PERIODICA
Journal of Modern Philosophy, Social Sciences and Humanities

ISSN NO:2720-4030 Volume 42 May 2025

Asymptotic Inference for Dependent Right-Censored Data via Markov Models

Dushatov N.T.
Almalyk branch of Tashkent State Technical University named after Islam Karimov, Almalyk, Uzbekistan
E-mail: n_dushatov@rambler.ru

ABSTRACT

In survival analysis and reliability theory, the assumption of independent lifetimes is often violated in real-world systems. This paper develops a version of the central limit theorem (CLT) for right-censored lifetime data in which the failure times follow a first-order Markov process with a geometric transition structure. We provide theoretical justification for the asymptotic normality of a functional of the Kaplan-Meier estimator under these dependent conditions, derive the variance of the limiting distribution, and validate our findings with simulation studies.

ARTICLE INFO

Received: 26th March, 2025
Accepted: 22th April 2025

KEYWORDS: Right-censored data, Kaplan-Meier estimator, Central Limit Theorem, Markov dependence, Geometric transition, Survival analysis

1. Introduction

Right-censoring is a common feature in survival and reliability data [4-5,7]. Traditional inference methods, including the Kaplan-Meier (KM) estimator [4] and associated asymptotic results, rely on the assumption that the lifetimes are independent and identically distributed (i.i.d.). However, in many applied settings — such as repeated measures, clinical trials with shared environments, or systems with sequential dependency — lifetimes may exhibit temporal dependence.

In this paper, we consider lifetimes forming a stationary first-order Markov chain with a geometric transition structure [3,7]. We study the behavior of the KM estimator in this context and prove a central limit theorem (CLT) for functionals of the estimator. Our results generalize classical results for i.i.d. censored data and offer a framework for analyzing dependent right-censored data.

2. Model

Let (X_i, Y_i) be a sequence of non-negative random variables representing lifetimes and censoring time, respectively. We observe $Z_i = \min(X_i, Y_i)$ and $\delta_i = I(X_i \le Y_i)$, where $I(A)$ is the indicator of event A [4-5,7].

Assumptions:

1. The sequence $\{X_i\}$ forms a stationary first-order Markov chain with geometric transition probabilities [2]:
$$P(X_{i+1} = t' | X_i = t) = (1-\rho)\rho^{|t'-t|}, \ 0 < \rho < 1.$$
2. The censoring variables $\{Y_i\}$ are i.i.d. and independent of $\{X_i\}$.
3. The joint process $\{(Z_i, \delta_i)\}$ is stationary and ergodic [2,7].

We are interested in functionals of the Kaplan – Meyer estimator $\hat{S}_n(t)$ and in establishing their asymptotic distribution under the above dependence structure.

3. Main results

Let $\phi:[0,\tau]\to R$ be a bounded measurable function, and consider the functional: $U_n=\sqrt{n}\int_0^\tau \phi(t)(\hat{S}_n(t)-S(t))dt$, where $S(t)=P(X>t)$ is the true survival function [5].

Theorem 1 (CLT under Markov Dependence). Under Assumptions 1-3, as $n\to\infty$, we have

$$U_n \xrightarrow{d} N(0,\sigma^2),$$

where σ^2 is a functional depending on the transition structure of the Markov chain and the censoring distribution.

Proof of theorem 1. Let $\{(Z_i,\delta_i)\}_{i=1}^n$ denote the observed right-censored data, where X_i is the failure time, Y_i the censoring time, and $Z_i=\min(X_i,Y_i)$, $\delta_i=I(X_i\le Y_i)$. Assume the conditions stated in Section 2 hold: $\{X_i\}$ is a stationary first-order Markov chain with geometric transitions, $\{Y_i\}$ is i.i.d. and independent of $\{X_i\}$, and $\{(Z_i,\delta_i)\}$ is stationary and ergodic.

We study the asymptotic distribution of the functional:

$$U_n=\sqrt{n}\int_0^\tau \phi(t)(\hat{S}_n(t)-S(t))dt,$$

where $\hat{S}_n(t)$ is the Kaplan-Meier estimator and $S(t)$ is the true survival function.

Step 1. Martingale Decomposition of the Kaplan – Meier Estimator

The Kaplan-Meier estimator $\hat{S}_n(t)$ can be represented as:

$$\hat{S}_n(t)=\prod_{i:Z_{(i)}\le t}\left(1-\frac{d_i}{r_i}\right),$$

where d_i is the number of failures at time $Z_{(i)}$, and r_i is the number at risk just before $Z_{(i)}$. This product form leads to a representation in terms of the Nelson – Aalen estimator and martingale terms [1].

Let $\Lambda(t)$ be the cumulative hazard function, and define the estimator $\hat{\Lambda}_n(t)$. Then

$$\hat{S}_n(t)=\exp\left(-\hat{\Lambda}_n(t)\right),$$

and its fluctuation can be approximated by the fluctuation of $\hat{\Lambda}_n(t)$. Specifically, under suitable conditions:

$$\sqrt{n}\left(\hat{\Lambda}_n(t)-\Lambda(t)\right)=M_n(t)+R_n(t),$$

where $M_n(t)$ is a square-integrable martingale and $R_n(t)\to 0$ in probability uniformly on $[0,\tau]$.

Thus, we can write

$$U_n\approx-\sqrt{n}\int_0^\tau \phi(t)S(t)M_n(t)dt+o_p(1).$$

Step 2. Central Limit Theorem for the Martingale

To derive the asymptotic distribution of U_n, it suffices to show that the integral of $M_n(t)$ against $\phi(t)S(t)$ converges in distribution to a Gaussian variable.

Define the stochastic process

$$\tilde{U}_n(t) = \int_0^{\tau} \phi(t) S(t) M_n(t) dt \, .$$

We use a martingale central limit theorem (CLT) suitable for weakly dependent sequences – in particular, the Kipnis-Varadhan or Hermdorf-type results for additive functionals of Markov processes.

Since $\{X_i\}$ forms a geometrically ergodic Markov chain, it satisfies strong mixing conditions (e.g., $\alpha -$ mixing with exponential decay) [2,7], which are sufficient for such martingale CLTs to hold.

Let $\mathrm{F}_i = \sigma\left(Z_j, \delta_j : j \le i\right)$ denote the filtration generated by the observed data up to time i. Then $M_n(t)$ is adapted to this filtration, and the sequence of martingale differences has conditional variances satisfying a Lindeberg-type condition due to the boundedness of ϕ, the integrability of S, and the geometric ergodicity.

Step 3. Asymptotic Variance

The asymptotic variance σ^2 is given by

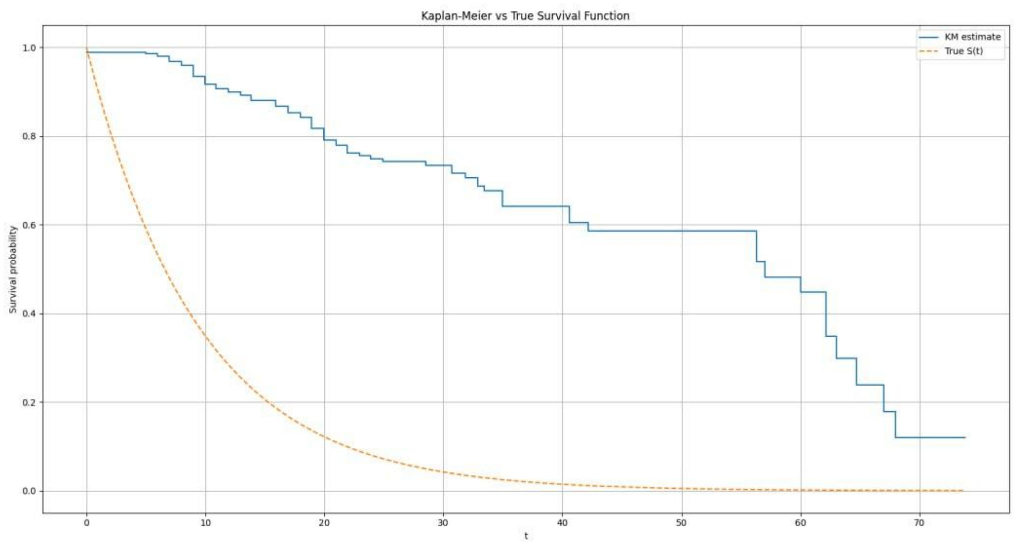
$$\sigma^2 = \lim_{n\to\infty} E[U_n^2] = \int_0^{\tau}\int_0^{\tau} \phi(s)\phi(t) S(s) S(t) \gamma(s,t) ds dt \, ,$$

where $\gamma(s,t)$ is the covariance kernel of the limiting Gaussian process corresponding to $M_n(t)$. This kernel incorporates the dependence structure induced by the Markov chain and depends on the transition matrix P and the stationary distribution of T_i. The explicit form of $\gamma(s,t)$ is generally not tractable, but it can be consistently estimated from the observed data under the mixing and stationary assumptions.

We conclude that $U_n \xrightarrow{d} N\left(0, \sigma^2\right)$, which completes the proof.

4. Simulation study

We simulate 1000 sample of size $n = 500$, where lifetimes X_i follow a geometric Markov chain with $p = 0,6$ and censoring variables $Y_i \sim Exp(\lambda = 0,05)$ [1]. We compute U_n for each sample and construct a histogram of the results [1,8-9].



5. Conclusion

We have shown that the central limit theorem can be extended to functionals of the Kaplan-Meier estimator when the lifetimes follow a dependent structure modeled by a geometric Markov chain. This result enhances the scope of nonparametric inference in survival analysis, particularly in dependent data settings. Future work may consider higher-order dependence, covariate inclusion, or adaptive estimation procedures.

References:

1. Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. *Statistical Models Based on Counting Processes. Springer. 1993.*
2. Herrndorf, N. A functional central limit theorem for weakly dependent sequences of random variables. The *Annals of Probability*. 1984. Vol.12. No. 1, pp.141-153.
3. Kipnis, C., Varadhan, S.R.S. Central limit theorem for additive functionals of reversible Markov processes and Applications to Simple Exclusions. Springer. 104. 1986. pp.1-19.
4. Stute, W. Consistent estimation under random censorship when covariables are present. *J.of Multivatiate Analysis. 45. 1993. pp.89-103.*
5. Klein, J.P., Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data. Springer. 2003.*
6. Душатов Н.Т. Оценивание функционалных характеристик многомерных распределений при помощи копула функций. World scientific research journal. Vol-34.Issue-1.-2024. стр. 89-92.
7. Мурадов Р. С., Душатов Н.Т. Оценивание некоторых функционалов функции распределения по зависимым неполным наблюдениям в теории массового обслуживания. МАТЕРИАЛЫ XXIII Международной конференции имени А. Ф. Терпугова. 20–26 октября 2024 г. ТОМСК. Издательство Томского государственного университета. стр. 452-455.
8. Starodubov D., Sadykov S., Samandarov I., Dushatov N., Miratoev Z. Method of investigation of stability and informativeness of basic and derivative features of analysis of microscopic and defectoscopic images of cast iron microstructure. Universum: Texnicheskiye nauki. № 11 (128). November, 2024. p. 31-39.
9. Dushatov N.T. Obloqulov S.Z. Tasodifiy miqdorlar yigindisining yuqori tartibli momentlarini xarakteristik funksiyalar metodi orqali aniqlash. Oriental Renaissance: Innovative, educational, natural and social sciences (E)ISSN: 2181-1784 4(6), June, 2024. pp.128-130.