$. 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

1. Alam K. and Saxena K.M.L. Positive dependent in multivariate distributions. *Communication in Statistics-Theory and Mathods*. **10**(12): 1183–1196 (1981).
2. Bagai I. and Parakasa Rao B.L.S. Estimation of the survival function for stationary associated processes. *Statistics & Probability Letters*. **12**: 385–391 (1991).
3. Brindley Jr E.C. and Thompson Jr W.A. Dependence and aging aspects of multivariate survival. *Journal of the American Statistical Association*. **23**(2): 301–306 (1972).
4. Cai Z.W. and Roussas G.G. Kaplan-Meier estimator under association. *Journal of Multivariate Analysis.* **67**: 318–348 (1998).
5. Gijbels I. and Wang J.L. Strong Representations of the Survival Function Estimator for Truncated and Censored Data with Applications. *Journal of Multivariate Analysis.* **47**: 210–229 (1993).
6. Gill R.D. Censoring and Stochastic Integrals. *Mathematical Centre tracts, Mathematish Centrum, Amsterdam*, No**. 124** (1980).
7. Gill R.D. Testing with replacement and the product limit estimator. *The Annals of Statistics*. **9**: 853–860 (1981).
8. Gill R.D. Large sample behavior of the product limit estimator on the whole line. *The Annals of Statistics.* **11**: 49–56 (1983).
9. Ghosh M. Multivariate negative dependence. *Communications in Statistics-Theory and Methods.* **10**(4): 307-337 (1981).
10. Joag-Dev K. and Proschan F. Negative association of random variables, with application. *The Annals of Statistics*. **11**: 286–295 (1983).
11. Lee S.Y. and Tsai W.Y. An estimator of the survival function based on the Semi-Markov model under dependent censoring. *Lifetime data analysis.* **11**: 193–211 (2005).
12. Li W. and Patilea V. A dimension reduction approach for conditional Kaplan-Meier estimators. *TEST*. (2015).
13. Newman C.M. Normal fluctuations and the FKG inequalities. *Communications in Mathematical.* **74**: 119–128 (1980).
14. Roussas G.G. Kernel estimates under association: Strong uniform consistency. *Statistics & Probability Letters.* **12**: 393–403 (1991).
15. Roussas G.G. Asymptotic normality of random fields of positively or negatively associated processes. *Journal of Multivariate Analysis*. **50**: 152–173 (1994).
16. Sadikova S.M. Two-dimensional analogies of an inequality of Esseen with applications to the central limit theorem. *Theory of Probability & Its Application.* **11**: 325–335 (1996).
17. Shabani A.H., Jabbari H. and Fakoor V. Weak convergence of TJW product-limit estimator under association. *Stat*. ***8***(1), e217 (2019).
18. Shen A. and Wang X. Kaplan-Meier estimator and hazard estimator for censored negatively superadditive dependent data. *Statistics*. **5**: 377–388 (2016).
19. Stute W. Convergence of Kaplan-Meier estimator in weighted sup-norms. *Statistics & Probability Letters*. **20**:

219–223 (1994).
20. Stute W. and Wang J.L. A strong law under random censorship. *The Annals of Statistics*. **21**: 1591–1607 (1993).
21. Sun L. and Zhou X. Survival function and density estimation for truncated dependent data. *Statistics & Probability Letters*. **52**: 47–57 (2001).
22. Tsai W.Y., Jewell N.P. and Wang M.C. A note on the product-limit estimator under right censoring and left truncation. *Biometrika*. **74**: 883–886 (1987).
23. Vershynin, R. High-dimensional probability: An introduction with applications in data science. Cambridge University Press (2018).
24. Wang A. Properties of the marginal survival functions for dependent censored data under an assumed Archimedean copula. *Journal of Multivariate Analysis.* **129**: 57–68 (2014).
25. Wang C., Sun J., Sun L., Zhou J. and Wang D. Nonparametric estimation of current status data with dependent censoring. *Lifetime Data Analysis.* **18**: 434–445 (2012).
26. Ying Z. and Wei L.J. The Kaplan-Meier estimate for dependent failure time observations. *Journal of Multivariate Analysis.* **50**: 17–29 (1994).
27. Yu H. A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences. *Probability theory and related fields*. **95**: 357–370 (1993).
28. Zhou Y. A note on the TJW product-limit estimator for truncated and censored data. *Statistics & Probability Letters*. **26**: 381–387 (1996).