Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

Strong representation of the Kaplan–Meier and hazard estimators for censored data with *m*-widely acceptable dependence

Ikhlasse Chebbab

To cite this article: Ikhlasse Chebbab (28 May 2026): Strong representation of the Kaplan–Meier and hazard estimators for censored data with *m*-widely acceptable dependence, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2026.2679126

To link to this article: https://doi.org/10.1080/24754269.2026.2679126

Published online: 28 May 2026.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2026.2679126

Strong representation of the Kaplan–Meier and hazard estimators for censored data with *m*-widely acceptable dependence

Ikhlasse Chebbab [a,b]

[a]National Higher School of Telecommunications and Information and Communication Technologies, Oran, Ageria; [b]Laboratory of Statistics and Stochastic Processes, Djillali Liabes University, Sidi Bell Abbes, Algeria

ABSTRACT

This paper investigates the asymptotic properties of the Kaplan–Meier and hazard estimators for censored survival time data. We conduct this analysis under the assumption of m-widely acceptable (m-WA) dependence, a generalized form of weak correlation. Using the Fuk–Nagaev inequality, we establish strong consistency and strong representation results for these estimators. Our findings show that the rate of strong consistency is near $O\left(\sqrt{\frac{g(n)\log n}{n}}\right)$ and the remainder term in the strong representation is of the same order. These results generalize and extend existing work for other types of dependent data, such as linearly extended negative quadrant-dependent (LENQD) and extended negative dependent (END) sequences, thereby broadening the theoretical foundation for these widely used statistical tools.

ARTICLE HISTORY

Received 6 August 2025
Revised 22 December 2025
Accepted 20 May 2026

KEYWORDS

Kaplan–Meier estimator; strong consistency; m-WA data; strong representation

1. Introduction

Survival analysis has emerged as a fundamental statistical methodology for time-to-event data across diverse scientific disciplines (Klein & Moeschberger, 2003). The Kaplan–Meier (Kaplan & Meier, 1958) and Nelson–Aalen (Aalen, 1978; Nelson, 1969) estimators have become cornerstone techniques for nonparametric survival estimation, particularly for right-censored observations. While classical asymptotic theory typically assumes independent and identically distributed data (Fleming & Harrington, 1991), modern applications increasingly reveal complex dependence structures that challenge this foundational assumption (Andersen et al., 1993). The theoretical framework for dependent survival data has evolved significantly since early work by Liebscher (2002), who established convergence rates under strong mixing conditions. Subsequent developments by Q. Y. Wu and Chen (2013) extended these results to negatively associated data, while Y. M. Li and Zhou (2020, 2024) advanced the theory for widely orthant and extended negatively dependent observations. These contributions collectively demonstrated the adaptability of survival estimators to various dependence

CONTACT Ikhlasse Chebbab ikhlasse.chebbab@ensttic.dz, chebbab-ikhlasse@outlook.fr National Higher School of Telecommunications and Information and Communication Technologies, Bp 1504, Oran, 31000, Algeria; Laboratory of Statistics and Stochastic Processes, Djillali Liabes University, Bp 89, Sidi Bell Abbes, 22000, Algeria

structures (Anevski, 2017). Despite these advances, two critical limitations persist in contemporary survival analysis. First, traditional linear expectation operators often fail to capture the distributional uncertainty prevalent in modern applications (Peng, 2006). Second, conventional dependence structures may impose unrealistic constraints on complex real-world data (Y. Wu et al., 2023). These challenges are particularly evident in clustered clinical trials (Shen & Wang, 2016), engineering reliability studies (Nematolahi et al., 2020) and longitudinal epidemiological research (Ahmed & Flandre, 2020).

The need for advanced dependence modelling is underscored by specific, high-impact application scenarios. For instance, in multicentre oncology trials, patient outcomes (e.g., time to progression) within the same medical centre are often positively dependent due to shared protocols, clinician expertise or local environmental factors, resulting in clustered, censored failure-time data. Similarly, in reliability engineering, the lifetimes of multiple components within a single system (e.g., an aircraft engine or a power grid module) are subject to common random stress loads, inducing a positive association structure in their failure times, which are frequently observed under type-I or type-II censoring. In these and other settings, the m-widely acceptable (m-WA) dependence structure provides a flexible and theoretically sound framework to capture the complex positive and negative associations that simpler models may miss. Developing robust nonparametric estimators under m-WA dependence is therefore not only a theoretical advancement but a practical necessity for accurate inference in these fields.

Recent work by Y. Li et al. (2024) has significantly advanced our understanding of the Kaplan–Meier and hazard estimators for censored linearly extended negatively quadrant dependent (LENQD) data. Building upon these foundations, the present work considers the more general framework of m-widely acceptable (m-WA) dependence. It is important to emphasize the hierarchical relationship between these structures: the LENQD class is a specific subset of the m-WA class. This framework is particularly robust as it encompasses negatively associated (NA) and extended negatively dependent (END) random variables as special cases. While LENQD requires a global linear dominating coefficient to bound joint probabilities, the m-WA structure allows for a more flexible partitioning of the sequence into m blocks, each satisfying widely acceptable properties. Consequently, our results provide a unified theoretical framework that not only encompasses the recent findings for LENQD data but also extends applicability to systems with shifting or non-linear dependence structures.

The Kaplan–Meier estimator and the hazard rate estimator are fundamental tools in survival analysis, particularly for handling censored data. While much of the literature focuses on independent or weakly dependent data, there is growing interest in extending these results to more general dependence structures.

We have n independent and identically distributed (i.i.d.) survival times, $X_1, \ldots, X_n$, which represent the time until an event occurs. These times come from a continuous probability distribution with an unknown distribution function $\Gamma(x) = P(X_i \leq x)$ and $\Gamma(0) = 0$. Additionally, we have n i.i.d. censoring times, $Y_1, \ldots, Y_n$, which also follow their own unknown continuous distribution function $G(y) = P(Y_i \leq y)$, with $G(0) = 0$.

A fundamental assumption is that the survival times $\{X_i\}$ and censoring times $\{Y_i\}$ are independent of each other. So given the observed data (U_i, δ),

$$U_i = X_i \wedge Y_i \quad \text{and} \quad \delta_i = I(X_i \leq Y_i), \tag{1}$$

where $a \wedge b$ signifies the minimum and $I(B)$ is the indicator of event B.

In survival analysis, the count of uncensored events up to time u is given by $N_n(u) = \sum_{i=1}^{n} I(U_i \le u, \delta_i = 1) = \sum_{i=1}^{n} I(X_i \le u \wedge Y_i)$, where $U_i = \min(X_i, Y_i)$ represents observed times, X_i event times, Y_i censoring times and δ_i event indicators. The at-risk population at time u includes all uncensored or censored observations still under study: $M_n(u) = \sum_{i=1}^{n} I(U_i \ge u)$. The Kaplan–Meier estimator (Kaplan & Meier, 1958) for the cumulative distribution function $\Gamma(x)$ is constructed as $\widehat{\Gamma}_n(x) = 1 - \prod_{v \le x}(1 - \frac{\mathrm{d}N_n(v)}{M_n(v)})$, with $\mathrm{d}N_n(v) = N_n(v) - N_n(v^-)$ capturing event jumps at time s. When $\Gamma(x)$ admits a density $f(x)$, the hazard rate $\lambda(x)$ is defined as $\lambda(x) = \frac{\mathrm{d}}{\mathrm{d}x}[-\log \overline{\Gamma}(x)] = \frac{f(x)}{\overline{\Gamma}(x)}$, where survival function is represented by $\overline{\Gamma}(x)$, which is defined as $1 - \Gamma(x)$, characterizing the instantaneous event risk conditional on survival up to time x.

The distribution function $K(u)$ and its empirical counterpart $K_n(u)$ for the observed times $\{U_i\}$ are

$$K(u) = 1 - \overline{\Gamma}(u)\overline{G}(u) = 1 - [1 - \Gamma(u)][1 - G(u)], \quad \overline{K}(u) = 1 - K(u), \tag{2}$$

$$K_n(u) = \frac{1}{n}\sum_{i=1}^{n} I(U_i < u) = 1 - \frac{M_n(u)}{n}, \quad \overline{K}_n(u) = 1 - K_n(u). \tag{3}$$

The cumulative hazard function, denoted as $\Upsilon(x)$, has an estimated version called the Nelson–Aalen estimator, $\widehat{\Upsilon}_n(x)$:

$$\Upsilon(x) = -\log \overline{\Gamma}(x) = \int_0^x \frac{\mathrm{d}\Gamma(u)}{\overline{\Gamma}(u)} = \int_0^x \frac{\mathrm{d}\Gamma_*(u)}{\overline{K}(u)}, \tag{4}$$

$$\widehat{\Upsilon}_n(x) = \int_0^x \frac{\mathrm{d}N_n(u)}{M_n(u)} = \int_0^x \frac{\mathrm{d}\Gamma_{*n}(u)}{\overline{K}_n(u)}, \tag{5}$$

where $F_*(u)$ and its empirical version $F_{*n}(u)$ are defined as

$$\Gamma_*(u) = P(U_1 \le u, \delta_1 = 1) = \int_0^{\infty} F(u \wedge y)\,\mathrm{d}G(y) = \int_0^u \overline{G}(y)\,\mathrm{d}\Gamma(y), \tag{6}$$

$$\Gamma_{*n}(u) = \frac{1}{n}\sum_{i=1}^{n} I(U_i \le u, \delta_i = 1) = \frac{N_n(u)}{n}. \tag{7}$$

As $N_n(u)$ is a step function that increases by $\delta_{(k)}$ at each ordered failure time $U_{(k)}$, the estimators can be expressed in a summation form. Specifically, $\widehat{\Upsilon}_n(x)$ is calculated as the sum of $I(U_{(k)} \le x, \delta_{(k)} = 1)$ divided by $(n - k + 1)$ for k from 1 to n. Similarly, $\widehat{F}_n(x)$ is determined by 1 minus the product of $(1 - \frac{\delta_{(k)}}{n-k+1})$ raised to the power of $I(U_{(k)} \le x, \delta_{(k)} = 1)$ for i from 1 to n, where $\delta_{(k)}$ is the censoring indicator for the kth ordered observation $U_{(k)}$. This paper's main achievement is establishing strong consistency and strong representation results for both the Kaplan–Meier estimator and the hazard rate estimator under conditions of m-WA dependence.

For this, we use the Fuk–Nagaev inequality, which is particularly well-suited to deal with sums of dependent random variables with complex structures. In this context, we assume that the underlying distribution functions are continuous, ensuring there are no ties in the observed data (i.e., all U_i are distinct with probability one).

2. Preliminary foundations

This section outlines fundamental definitions and preliminary lemmas that serve as the basis for proving the main results.

Definition 2.1 (Y. Wu et al., 2023): Let $(\Omega, \mathcal{H}, \widehat{\mathbb{E}})$ be a sublinear expectation space. A sequence $\{X_k\}_{k\geq 1}$ of random variables is called *widely acceptable (WA)* if there exists a positive nondecreasing sequence $\{g(n)\}_{n\geq 1}$ of dominating coefficients satisfying

$$\sum_{n=1}^{\infty} \frac{g(n)}{n} < \infty,$$

and such that for every $n \geq 1$, any nonnegative constants $\{a_{nk}\}_{1\leq k\leq n}$, and any family $\{\phi_k\} \subset C_{b,\mathrm{Lip}}(\mathbb{R})$ of uniformly monotone (i.e., either all nondecreasing or all nonincreasing) bounded Lipschitz functions,

$$\widehat{\mathbb{E}}\left[\exp\left(\sum_{k=1}^{n} a_{nk}\phi_k(X_k)\right)\right] \leq g(n) \prod_{k=1}^{n} \widehat{\mathbb{E}}\left[\exp\left(a_{nk}\phi_k(X_k)\right)\right].$$

Definition 2.2 (Y. Wu et al., 2023): For a given fixed integer $m \geq 1$, a sequence of random variables $\{X_k\}_{k\geq 1}$ is termed m-widely acceptable (m-WA) if it satisfies the following criterion: for every integer $k \geq 2$, and for any set of indices $n_1, \ldots, n_k$ where the minimum absolute difference between any two distinct indices is at least m (i.e., $\min_{j\neq l} |n_j - n_l| \geq m$), the subset of random variables $\{X_{n_1}, \ldots, X_{n_k}\}$ constitutes a widely acceptable (WA) family.

Remark 2.1: It is clear that m-WA random variables are a more general category than WA random variables. By definition, the subsequences

$$\{X_1, X_{1+m}, X_{1+2m}, \ldots\}, \{X_2, X_{2+m}, X_{2+2m}, \ldots\}, \ldots, \{X_m, X_{2m}, X_{3m}, \ldots\}$$

are **WA**. When $m = 1$, ***m*-WA** reduces to **WA**. Since the class of ***m*-WA** random variables generalizes several dependence concepts including **negatively dependent (ND)**, **extended negatively dependent (END)**, **widely negative orthant dependent (WOD)**, ***m*-END** and ***m*-WOD**, research on their probability limit theories holds considerable importance and broad relevance.

Lemma 2.3 (Y. Wu et al., 2023): *Consider a sequence $\{X_k\}_{k\geq 1}$ of m-dependent Widely Acceptable (m-WA) random variables defined on a sublinear expectation space $(\Omega, \mathcal{H}, \widehat{\mathbb{E}})$, characterized by the dominating coefficients $\{g(n)\}_{n\geq 1}$ (satisfying $\sum_{n=1}^{\infty} g(n)/n < \infty$). If $\{\varphi_k(\cdot), k \geq 1\} \subset C_{b,\mathrm{Lip}}(\mathbb{R})$ is a sequence of functions that are either all nondecreasing or all nonincreasing, then the transformed sequence $\{\varphi_k(X_k), k \geq 1\}$ remains m-WA with the same dominating coefficients $\{g(n)\}_{n\geq 1}$.*

Lemma 2.4 (Y. Wu et al., 2023): *Consider a sequence $\{X_k\}_{k\geq 1}$ of m-widely acceptable (m-WA) random variables in the sublinear expectation space $(\Omega, \mathcal{H}, \widehat{E})$, where $\widehat{E}[X_k] \leq 0$ for all*

$k \geq 1$. Then, for any threshold $x > 0$ and parameter $d > 0$, the tail probability obeys the following bound:

$$\mathbb{V}\left(\sum_{k=1}^{n} X_k > x\right) \leq m\mathbb{V}\left(\max_{1\leq k\leq n} X_k > \frac{d}{m}\right) + mg(n)\exp\left(\frac{x}{d} - \frac{x}{d}\ln\left(1 + \frac{xd/m}{2\sum_{k=1}^{n}\widehat{E}|X_k|^2}\right)\right).$$

Lemma 2.5: *For a fixed integer $m \geq 1$, if two sequences of m-WA random variables, $\{X_i\}_{i\geq 1}$ and $\{Y_i\}_{i\geq 1}$, are independent within the sublinear expectation space $(\Omega, \mathcal{H}, \widehat{E})$, then their combined sequence $\{Z_k\}_{k\geq 1}$ (where each Z_k is drawn from either an X_i or a Y_j) also possesses the m-WA property.*

Proof: Let $\{Z_k\}_{k\geq 1}$ be the combined sequence formed by interleaving the sequences $\{X_i\}_{i\geq 1}$ and $\{Y_j\}_{j\geq 1}$. To establish that $\{Z_k\}_{k\geq 1}$ is m-widely acceptable (m-WA), we must demonstrate that for any arbitrarily chosen subset $\{Z_{n_1}, \ldots, Z_{n_k}\}$ whose indices satisfy the minimum spacing condition $\min_{p\neq q}|n_p - n_q| \geq m$, this subset itself satisfies the widely acceptable (WA) condition.

We proceed by analysing two distinct cases based on the composition of the selected subset.

Case 1: Homogeneous Subset. If all random variables $\{Z_{n_i}\}_{i=1}^{k}$ within the chosen subset originate exclusively from one of the parent sequences (i.e., either all from $\{X_i\}$ or all from $\{Y_j\}$), then the inherent m-WA property of the original sequences directly applies. Consequently, there exists a dominating coefficient, denoted as $g_1(k)$ (or $g_2(k)$ if from $\{Y_j\}$), such that the WA inequality is satisfied:

$$\widehat{\mathbb{E}}\left[\exp\left(\sum_{i=1}^{k} a_i\phi_i(Z_{n_i})\right)\right] \leq g_1(k)\prod_{i=1}^{k}\widehat{\mathbb{E}}[\exp(a_i\phi_i(Z_{n_i}))].$$

Case 2: Mixed Subset. This case addresses subsets $\{Z_{n_i}\}_{i=1}^{k}$ that contain elements drawn from both the $\{X_i\}$ and $\{Y_j\}$ sequences. Let $I = \{i : Z_{n_i}\text{originated from } \{X_i\}\}$ and $J = \{j : Z_{n_j}\text{originated from } \{Y_j\}\}$ be the respective sets of indices. Due to the $\min_{p\neq q}|n_p - n_q| \geq m$ spacing condition imposed on the combined sequence, it inherently follows that the sub-collections $\{X_i\}_{i\in I}$ and $\{Y_j\}_{j\in J}$ also maintain this minimum spacing within their indices.

Crucially, leveraging the m-WA property for each individual sequence and the independence between the $\{X_i\}$ and $\{Y_j\}$ sequences, we can factorize the sublinear expectation of the exponential sum:

$$\widehat{\mathbb{E}}\left[\exp\left(\sum_{i=1}^{k} a_i\phi_i(Z_{n_i})\right)\right] = \widehat{\mathbb{E}}\left[\exp\left(\sum_{i\in I} a_i\phi_i(X_i)\right)\right]\widehat{\mathbb{E}}\left[\exp\left(\sum_{j\in J} a_j\phi_j(Y_j)\right)\right].$$

By then applying the WA condition to each factored term, the above equation satisfies

$$\leq g_1(|I|)g_2(|J|)\prod_{i\in I}\widehat{\mathbb{E}}[\exp(a_i\phi_i(X_i))]\prod_{j\in J}\widehat{\mathbb{E}}[\exp(a_j\phi_j(Y_j))].$$

This expression can be further bounded to satisfy the general WA condition for the combined sequence:

$$\leq g(k) \prod_{i=1}^{k} \widehat{\mathbb{E}}[\exp(a_i \phi_i(Z_{n_i}))].$$

Here, $g(k)$ serves as the new dominating coefficient for the combined sequence, defined as $g(k) = \max\{g_1(|I|)g_2(|J|) : |I| + |J| = k\}$, where the maximum is taken over all possible partitions of the k elements into subsets I and J.

In both cases, the WA condition holds for all sufficiently spaced subsets of $\{Z_k\}$, proving the m-WA property is preserved. ■

Lemma 2.6: *Let $\{X_n\}_{n\geq 1}$ be a sequence of m-WA random variables in the sublinear expectation space $(\Omega, \mathcal{H}, \widehat{\mathbb{E}})$, sharing common dominating coefficients $\{g(n)\}_{n\geq 1}$. We assume $g(n)$ is a positive sequence such that $\sum_{n=1}^{\infty} g(n)/n < \infty$. Assume the distribution function $\Gamma(x) = \widehat{\mathbb{E}}[\mathcal{I}(X_n \leq x)]$ has a density $f(x)$ that is both bounded and bounded away from zero over the support of X_n, i.e., there exist constants $C \geq c > 0$ such that $c \leq f(x) \leq C$. Defining the empirical process $\Gamma_n(x) = n^{-1} \sum_{i=1}^{n} \mathcal{I}(X_i \leq x)$ and the scaling factor τ_n as*

$$\tau_n = \sqrt{\frac{g(n)\log n}{n}},$$

we establish the uniform convergence:

$$\sup_x |\Gamma_n(x) - \Gamma(x)| = O(\tau_n) \quad \textit{quasi-surely.}$$

Proof: For $n \geq 3$, choose points $\{x_{n,k}\}_{k=1}^{n-1}$ such that

$$\Gamma(x_{n,k}) = \frac{k}{n}.$$

Since $c \leq f(x) \leq C$, the Mean Value Theorem yields

$$\frac{1}{nC} \leq x_{n,k+1} - x_{n,k} \leq \frac{1}{nc}.$$

Hence,

$$\max_{1\leq k\leq n-2} |x_{n,k+1} - x_{n,k}| = O\left(\frac{1}{n}\right).$$

For any $x \in \mathbb{R}$, let k satisfy $x_{n,k} \leq x < x_{n,k+1}$. Then

$$\begin{aligned}\Gamma_n(x) - \Gamma(x) &\leq \Gamma_n(x_{n,k+1}) - \Gamma(x_{n,k}) \\ &= \big[\Gamma_n(x_{n,k+1}) - \Gamma(x_{n,k+1})\big] + \big[\Gamma(x_{n,k+1}) - \Gamma(x_{n,k})\big].\end{aligned}$$

Because $|\Gamma(x_{n,k+1}) - \Gamma(x_{n,k})| \leq C|x_{n,k+1} - x_{n,k}| = O(1/n)$, we obtain $\Gamma_n(x) - \Gamma(x) \leq \max_k |\Gamma_n(x_{n,k}) - \Gamma(x_{n,k})| + O(\frac{1}{n})$, and the same bound holds from below. Consequently,

$$\sup_x |\Gamma_n(x) - \Gamma(x)| \leq \max_{1\leq k\leq n-1} |\Gamma_n(x_{n,k}) - \Gamma(x_{n,k})| + O\left(\frac{1}{n}\right). \tag{8}$$

Fix k and define $\xi_{j,k} = I(X_j \le x_{n,k}) - \mathbb{E}[I(X_j \le x_{n,k})]$. Then $|\xi_{j,k}| \le 2$, $\mathbb{E}[\xi_{j,k}] \le 0$, and $\sum_{j=1}^{n} \mathbb{E}[|\xi_{j,k}|^2] \le n$.

Applying Lemma 2.4 (exponential inequality for m-WA sums) with $S_n = \sum_{j=1}^{n} \xi_{j,k}$, $x = \varepsilon n \tau_n$, $d = 2m\varepsilon\tau_n$, where $\varepsilon > 0$ is arbitrary, we obtain for large n

$$\mathbb{V}\left(S_n > \varepsilon n \tau_n\right) \le m\, g(n) \exp\left(-\frac{\varepsilon^2 \log n}{2}\right) = m\, g(n)\, n^{-\varepsilon^2/2}. \tag{9}$$

The same bound holds for $\mathbb{V}(-S_n > \varepsilon n \tau_n)$ by symmetry.

We obtain the union bound over all grid points from (9),

$$\begin{aligned}\mathbb{V}\left(\max_k |\Gamma_n(x_{n,k}) - \Gamma(x_{n,k})| > \varepsilon\tau_n\right) &\le \sum_{k=1}^{n-1} \mathbb{V}\left(|S_n^{(k)}| > \varepsilon n \tau_n\right) \le 2n \cdot m\, g(n)\, n^{-\varepsilon^2/2} \\ &= 2m\, g(n)\, n^{1-\varepsilon^2/2}, \end{aligned} \tag{10}$$

where $S_n^{(k)} = \sum_{j=1}^{n} \xi_{j,k}$.

Choose ε sufficiently large so that $\varepsilon^2/2 > 1$. Then $n^{1-\varepsilon^2/2} = O(n^{-\delta})$ for some $\delta > 0$. Since $\sum g(n)/n < \infty$, we have $\sum_{n=1}^{\infty} g(n)\, n^{1-\varepsilon^2/2} < \infty$.

From (8) and the fact that $1/n = o(\tau_n)$ (because $\tau_n = \sqrt{g(n) \log n/n}$ and $g(n) \ge 1$), there exists $\varepsilon' > \varepsilon$ such that for large n,

$$\mathbb{V}\left(\sup_x |\Gamma_n(x) - \Gamma(x)| > \varepsilon'\tau_n\right) \le \mathbb{V}\left(\max_k |\Gamma_n(x_{n,k}) - \Gamma(x_{n,k})| > \varepsilon\tau_n\right). \tag{11}$$

Combining (10) and (11) yields $\sum_{n=1}^{\infty} \mathbb{V}(\sup_x |\Gamma_n(x) - \Gamma(x)| > \varepsilon'\tau_n) < \infty$.

By the sublinear Borel–Cantelli lemma, $\sup_x |\Gamma_n(x) - \Gamma(x)| \le \varepsilon'\tau_n$ quasi-surely for all large n, which is precisely $\sup_x |\Gamma_n(x) - \Gamma(x)| = O(\tau_n)$ quasi-surely. ■

3. Main results

Theorem 3.1: *Consider two sequences $\{X_n\}_{n\ge1}$ and $\{Y_n\}_{n\ge1}$ of m-Widely Acceptable (m-WA) random variables in a sublinear expectation space $(\Omega, \mathcal{H}, \widehat{\mathbb{E}})$, sharing common dominating coefficients $\{g(n)\}_{n\ge1}$ (satisfying $\sum_{n=1}^{\infty} g(n)/n < \infty$). Define the rate function $\tau_n = \sqrt{\frac{g(n)\log n}{n}}$. Then, for any threshold parameter τ satisfying $0 < \tau < \tau_K$ (where τ_K is a critical bound), the following holds:*

$$\sup_{0\le x\le\tau} \left|\widehat{\Upsilon}_n(x) - \Upsilon(x)\right| = O(\tau_n) \quad q.s., \tag{12}$$

$$\sup_{0\le x\le\tau} \left|\widehat{\Gamma}_n(x) - \Gamma(x)\right| = O(\tau_n) \quad q.s. \tag{13}$$

Proof: We first establish that the transformed sequences, specifically the observed times $\{U_i = X_i \wedge Y_i\}$ and censoring indicators $\{\delta_i = I(X_i \le Y_i)\}$, retain the m-WA property. This follows from Lemma 2.3, as their transformations from (X_i, Y_i) are coordinate-wise monotonic, and Lemma 2.5 further guarantees that their union is m-WA with the same dominating coefficients $\{g(n)\}$. Consequently, applying Lemma 2.6 to these m-WA sequences, $\{I(U_i \le x, \delta_i = 1)\}$ and $\{I(U_i \ge x)\}$, we obtain the uniform convergence rates:

$$\sup_{x\ge 0} |\widehat{\Gamma}_n^*(x) - \Gamma^*(x)| = O(\tau_n) \quad \text{q.s. and} \sup_{x\ge 0} |\widehat{K}_n(x) - K(x)| = O(\tau_n) \quad \text{q.s.,}$$

where $\Gamma^*(x) = E[I(U_i \le x, \delta_i = 1)]$ and $K(x) = E[I(U_i \ge x)]$. To analyse the total error $\widehat{\Upsilon}_n(x) - \Upsilon(x)$, we establish a structural algebraic decomposition by adding and subtracting the term $\int_0^x \widehat{K}_n^{-1} \mathrm{d}\Gamma^*$. This identity isolates the error of the counting process from the error in the at-risk weights:

$$\begin{aligned}\widehat{\Upsilon}_n(x) - \Upsilon(x) &= \int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}\widehat{\Gamma}_n^* - \int_0^x \frac{1}{K} \mathrm{d}\Gamma^* \\ &= \left(\int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}\widehat{\Gamma}_n^* - \int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}\Gamma^* \right) + \left(\int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}\Gamma^* - \int_0^x \frac{1}{K} \mathrm{d}\Gamma^* \right).\end{aligned}$$

By the linearity of the integral, this simplifies to

$$\widehat{\Upsilon}_n(x) - \Upsilon(x) = \int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}(\widehat{\Gamma}_n^* - \Gamma^*) + \int_0^x \left(\frac{1}{\widehat{K}_n} - \frac{1}{K} \right) \mathrm{d}\Gamma^*.$$

Bounding the first term involves integration by parts and the uniform convergence of $\widehat{\Gamma}_n^*$. Given that $\inf_x \widehat{K}_n(x) \ge c/2 > 0$ almost surely for large n, we have

$$\left| \int_0^x \frac{1}{\widehat{K}_n} \mathrm{d}(\widehat{\Gamma}_n^* - \Gamma^*) \right| \le \frac{\sup_x |\widehat{\Gamma}_n^*(x) - \Gamma^*(x)|}{\inf_x \widehat{K}_n(x)} = O(\tau_n) \quad \text{q.s.}$$

For the second term, we utilize the condition $\inf_x \widehat{K}_n(x) \ge c/2 > 0$ and the mean value theorem to bound the difference in reciprocals. The integral is dominated by the supremum of the at-risk process deviation:

$$\left| \int_0^x \frac{K(x) - \widehat{K}_n(x)}{\widehat{K}_n(x) K(x)} \mathrm{d}\Gamma^*(x) \right| \le \frac{\sup_x |\widehat{K}_n(x) - K(x)|}{c^2/2} = O(\tau_n) \quad \text{q.s.}$$

Thus $\sup_{0\le x\le\tau} |\widehat{\Upsilon}_n(x) - \Upsilon(x)| = O(\tau_n)$ q.s. The convergence of the Kaplan–Meier estimator $\widehat{\Gamma}_n(x)$ follows from its product-limit representation:

$$\widehat{\Gamma}_n(x) - \Gamma(x) = \mathrm{e}^{-\Upsilon(x)} (\widehat{\Upsilon}_n(x) - \Upsilon(x)) + R_n(x).$$

The remainder $R_n(x)$ satisfies $|R_n(x)| \le \sum \delta_i I(U_i \le x)/(n\widehat{K}_n(U_i))^2 = O(n^{-1})$ q.s., by the tail bounds for m-WA sums in Lemma 2.4. Since τ_n dominates n^{-1}, we conclude $\sup_{0\le x\le\tau} |\widehat{\Gamma}_n(x) - \Gamma(x)| = O(\tau_n)$ q.s. ■

Theorem 3.2: *Under Theorem* 3.1 *assumptions, we obtain the quasi-sure (q.s.) decompositions:*

$$\widehat{\Upsilon}_n(x) - \Upsilon(x) = -\frac{1}{n} \sum_{i=1}^n \eta(U_i, \delta_i, x) + R_{1n}(x), \tag{14}$$

$$\widehat{\Gamma}_n(x) - \Gamma(x) = -F(x) \left(\frac{1}{n} \sum_{i=1}^n \eta(U_i, \delta_i, x) \right) + R_{2n}(x), \tag{15}$$

where for $j = 1, 2$, *the remainders satisfy* $\sup_{0\le x\le\tau} |R_{jn}(x)| = O(\tau_n)$ *q.s. The influence function* η *takes the form*

$$\eta(u,\delta,x) = \tilde{g}(u\wedge x) - \frac{I(u\le x,\delta=1)}{\overline{K}(u)}$$

with $\tilde{g}(x) = \int_0^x \frac{dF^*(u)}{\overline{K}^2(u)}$ *representing the integrated inverse censoring probability.*

Proof: We begin by examining the difference between the estimated and true cumulative hazard functions, $\widehat{\Upsilon}_n(x) - \Upsilon(x)$. We can express this difference through the following decomposition, a common approach in analysing such estimators:

$$\widehat{\Upsilon}_n(x) - \Upsilon(x) = \left(\int_0^x \frac{\mathrm{d}\Gamma_{*n}(u)}{\overline{K}_n(u)} - \int_0^x \frac{\mathrm{d}\Gamma_*(u)}{\overline{K}(u)}\right).$$

This expression can be further decomposed into three distinct terms for analytical convenience:

$$\left(\int_0^x \frac{\mathrm{d}\Gamma_{*n}(u)}{\overline{K}(u)} - \int_0^x \frac{\overline{K}_n(u)}{\overline{K}^2(u)}\mathrm{d}\Gamma_*(u)\right) + \int_0^x \left(\frac{1}{\overline{K}_n(u)} - \frac{1}{\overline{K}(u)}\right)\mathrm{d}(\Gamma_{*n}(u) - \Gamma_*(u))$$
$$+ \int_0^x \frac{(\overline{K}_n(u) - \overline{K}(u))^2}{\overline{K}^2(u)\overline{K}_n(u)}\mathrm{d}\Gamma_*(u) = A_1(x) + A_2(x) + A_3(x). \tag{16}$$

For the term $A_1(x)$, we note that $F_{*n}(t) = \frac{1}{n}\Sigma_{i=1}^n I(U_i \le t, \delta_i = 1)$. We can rewrite $A_1(x)$ in terms of $\eta(t,\delta,x)$ as

$$A_1(x) = \frac{1}{n}\sum_{i=1}^n \frac{I(U_i\le x,\delta_i=1)}{\overline{K}(U_i)} - \frac{1}{n}\sum_{i=1}^n \tilde{g}(x\wedge U_i) = -\frac{1}{n}\sum_{i=1}^n \eta(U_i,\delta_i,x). \tag{17}$$

Thus $A_1(x)$ represents the leading-order component, which is a sum of centred, m-WA random variables.

Next, we analyse the second term, $A_2(x)$, expressed as

$$A_2(x) = \int_0^x \left(\frac{1}{\widehat{K}_n(u)} - \frac{1}{\overline{K}(u)}\right)\mathrm{d}(\Gamma_{*_n}(u) - \Gamma_*(u)).$$

We bound this integral by employing a standard discretization technique. We divide the interval $[0,\tau]$ into k_n subintervals $[x_j, x_{j+1}]$ such that the variation of $\Upsilon(x)$ over each subinterval is $O(\tau_n)$. This choice implies that $k_n = O(\tau_n^{-1})$. By utilizing the uniform convergence results from Theorem 1 (which relies on Lemma 2.4 for empirical processes) and properties of $\overline{K}_n(u)$ and $\overline{F}_{*n}(t)$, we can bound $A_2(x)$ by

$$A_2(x) \le c \max_{1\le i\le k_n} \sup_{y\in[x_i,x_{i+1}]} \left|\frac{1}{\overline{K}_n(y)} - \frac{1}{\overline{K}_n(x_i)} - \frac{1}{\overline{K}(y)} + \frac{1}{\overline{K}(x_i)}\right|$$
$$+ c \max_{1\le i\le k_n} |\Gamma_{*n}(x_{i+1}) - \Gamma_{*n}(x_i) - \Gamma_*(x_{i+1}) + \Gamma_*(x_i)| + O(\tau_n^2).$$

We denote these parts as A_{21} and A_{22}, respectively:

$$A_{21} + A_{22} + O(\tau_n^2). \tag{18}$$

From Theorem 1, we know that $\sup_x |\overline{K}_n(x) - \overline{K}(x)| = O(\tau_n)$ q.s., which implies that for sufficiently large n, $\overline{K}_n(t)$ is bounded away from zero on $[0, \tau]$ quasi-surely.

To handle A_{21}, we further subdivide each interval $[x_i, x_{i+1}]$ into b_n smaller subintervals $[x_{ij}, x_{i(j+1)}]$, chosen such that the true survival probability $\overline{K}(t)$ varies by at most $O(\tau_n^{3/2})$ within each subinterval. This particular choice leads to $b_n = O(\tau_n^{-1/2})$. Leveraging Theorem 1 (uniform convergence of $\overline{K}_n(y)$ to $\overline{K}(y)$) and the empirical process property that $|\overline{K}_n(y) - \overline{K}_n(x_{ij})| \leq 1/n$ for $y \in [x_{ij}, x_{i(j+1)}]$, we are able to bound A_{21} in terms of sums of centred indicator functions:

$$A_{21} \leq \max_{1\leq i\leq k_n} \max_{1\leq j\leq b_n} \left| \frac{1}{n} \sum_{k=1}^{n} (\eta_{ik} + \zeta_{ijk}) \right| + O(\tau_n^{3/2}), \tag{19}$$

where $\eta_{ik} = I(U_k \geq x_i) - \widehat{\mathbb{E}}[I(U_k \geq x_i)]$ and $\zeta_{ijk} = I(U_k \geq x_{ij}) - \widehat{\mathbb{E}}[I(U_k \geq x_{ij})]$. These are centred indicator functions. Given that these sequences are m-WA, we apply Lemma 2.4 (a Fuk–Nagaev-type inequality) to the sum $\sum_{k=1}^{n}(\eta_{ik} + \zeta_{ijk})$. The detailed application of this provides a bound on the capacity of the sum, which in turn implies the desired convergence rate. For appropriate choices of parameters and τ_n (e.g., related to $\log n/n$), we can show that

$$\mathbb{V}\left(\max_{1\leq i\leq k_n} \max_{1\leq j\leq b_n} \left| \frac{1}{n} \sum_{k=1}^{n} (\eta_{ik} + \zeta_{ijk}) \right| \geq C\tau_n \right) \leq C' n^{-c} \text{for } c > 1.$$

This capacity bound, together with the generalized Borel–Cantelli lemma for capacities, implies the quasi-sure convergence:

$$A_{21} = O(\tau_n), \quad \text{q.s.} \tag{20}$$

Next, considering A_{22}, which involves differences of the empirical sub-distribution functions, similar arguments based on Theorem 1 (uniform convergence of Γ_{*n}) yield

$$A_{22} = O(\tau_n). \tag{21}$$

Combining (18), (20), and (21), we conclude

$$A_2(x) = O(\tau_n), \quad \text{q.s.} \tag{22}$$

Finally, we consider $A_3(x)$:

$$A_3(x) = \int_0^x \frac{(\overline{K}_n(u) - \overline{K}(u))^2}{\overline{K}^2(u)\overline{K}_n(u)} \mathrm{d}\Gamma^*(u).$$

From Theorem 1, we know that $\sup_{0\leq t\leq \tau} |\overline{K}_n(u) - \overline{K}(u)| = O(\tau_n)$ q.s. Since $\overline{K}(u)$ is bounded away from zero on $[0, \tau]$ (by definition of τ_L) and $\overline{K}_n(u)$ also converges uniformly to $\overline{K}(u)$ (and thus is bounded away from zero q.s. for large n), the denominator $\overline{K}^2(u)\overline{K}_n(u)$ is bounded below by a positive constant. Therefore, we can conclude that the term involving a squared difference is a higher-order remainder:

$$A_3(x) = O(\tau_n^2), \quad \text{q.s.} \tag{23}$$

By combining the results for $A_1(x)$ in (17), $A_2(x)$ in (22) and $A_3(x)$ in (23) into the initial decomposition (16), we obtain the expansion for the Nelson–Aalen estimator:

$$\widehat{\Upsilon}_n(x) - \Upsilon(x) = -\frac{1}{n}\sum_{i=1}^{n} \eta(U_i, \delta_i, x) + O(\tau_n) \quad \text{q.s.}$$

This result establishes Equation (13) of Theorem 2, where $R_{1n}(x)$ comprises the higher-order terms $A_2(x) + A_3(x)$, which sum up to $O(\tau_n)$ q.s. (since $O(\tau_n) + O(\tau_n^2) = O(\tau_n)$).

We then extend this result to the Kaplan–Meier estimator. The relationship between the Kaplan–Meier estimator and the Nelson–Aalen estimator can be expressed through a functional Taylor expansion:

$$\widehat{\Gamma}_n(x) - \Gamma(x) = -\Gamma(x)(\widehat{\Upsilon}_n(x) - \Upsilon(x)) + R'_{2n}(x),$$

where $R'_{2n}(x)$ is a higher-order remainder term, typically $O(\tau_n^2)$ or $O(n^{-1})$. Substituting the expansion for $\widehat{\Upsilon}_n(x) - \Upsilon(x)$ from Equation (10) into this relation, we get

$$\widehat{\Gamma}_n(x) - \Gamma(x) = -\Gamma(x)\left(-\frac{1}{n}\sum_{i=1}^{n} \eta(U_i, \delta_i, x) + R_{1n}(x)\right) + R'_{2n}(x) \tag{24}$$

$$= -\Gamma(x)\left(\frac{1}{n}\sum_{i=1}^{n} \eta(U_i, \delta_i, x)\right) - \Gamma(x)R_{1n}(x) + R'_{2n}(x). \tag{25}$$

By grouping the remainder terms, we define $R_{2n}(x) = -\Gamma(x)R_{1n}(x) + R'_{2n}(x)$. Since $R_{1n}(x) = O(\tau_n)$ q.s. (as now established) and $R'_{2n}(x)$ is at least $O(\tau_n^2)$ q.s., we conclude that $R_{2n}(x) = O(\tau_n)$ quasi-surely. ■

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Ikhlasse Chebbab http://orcid.org/0000-0002-6409-7943

References

Aalen, O. O (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics*, *6*(4), 701–726. https://dx.doi.org/10.1214/aos/1176344247

Ahmed, I., & Flandre, P. (2020). Weighted Kaplan–Meier estimators motivating to estimate HIV-1 RNA reduction censored by a limit of detection. *Statistics in Medicine*, *39*(7), 968–983. https://doi.org/10.1002/sim.8455

Andersen, P. K., Borgan, O., Gill, R. D., & Keiding, N (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag.

Anevski, D. (2017). Functional central limit theorems for the Nelson–Aalen and Kaplan–Meier estimators for dependent stationary data. *Statistics & Probability Letters*, *124*(1), 83–91. https://doi.org/10.1016/j.spl.2017.01.005

Fleming, T. R., & Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. Wiley.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452

Klein, J. P., & Moeschberger, M. L (2003). *Survival Analysis: Techniques for Censored and Truncated Data* (Vol. 1230). Springer.
Li, Y., Pang, W., Feng, Z., & Li, N. (2024). Asymptotic properties of Kaplan–Meier estimator and hazard estimator for censored survival time with LENQD data. *Statistical Theory and Related Fields*, *8*(2), 107–116. https://doi.org/10.1080/24754269.2024.2302754
Li, Y. M., & Zhou, Y. (2020). The Kaplan–Meier estimator and hazard estimator for censored END survival time observations. *Communications in Statistics – Theory and Methods*, *49*(11), 2690–2702. https://doi.org/10.1080/03610926.2019.1580737
Li, Y. M., & Zhou, Y. (2024). Asymptotic properties of Kaplan–Meier estimator and hazard estimator for censored survival time with LENQD data. *Statistical Theory and Related Fields*, *8*(2), 107–116. https://doi.org/10.1080/24754269.2024.2302754
Liebscher, E. (2002). Kernel density and hazard rate estimation for censored data under α-mixing condition. *Annals of the Institute of Statistical Mathematics*, *54*(1), 19–28. https://doi.org/10.1023/A:1016157519826
Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology*, *1*(1), 27–52. https://doi.org/10.1080/00224065.1969.11980344
Nematolahi, S., Nazari, S., Shayan, Z., Ayatollahi, S. M. T., & Amanati, A. (2020). Improved Kaplan–Meier estimator in survival analysis based on partially rank-ordered set samples. *Computational and Mathematical Methods in Medicine*, *2020*, 1–11. https://doi.org/10.1155/2020/7827434
Peng, S. G. (2006). G-expectation, G-Brownian motion and related stochastic calculus of Itô's type G-expectation, G-Brownian motion and related stochastic calculus of Itô's type. In H. Holden, K. H. Karlsen, K. A. Lie, & N. H. Risebro (Eds.), *The Abel Symposium 2005, Abel Symposia* (pp. 541–567). Springer-Verlag.
Shen, A. T., & Wang, X. H. (2016). Kaplan–Meier estimator and hazard estimator for censored negatively superadditive dependent data. *Statistics*, *50*(2), 377–388. https://doi.org/10.1080/02331888.2015.1038269
Wu, Q. Y., & Chen, P. Y. (2013). Strong representation results of the Kaplan–Meier estimator for censored negatively associated data. *Journal of Inequalities and Applications*, *2013*(1), 1–9. https://doi.org/10.1186/1029-242X-2013-1
Wu, Y., Deng, X., & Wang, X. J. (2023). Capacity inequalities and strong laws for m-widely acceptable random variables under sub-linear expectations. *Journal of Mathematical Analysis and Applications*, *525*(1), 127–282. https://doi.org/10.1016/j.jmaa.2023.127282