STATISTICA, anno LXXXIII, n. 4, 2023

NON-PARAMETRIC KERNEL ESTIMATION OF WEIGHTED DYNAMIC CUMULATIVE PAST INACCURACY MEASURE BASED ON CENSORED DATA

K.V. Viswakala
SQC & OR Unit, Indian Statistical Institute, Bangalore, Karnataka, India
E.I. Abdul Sathar [1]
Department of Statistics, University of Kerala, Thiruvananthapuram, Kerala, India.

SUMMARY

The inaccuracy measure has recently become a valuable tool for detecting errors in experimental data. This measure applies only when random variables have density functions. To circumvent this constraint, the cumulative inaccuracy measure is a commonly used alternative measure of inaccuracy in the literature. When the observations generated by a stochastic process are recorded using a weight function, weighted distributions are established. Based on right-censored dependent data, we provide a nonparametric estimate for the weighted dynamic cumulative past inaccuracy measure in this study. The proposed estimator's asymptotic characteristics have been examined, and its performance demonstrated through simulated and real-world data sets.

Keywords: Alpha-mixing; Information measures; Recursive kernel density estimator; Right-censored data; Weighted dynamic cumulative past inaccuracy measure.

1. INTRODUCTION

Suppose an investigator mistakenly uses β_j, the probability of the occurrence of the j^{th} event, instead of its true probability θ_j, where $1 \leq j \leq n$. Kerridge (1961) proposed an inaccuracy measure that can be computed as $-\sum_{j=1}^{n} \theta_j \log \beta_j$, provided $\sum_{j=1}^{n} \theta_j = \sum_{j=1}^{n} \beta_j = 1$. This measure reduces to the uncertainty measure $-\sum_{j=1}^{n} \theta_j \log \theta_j$, also known as the entropy measure, proposed by Shannon (1948), when the investigator's decision is correct. Shannon's entropy quantifies the average uncertainty associated with an event's occurrence.

Let $g_1(t)$ and $g_2(t)$ denote the probability density functions (pdfs) of the continuous non-negative random variables T_1 and T_2, with distribution functions $G_1(x)$

[1] Corresponding Author. E-mail: sathare@gmail.com

$= P(T_1 \leq x)$ and $G_2(x) = P(T_2 \leq x)$, respectively. Nath (1968) proposed the following measure to quantify the inaccuracy between random variables with pdfs $g_1(t)$ and $g_2(t)$:

$$I(T_1, T_2) = -\int_0^{\infty} g_1(t) \log g_2(t) dt = -E_{g_1}\left[\log g_2(T)\right]. \tag{1}$$

The measure in Eq. (1) helps evaluate the accuracy of experimental results when random variables follow continuous pdfs. It is widely recognized for its usefulness. We explore various applications and advancements in the literature pertaining to this measure, as discussed below.

Nonadditive measures of relative information and precision were proposed by Hooda and Tuteja (1985). Theorems for subjective probability codes for nonadditive measures of inaccuracy were developed by Dial (1987). The quantitative-qualitative measure of inaccuracy was created by Bhatia and Taneja (1993) using reversible symmetry. The measure of inaccuracy between n^{th} record value distributions was studied and explicated by Goel *et al.* (2018). James and Anita (2006) used inaccuracy measures in their demand analysis, paving the way to demonstrate its application in the field of economics. Rajesh *et al.* (2017) and Sathar *et al.* (2019) proposed nonparametric methods for estimating inaccuracy measures under right-censoring, focusing on the residual and past lifetimes of random variables, respectively.

Kumar and Taneja (2015) and Kundu *et al.* (2016) separately developed the measure of past cumulative inaccuracy, which can be used to assess inaccuracy in cases where the probability density functions do not exist, and is defined as follows:

$$\bar{\xi}_{G_1 G_2} = -\int_0^{\infty} G_1(t) \ln G_2(t) dt, \tag{2}$$

where the experimenter's proposed cumulative distribution function ($G_2(t)$) is used in place of the actual cumulative distribution function ($G_1(t)$) due to inaccurate or missing data in the experiment. The cumulative equivalent of the inaccuracy measure described in Eq. (1) is represented by Eq. (2). The quantile variants of the cumulative residual (past) inaccuracy measures, along with their dynamic forms, were introduced by Kayal (2018). In some regularity scenarios, Viswakala and Sathar (2024) investigated nonparametric estimate of the cumulative past inaccuracy measure. The bivariate extension of past cumulative inaccuracy measure and its characteristics were examined by Ghosh and Kundu (2020). Additionally, some findings were derived using the proportional reversed hazard rate models.

Unequal sampling probabilities refer to situations where different members of a population have different chances of being selected for a sample. This scenario occurs in contexts like stratified sampling, where different subgroups are sampled at varying rates, and in cluster sampling, where individuals in larger clusters may have lower probabilities of selection compared to those in smaller clusters. Unequal sampling probabilities are also present when certain subpopulations are over sampled to ensure sufficient representation. In such cases, the assumption of equal probability is violated, necessitating adjustments to maintain the validity of the analysis. Consequently, Equations (1) and (2) are not valid if the random variables posses unequal sampling probabilities. In this context, Kumar *et al.* (2010) proposed a length-biased weighted form of the inaccuracy measure, which is defined as follows:

$$I^W(T_1,T_2) = -\int_0^\infty t\, g_1(t) \log g_2(t) dt,$$

and discussed its properties in residual lifetime scenario.

Viswakala and Sathar (2021) explored and examined the properties of a nonparametric estimator of the weighted residual inaccuracy measure. Within the setting of length biased samples, Rajesh *et al.* (2021) and Richu *et al.* (2022) suggested kernel estimation for the entropy measure and extropy measure, respectively. Daneshi *et al.* (2019) proposed an alternative measure to Eq. (2), namely the weighted (length-biased) cumulative past inaccuracy (WCPI), which is defined as

$$\bar{\xi}^{\mathscr{W}}_{G_1G_2} = -\int_0^\infty t\, G_1(t) \ln G_2(t) dt.$$

This paper also suggested an empirical estimator for the measure and investigated its almost sure convergence property with respect to record values.

In many realistic situations, uncertainty relates to the past. If a random variable, T, is found to be down at time x, then $[x - T | T \leq x]$ describes the time elapsed between the failure of the random variable and the time. Jalayeri and Khorashadizadeh (2017) defined weighted (length-biased) dynamic cumulative past inaccuracy (WDCPI), as

$$\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x) = -\int_0^x t \frac{G_1(t)}{G_1(x)} \ln \frac{G_2(t)}{G_2(x)} dt. \tag{3}$$

Equation(3) can equivalently be expressed as

$$\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x) = \ln G_2(x) \bar{\delta}^{\mathscr{W}}_{T_1}(x) - \frac{1}{G_1(x)} \int_0^x t\, G_1(t) \ln G_2(t) dt, \tag{4}$$

where $\bar{\delta}^{\mathscr{W}}_{T_1}(x) = \dfrac{\int_0^x t \bar{G}_1(t) dt}{\bar{G}_1(x)}$ is the weighted mean past life (WMPL) function. Jalayeri and Khorashadizadeh (2017) explored and identified some properties of the extended weighted cumulative past inaccuracy for doubly truncated random variables but did not provide any method for estimating this measure. Additionally, dependent random variables are encountered more frequently than independent ones in many practical scenarios. The recursive structure of the kernel form enhances the estimator's ability to model complex patterns, making it more effective than non-recursive kernels. These considerations motivated us to develop a non-parametric recursive kernel estimator of WDCPI under dependent conditions.

We discuss an example of a WDCPI measure, and the variation of the measure with respect to x and parameter λ is shown in Figure 1.

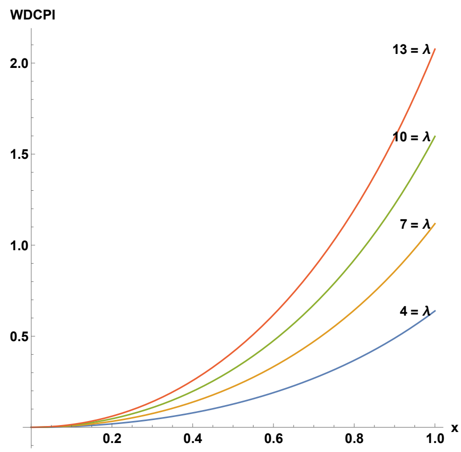


Figure 1 – Plot of $\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)$ against $x \in [0,1]$ for different parameter λ.

EXAMPLE 1. *Let the random variable T_1 has distribution function $G_1(x) = 2x - x^2$, and the random variable T_2 follows power distribution with distribution function $G_2(x) = x^{\lambda}$. Then, for $x \in [0,1]$, the WDCPI measure, $\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)$ is obtained as*

$$\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x) = \frac{\lambda x^2(9x-32)}{144(x-2)}.$$

The plot of the WDCPI measure, $\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)$, for $x \in [0,1]$ and $\lambda \in \{4,\ 7,\ 10,\ 13\}$ is shown in Figure 1. The graph clearly shows that the WDCPI measure exhibits a monotonically increasing trend with respect to both x and λ. This increase is particularly rapid for x values exceeding 0.4.

The outline for this paper is as follows: We discuss the nonparametric recursive kernel estimation of the WDCPI measure based on right-censored samples in Section 2, and Section 3 discusses some of its properties. In Section 4, we carried out a numerical study to assess the performance of the proposed estimator.

2. ESTIMATION OF WDCPI MEASURE

This Section presents a nonparametric estimator for Eq. (4), based on right-censored dependent data, assuming that the underlying lifetimes are α—mixing.

2.1. *Basic concepts*

α—mixing: Rosenblatt (1956) Let $\left(\Omega,\xi,P\right)$ be a probability space and ξ_q^s be the σ-field of events generated by random variables $\{T_r\}_{q\le r\le s}$. Then, the stationary process $\{T_r\}$ is under α—mixing if

$$\alpha(s)=\sup_{\substack{A\in\xi_{-\infty}^{q}\\ B\in\xi_{q+s}^{\infty}}}\Big|P(A\cap B)-P(A)P(B)\Big|\downarrow 0,\text{as } s\to\infty.$$

Recursive kernel density function: Let $\{T_{1r}\}_{1\le r\le n}$ be a sequence of random variables having common density function $g_1(t)$, then the kernel estimator for $g_1(t)$ introduced by Wolverton and Wagner (1969) as

$$g_{1n}(t)=\frac{1}{n}\sum_{r=1}^{n}K_r(t-T_{1r}),\tag{5}$$

where $K_r(u)=\dfrac{1}{h_r}K\left(\frac{u}{h_r}\right)$ and $\{h_n\}$ is a sequence of positive constant, called bandwidth, satisfies $\dfrac{1}{n}\sum\limits_{r=1}^{n}\left(\frac{h_r}{h_n}\right)^i\to\gamma_i<\infty$, as n goes to infinity and $i\in\{1,2,\ldots\}$. The estimator in Eq. (5), which recursively estimates at the point t, can be written as

$$g_{1n}(t)=\frac{n-1}{n}g_{1(n-1)}(t)+\frac{1}{n}K_n(t-T_{1n}),\qquad \text{provided}\quad g_{10}(x)=0.$$

Let the right-censoring random variables $\{X_{1r}\}_{1\le r\le n}$ are i.i.d. with common continuous distribution $R_1(x)$ and are independent of the random variables, $\{T_{1r}\}_{1\le r\le n}$. Let $Y_{1r}=\min\{T_{1r},X_{1r}\}$ and $\Delta_r=I(T_{1r}\le X_{1r})$. An alternative kernel density estimator extending to right-censored sample is given by

$$g_{1n}^*(t)=\frac{1}{n}\sum_{r=1}^{n}\frac{K_r(t-Y_{1r})\Delta_r}{1-R_1(Y_{1r})},\tag{6}$$

and can be recursively estimated at the point t as

$$g_{1n}^*(t)=\frac{n-1}{n}g_{1(n-1)}^*(t)+\frac{1}{n}\frac{K_n(t-Y_{1n})}{1-R_1(Y_{1n})},\qquad \text{provided}\quad g_{10}^*(t)=0.$$

The kernel estimate of distribution function is given in Nadaraya (1965) as

$$G_{1n}(x)=\int_{-\infty}^{x}g_{1n}(t)dt=\frac{1}{n}\sum_{r=1}^{n}\int_{-\infty}^{x}K_r(t-T_{1r})dt,$$

and can be extended to the right-censored case as

$$G_{1n}^{*}(x) = \int_{-\infty}^{x} g_{1n}^{*}(t)dt. \tag{7}$$

If the random variable $\{T_{1r}\}_{1 \leq r \leq n}$ is under α—mixing dependent condition, we get

$$\begin{aligned}
\text{Bias}[G_{1n}^{*}(x)] &= \frac{h_n^2 C_{1+}\gamma_2}{2}\int_0^x g_1^{(2)}(t)dt + O(h_n^2), \\
\text{Var}[G_{1n}^{*}(x)] &\asymp \frac{C_{2+}\eta_1}{nh_n}\int_0^x \frac{g_1(t)}{1-R_1(t)}dt
\end{aligned}$$

and

$$\lim_{n\to\infty} nh_n \text{Cov}(G_{1n}^{*}(x), G_{1n}^{*}(y)) = 0, \quad x \neq y,$$

where $C_{1+} = \int_{\mathbb{R}^*} \alpha^2 K(\alpha)d\alpha$, $C_{2+} = \int_{\mathbb{R}^*} K^2(\alpha)d\alpha$; $\mathbb{R}^* = [0, \infty]$

If the random variables T_1 and T_2 are under α—mixing condition, then the kernel estimator of Eq. (4) under right-censoring is obtained as

$$\bar{\xi}^{\mathscr{W}}{}_{n}^{*}(x) = \ln G_{2n}^{*}(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) - \frac{1}{G_{1n}^{*}(x)}\int_0^x G_{1n}^{*}(t)\ln G_{2n}^{*}(t)dt, \tag{8}$$

where $G_{1n}^{*}(t)$ and $G_{2n}^{*}(t)$ are the nonparametric estimator for $G_1(t)$ and $G_2(t)$ under right-censoring defined in Eq. (7) and

$$\bar{\delta}_{1n}^{\mathscr{W}*}(x) = \frac{1}{G_{1n}^{*}(x)}\int_0^x tG_{1n}^{*}(t)dt,$$

is the nonparametric estimator of $\bar{\delta}_{T_1}^{\mathscr{W}}(x)$ under right-censoring.

3. Properties of the estimators

This Section examines the recursive and asymptotic properties of the nonparametric kernel estimator for the WDCPI measure under right-censoring.

The following theorem outlines the recursive property of the kernel estimator for Eq. (4), demonstrating its recursive calculation.

THEOREM 2. *Given $\bar{\xi}^{\mathscr{W}}*_n(x)$ is the kernel estimator of $\bar{\xi}_{G_1G_2}^{\mathscr{W}}(x)$ under the right-censored α—mixing condition, then the estimator exhibits the recursive property, as expressed in Eq. (9):*

$$
\begin{aligned}
\bar{\xi}^{\mathscr{W}}{}_{n}^{*}(x) &= \ln G_{2n}^{*}(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) + \frac{n-1}{n}\frac{G_{1(n-1)}^{*}(x)}{G_{1n}^{*}(x)} \\
&\times\Big[\bar{\xi}^{\mathscr{W}}{}_{n-1}^{*}(x) - \ln G_{2(n-1)}^{*}(x)\bar{\delta}_{1(n-1)}^{\mathscr{W}*}(x)\Big] \\
&-\frac{n-1}{n}\ln\Big(\frac{n-1}{n}\Big)\frac{1}{G_{1n}^{*}(x)}\int_0^x tG_{1(n-1)}^{*}(t)dt \\
&-\frac{1}{nG_{1n}^{*}(x)}\int_0^x t\bar{I}_{Y_1}(t)\ln\Big(\frac{\bar{I}_{Y_2}(t)}{n}\Big)dt \\
&-\frac{1}{nG_{1n}^{*}(x)}\Big[(n-1)S_1+S_2\Big],
\end{aligned}
\tag{9}
$$

where

$$
\begin{aligned}
S_1 &= \sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\int_0^x\Big[\frac{\bar{I}_{Y_2}(t)}{(n-1)G_{2(n-1)}^{*}(t)}\Big]^i tG_{1(n-1)}^{*}(t)dt, \\
S_2 &= \sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\int_0^x\Big[\frac{(n-1)G_{2(n-1)}^{*}(t)}{\bar{I}_{Y_2}(t)}\Big]^i t\bar{I}_{Y_1}(t)dt, \\
\bar{I}_{Y_j}(t) &= \int_0^t\int_0^u\frac{K_n(v-Y_{in})}{1-R_i(Y_{in})}dvdu,\ j\in\{1,2\}.
\end{aligned}
$$

PROOF. From Eq. (8), we have

$$
G_{1n}^{*}(x)\big[\ln G_{2n}^{*}(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) - \bar{\xi}^{\mathscr{W}}{}_{n}^{*}(x)\big] = \int_0^x tG_{1n}^{*}(t)\ln G_{2n}^{*}(t)dt. \tag{10}
$$

By substituting n with $n-1$ in Eq. (10), we obtain:

$$
\begin{aligned}
&G_{1(n-1)}^{*}(x)\Big[\ln G_{2(n-1)}^{*}(x)\bar{\delta}_{1(n-1)}^{\mathscr{W}*}(x) - \bar{\xi}^{\mathscr{W}}{}^{*}{}_{n-1}(x)\Big] \\
&= \int_0^x tG_{1(n-1)}^{*}(t)\ln G_{2(n-1)}^{*}(t)dt.
\end{aligned}
\tag{11}
$$

For $j\in\{1,\ 2\}$, using the recursive formula of $G_{jn}^{*}(t)$, we get

$$G^*_{1n}(x)\big[\ln G^*_{2n}(x)\bar{\delta}^{\mathscr{W}*}_{1n}(x)-\bar{\xi}^{\mathscr{W}}{}_{\underset{n}{*}}(x)\big]$$

$$\begin{aligned}&= \int_0^x \frac{n-1}{n} t G^*_{1(n-1)}(t)\ln\left\{\frac{n-1}{n}G^*_{2(n-1)}(t)+\frac{\bar{I}_{Y_2}(t)}{n}\right\}dt\\&\quad+\frac{1}{n}\int_0^x t\bar{I}_{Y_1}(t)\ln\left\{\frac{n-1}{n}G^*_{2(n-1)}(t)+\frac{\bar{I}_{Y_2}(t)}{n}\right\}dt,\\&= E_6+F_6,\end{aligned} \tag{12}$$

where

$$E_6=\int_0^x \frac{n-1}{n} t G^*_{1(n-1)}(t)\ln\left\{\frac{n-1}{n}G^*_{2(n-1)}(t)+\frac{\bar{I}_{Y_2(t)}}{n}\right\}dt$$

and

$$F_6=\frac{1}{n}\int_0^x t\bar{I}_{Y_1}(t)\ln\left\{\frac{n-1}{n}G^*_{2(n-1)}(t)+\frac{\bar{I}_{Y_2}(t)}{n}\right\}dt.$$

Applying Eq. (11), we derive:

$$\begin{aligned}E_6 &= \frac{n-1}{n}G^*_{1(n-1)}(x)\Big[\ln G^*_{2(n-1)}(x)\bar{\delta}^{\mathscr{W}*}_{1(n-1)}(x)-\bar{\xi}^{\mathscr{W}}{}_{\underset{n-1}{*}}(x)\Big]\\&\quad+\frac{n-1}{n}\ln\left(\frac{n-1}{n}\right)\int_0^x tG^*_{1(n-1)}(t)dt+\frac{n-1}{n}S_1\end{aligned} \tag{13}$$

and

$$\begin{aligned}F_6 &= \frac{1}{n}\int_0^x t\bar{I}_{Y_1}(t)\ln\left\{\frac{\bar{I}_{Y_2}(t)}{n}\left[1+\frac{(n-1)G^*_{2(n-1)}(t)}{\bar{I}_{Y_2}(t)}\right]\right\}dt,\\&= \frac{1}{n}\int_0^x t\bar{I}_{Y_1}(t)\ln\left\{\frac{\bar{I}_{Y_2}(t)}{n}\right\}dt+\frac{S_2}{n}.\end{aligned} \tag{14}$$

By substituting Equations (13) and (14) into Eq. (12) and rearranging terms, we arrive at Eq. (9). □

The following theorem discusses the consistency property of kernel estimator of Eq. (4) based on right-censored case.

THEOREM 3. *For some positive constants m and N, if*

1. *the second order kernel $K(\cdot)$ satisfies the following conditions:*

   A1: $K(\cdot)$ is symmetric probability function.

   A2: For $\alpha\in\{\beta:K(\beta)\neq 0\}$; $m<K(\alpha)<N$.

A3: $\int_{-\infty}^{\infty} K(t)dt = 1, \quad \int_{-\infty}^{\infty} K^2(t)dt < \infty$ *and* $\int_{-\infty}^{\infty} t^2 K(t)dt < \infty.$

A4: $\int_{-\infty}^{\infty} |K(t)|dt < \infty, \quad \sup |K(t)| < \infty$ *and* $\lim_{|t|\to\infty} |tK(t)| = 0,$

2. *suppose the derivatives of the density functions,* $g_i^{(1)}(t)$, $g_i^{(2)}(t)$ *exist and* $g_i^{(3)}(t)$ *is bounded, for* $i = 1, 2$,

3. *the joint probability density* $g_1(x, y; s)$ *of the random variables* T_{1r} *and* $T_{1(r+s)}$ *exists and satisfies*

$$|g_1(x,y;s) - g_1(x)g_1(y)| \leq M < \infty, \quad \text{for all } x,\ y \text{ and } s \geq 1,$$

where $\{T_{1r}\}$ *is strong mixing, such that* $\sum_{s=1}^{\infty} [\alpha(s)]^a < \infty$, *for* $0 < a < 1/2$ *and a similar assumption is made for the random vector* $\{T_{2r}\}$,

4. *the bandwidth parameters satisfy* $\frac{1}{n}\sum_{r=1}^{n}\left(\frac{h_r}{h_n}\right)^2 \to \gamma_2 < \infty$, $\frac{1}{n}\sum_{r=1}^{n}\frac{h_r}{h_n} \to \eta_1 < \infty$ *and* $h_n \to 0$, nh_n *tends to infinity, when* n *goes to infinity,*

then Eq. (8) is a consistent estimator of Eq. (4).

PROOF. Define

$$\mathcal{V}(x) = \int_0^x tG_1(t)\ln G_2(t)dt$$

and

$$\mathcal{V}_n^*(x) = \int_0^x tG_{1n}^*(t)\ln G_{2n}^*(t)dt,$$

to simplify the notation. With these definitions, we can express

$$\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x) = \ln G_2(x)\bar{\delta}^{\mathcal{W}}_{T_1}(x) - \frac{\mathcal{V}(x)}{G_1(x)}$$

and

$$\bar{\xi}^{\mathcal{W}}_{n}{}^{*}(x) = \ln G_{2n}^*(x)\bar{\delta}^{\mathcal{W}*}_{1n}(x) - \frac{\mathcal{V}_n^*(x)}{G_{1n}^*(x)}.$$

Using Taylor's series expansion, we get

$$\ln G_{2n}^*(t) = \ln G_2(t) + \frac{G_{2n}^*(t) - G_2(t)}{G_2(t)} + \mathbb{E}_\star,$$

where

$$\mathbb{E}_\star = -\int_0^1 \frac{(1-\epsilon)\left[G_{2n}^*(t)-G_2(t)\right]^2}{\left\{G_2(t)+\epsilon[G_{2n}^*(t)-G_2(t)]\right\}^2} d\epsilon.$$

Hence, we get

$$t\ln G_{2n}^*(t)G_{1n}^*(t) = t\ln G_2(t)G_{1n}^*(t) + tG_{1n}^*(t)\frac{G_{2n}^*(t)-G_2(t)}{G_2(t)} + tG_{1n}^*(t)\mathbb{E}_\star,$$

and

$$\ln G_{2n}^*(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) = \ln G_2(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) + \bar{\delta}_{1n}^{\mathscr{W}*}(x)\frac{G_{2n}^*(x)-G_2(x)}{G_2(x)} + \bar{\delta}_{1n}^{\mathscr{W}*}(x)\mathbb{E}_\star.$$

Then, we get

$$\begin{aligned} & tG_{1n}^*(t)\ln G_{2n}^*(t) - tG_1(t)\ln G_2(t) \\ &= [G_{1n}^*(t)-G_1(t)]t\ln G_2(t) + t\frac{[G_{1n}^*(t)-G_1(t)][G_{2n}^*(t)-G_2(t)]}{G_2(t)} \\ &\quad + [G_{2n}^*(t)-G_2(t)]\frac{tG_1(t)}{G_2(t)} + [G_{1n}^*(t)-G_1(t)]t\mathbb{E}_\star + tG_1(t)\mathbb{E}_\star \end{aligned} \tag{15}$$

and

$$\begin{aligned} & \ln G_{2n}^*(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x) - \ln G_2(x)\bar{\delta}_{T_1}^{\mathscr{W}}(x) \\ &= [\bar{\delta}_{1n}^{\mathscr{W}*}(x)-\bar{\delta}_{T_1}^{\mathscr{W}}(x)]\ln G_2(x) + \frac{[\bar{\delta}_{1n}^{\mathscr{W}*}(x)-\bar{\delta}_{T_1}^{\mathscr{W}}(x)][G_{2n}^*(x)-G_2(x)]}{G_2(x)} \\ &\quad + [G_{2n}^*(x)-G_2(x)]\frac{\bar{\delta}_{T_1}^{\mathscr{W}}(x)}{G_2(x)} + [\bar{\delta}_{1n}^{\mathscr{W}*}(x)-\bar{\delta}_{T_1}^{\mathscr{W}}(x)]\mathbb{E}_\star + \bar{\delta}_{T_1}^{\mathscr{W}}(x))\mathbb{E}_\star. \end{aligned} \tag{16}$$

Hence, we get bias, variance and mean squared error (MSE) of $\mathscr{V}_n^*(x)$ as

$$\begin{aligned} \mathrm{Bias}[\mathscr{V}_n^*(x)] &= \frac{h_n^2 C_{1+\gamma_2}}{2}\int_0^x \left[t\ln G_2(t)\int_0^t g_1^{(2)}(u)du + \frac{tG_1(t)}{G_2(t)}\int_0^t g_2^{(2)}(u)du\right]dt \\ &\quad + O\left(\frac{1}{nh_n}\right) + O(h_n^2), \end{aligned} \tag{17}$$

$$\begin{aligned} \mathrm{Var}[\mathscr{V}_n^*(x)] &= \frac{C_{2+}\eta_1}{nh_n}\int_0^x \left[t^2\ln^2 G_2(t)\int_0^t \frac{g_1(u)}{1-R_1(u)}du\right]dt \\ &\quad + \frac{C_{2+}\eta_1}{nh_n}\int_0^x \left[t\frac{G_1(t)}{G_2(t)}\right]^2\int_0^t \frac{g_2(u)}{1-R_2(u)}du\Big]dt + O(h_n^4) \end{aligned} \tag{18}$$

and

$$\mathrm{MSE}[\mathscr{V}_n^*(x)] \;=\; O\left(\frac{1}{nh_n}\right)+O(h_n^4)\longrightarrow 0, \quad \text{as } n \text{ goes to } \infty. \tag{19}$$

Also, we have

$$\bar{\delta}_{1n}^{\mathscr{W}*}(x)-\bar{\delta}_{T_1}^{\mathscr{W}}(x) \;=\; \frac{\int_0^x tG_{1n}^*(t)dt-\bar{\delta}_{1n}^{\mathscr{W}*}(x)G_{1n}^*(x)}{G_1(x)}\Big[1+O_p(1)\Big]. \tag{20}$$

Then, we get

$$\mathrm{Bias}\Big[\bar{\delta}_{1n}^{\mathscr{W}*}(x)\Big] \;=\; \frac{h_n^2C_{1+\gamma_2}}{2G_1(x)}\Big[\int_0^x t\int_0^t g_1^{(2)}(u)du\,dt-\bar{\delta}_{T_1}^{\mathscr{W}}(x)\int_0^x g_1^{(2)}(t)dt\Big] +O(h_n^2) \tag{21}$$

and

$$\mathrm{Var}\Big[\bar{\delta}_{1n}^{\mathscr{W}*}(x)\Big] \;\asymp\; \frac{C_{2+\eta_1}}{nh_n}\frac{1}{G_1^2(x)}\Big[\int_0^x t^2\int_0^t \frac{g_1(u)du}{1-R_1(u)}dt+\big[\bar{\delta}_{T_1}^{\mathscr{W}}(x)\big]^2 \times\int_0^x \frac{g_1(t)}{1-R_1(t)}dt\Big]. \tag{22}$$

Using Eq.(16), we get bias and variance of $\ln G_{2n}^*(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x)$ as

$$\mathrm{Bias}\Big[\ln G_{2n}^*(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x)\Big]$$

$$= \frac{h_n^2C_{1+\gamma_2}}{2}\frac{\ln G_2(x)}{G_1(x)}\left\{\int_0^x t\int_0^t g_1^{(2)}(u)du\,dt-\bar{\delta}_{T_1}^{\mathscr{W}}(x)\int_0^x g_1^{(2)}(t)dt\right\} +\frac{h_n^2C_{1+\gamma_2}}{2}\frac{\bar{\delta}_{T_1}^{\mathscr{W}}(x)}{G_2(x)}\int_0^x g_2^{(2)}(t)dt+\left(\frac{1}{nh_n}\right)+O(h_n^2), \tag{23}$$

$$\mathrm{Var}\Big[\ln G_{2n}^*(x)\bar{\delta}_{1n}^{\mathscr{W}*}(x)\Big]$$

$$\asymp \frac{C_{2+\eta_1}}{nh_n}\left\{\frac{\ln^2 G_2(x)}{\bar{G}_1^2(x)}\Big[\int_0^x t^2\int_0^t \frac{g_1(u)du}{1-R_1(u)}dt+\big[\bar{\delta}_{T_1}^{\mathscr{W}}(x)\big]^2\int_0^x \frac{g_1(t)}{1-R_1(t)}dt\Big] +\Big[\frac{\bar{\delta}_{T_1}^{\mathscr{W}}(x)}{G_2(x)}\Big]^2\int_0^x \frac{g_2(t)}{1-R_2(t)}dt\right\} \tag{24}$$

and hence the MSE of $\ln G^*_{2n}(x)\bar{\delta}^{\mathscr{W}*}_{1n}(x)$ is obtained as

$$\mathrm{MSE}[\ln G^*_{2n}(x)\bar{\delta}^{\mathscr{W}*}_{1n}(x)] = O\left(\frac{1}{nh_n}\right)+O(h_n^4)\longrightarrow 0. \quad (25)$$

We have $\bar{\xi}^{\mathscr{W}}{}_n^*(x) = \ln G^*_{2n}(x)\bar{\delta}^{\mathscr{W}*}_{1n}(x) - \frac{\mathscr{V}^*_n(t)}{G^*_{1n}(x)}$, as n goes to infinity, using Equations (19), (25) and Slutsky's theorem the desired result obtain. □

In the following theorem, we demonstrate that the MSE of the estimator in Eq. (4), based on right-censored case, goes to zero when n goes to infinity.

THEOREM 4. *Under the assumptions given in Theorem 3, the MSE of the estimator in Eq. (8) approaches to zero, when n goes to infinity.*

PROOF. We have,

$$\frac{\mathscr{V}^*_n(x)}{G^*_{1n}(x)} - \frac{\mathscr{V}(x)}{G_1(x)} = \frac{\mathscr{V}^*_n(x) - \frac{\mathscr{V}(x)}{G_1(x)}G^*_{1n}(x)}{G_1(x)}\Big[1+O_p(1)\Big]. \quad (26)$$

Hence, using Equations (17), (18), (23), (24) and (26), we get

$$\begin{aligned}\mathrm{Bias}[\bar{\xi}^{\mathscr{W}}{}_n^*(x)] &\\ = {} & \frac{h_n^2 C_{1+\gamma_2}}{2}\frac{\ln G_2(x)}{G_1(x)}\left\{\int_0^x t\int_0^t g_1^{(2)}(u)du\,dt - \bar{\delta}^{\mathscr{W}}_{T_1}(x)\int_0^x g_1^{(2)}(t)dt\right\}\\ &+\frac{h_n^2 C_{1+\gamma_2}}{2}\left\{\frac{\bar{\delta}^{\mathscr{W}}_{T_1}(x)}{G_2(x)}\int_0^x g_2^{(2)}(t)dt - \frac{1}{G_1(x)}\int_0^x t\ln G_2(t)\int_0^t g_1^{(2)}(u)du\,dt\right\}\\ &+\frac{h_n^2 C_{1+\gamma_2}}{2G_1(x)}\left\{\int_0^x \frac{tG_1(t)}{G_2(t)}\int_0^t g_2^{(2)}(u)du\,dt + \frac{\mathscr{V}(x)}{G_1(x)}\int_0^x g_1^{(2)}(t)dt\right\}\\ &+\left(\frac{1}{nh_n}\right)+O(h_n^2), \end{aligned} \quad (27)$$

$$
\begin{aligned}
&\operatorname{Var}[\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)] \\
&= \frac{C_{2+\eta_1}}{nh_n}\left\{\frac{\ln^2 G_2(x)}{G_1^2(x)}\Big[\int_0^x t^2\int_0^t \frac{g_1(u)du}{1-R_1(u)}dt+[\bar{\delta}^{\mathscr{W}}_{T_1}(x)]^2\int_0^x \frac{g_1(t)dt}{1-R_1(t)}\Big]\right. \\
&\left.+\Big[\frac{\bar{\delta}^{\mathscr{W}}_{T_1}(x)}{G_2(x)}\Big]^2\int_0^x \frac{g_2(t)dt}{1-R_2(t)}\right\}+\frac{C_{2+\eta_1}}{nh_nG_1^2(x)}\int_0^x t^2\ln^2 G_2(t)\int_0^t\frac{g_1(u)du}{1-R_1(u)}dt \\
&+\frac{C_{2+\eta_1}}{nh_nG_1^2(x)}\left\{\int_0^x\Big[\frac{tG_1(t)}{G_2(t)}\Big]^2\int_0^t\frac{g_2(u)du}{1-R_2(u)}dt+\Big[\frac{\mathscr{V}(x)}{G_1(x)}\Big]^2\int_0^x\frac{g_1(t)dt}{1-R_1(t)}\right\} \\
&+O(h_n^4). \qquad (28)
\end{aligned}
$$

When n goes to infinity, $nh_n \to \infty$ and $h_n \to 0$ in Equations (27) and (28), complete the proof. □

In the following theorem, we prove that the mean integrated squared error of the estimator in Eq. (4), based on right-censored case, goes to zero when n goes to infinity.

THEOREM 5. *Under the assumption given in Theorem 3, the mean integrated squared error of the estimator in Eq. (8) tends to zero, when n goes to infinity.*

PROOF. We have,

$$
\begin{aligned}
\operatorname{MISE}\Big[\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)\Big] &= E\int[\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)-\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)]^2dx, \\
&\leq \int E[\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)-\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)]^2dx=\int \operatorname{MSE}[\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)]dx. \qquad (29)
\end{aligned}
$$

Using Theorem 4 in Eq. (29), when n goes to infinity, the proof follows. □

In the following theorem, we provided the almost sure convergence property of the estimator in Eq. (4), based on the right-censored case.

THEOREM 6. *Let $\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)$ be the nonparametric estimator in Eq. (4) based on right-censored sample. Suppose $G_1(x)$ and $G_2(x)$ satisfy the Lipschitz condition and for some fixed point λ, $0<\lambda<\infty$, the distribution function of censored variables, $L_i(\cdot)$, satisfies $L_i(\lambda)<1$; i=1,2. Moreover, if the second order kernel $K(x)$ satisfies assumptions stated in Theorem 3, then*

$$
\sup_{0\leq x\leq\lambda}|\bar{\xi}^{\mathscr{W}}_{n}{}_{*}(x)-\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)| \;\to\; 0 \quad a.s.
$$

PROOF. We have,

$$\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)-\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x) = \left[\ln G_{2n}^{*}(x)\bar{\delta}_{1n}^{\mathcal{W}*}(x)-\ln G_2(x)\bar{\delta}_{T_1}^{\mathcal{W}}(x)\right]-\left[\frac{\mathcal{V}_n^{*}(x)}{G_{1n}^{*}(x)}-\frac{\mathcal{V}(x)}{G_1(x)}\right].$$

Using Equations (15) and (16), we get

$$\left|\mathcal{V}_n^{*}(x)-\mathcal{V}(x)\right| \approx \int_0^x |t\ln G_2(t)||G_{1n}^{*}(t)-G_1(t)|\,dt+\int_0^x \frac{tG_1(t)}{G_2(t)}|G_{2n}^{*}(t)-G_2(t)|\,dt. \tag{30}$$

Also, we have

$$\left|\bar{\delta}_{1n}^{\mathcal{W}*}(x)-\bar{\delta}_{T_1}^{\mathcal{W}}(x)\right| = \frac{\int_0^x t\left|G_{1n}^{*}(t)-G_1(t)\right|dt}{G_{1n}^{*}(x)}+\bar{\delta}_{T_1}^{\mathcal{W}}(x)\frac{\left|G_{1n}^{*}(x)-G_1(x)\right|}{G_{1n}^{*}(x)}. \tag{31}$$

Moreover, we have

$$\left|\frac{\mathcal{V}_n^{*}(x)}{G_{1n}^{*}(x)}-\frac{\mathcal{V}(x)}{G_1(x)}\right| = \frac{|\mathcal{V}_n^{*}(x)-\mathcal{V}(x)|}{G_{1n}^{*}(x)}+\frac{\mathcal{V}(x)}{G_1(x)}\frac{\left|G_{1n}^{*}(x)-G_1(x)\right|}{G_{1n}^{*}(x)}. \tag{32}$$

Also, we can express

$$\left|\ln G_{2n}^{*}(x)\bar{\delta}_{1n}^{\mathcal{W}*}(x)-\ln G_2(x)\bar{\delta}_{T_1}^{\mathcal{W}}(x)\right| \approx \left|\bar{\delta}_{1n}^{\mathcal{W}*}(x)-\bar{\delta}_{T_1}^{\mathcal{W}}(x)\right||\ln G_2(x)|+|G_{2n}^{*}(x)-G_2(x)|\left|\frac{\bar{\delta}_{T_1}^{\mathcal{W}}(x)}{G_2(x)}\right|. \tag{33}$$

Using the almost sure convergence property of the kernel estimator suggested by Cai and Roussas (1992) and Cai (1998), we get

$$\sup_{0\le x\le\lambda}\left|G_{in}^{*}(x)-G_i(x)\right| \;\to\; 0 \quad a.s.; \quad i=1,2.$$

Using this, the proof immediately follows from Equations (30)–(33). □

The following theorem discusses the asymptotic normality of kernel estimator in Eq. (4) based on right-censored case.

THEOREM 7. *Assume that the assumptions stated in Theorem 3 hold true. If x is a point of continuity of $G_1(x)$ and $G_2(x)$ with $G_1(x)>0$ and $G_2(x)>0$ and for some fixed point λ, $0<\lambda<\infty$ such that if $x\in[0,\lambda]$,*

$$\sqrt{nh_n}\left(\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)-\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x)\right)\to\mathcal{N}\left(0,\sigma^2_{\bar{\xi}^{\mathcal{W}}{}_{n}^{*}}\right),$$

with $\sigma^2_{\bar{\xi}^{\mathcal{W}}{}_{n}^{*}}=nh_n Var[\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)]$.

PROOF. We have,

$$\sqrt{nh_n}\Big(\bar{\xi}^{\mathscr{W}}{}_{n}^{*}(x)-\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)\Big)$$

$$= \sqrt{nh_n}\left\{\left[\ln G_{2n}^{*}(x)\bar{\delta}^{\mathscr{W}*}_{1n}(x)-\ln G_2(x)\bar{\delta}^{\mathscr{W}}_{T_1}(x)\right]-\left[\frac{\mathscr{V}_n^{*}(x)}{G_{1n}^{*}(x)}-\frac{\mathscr{V}(x)}{G_1(x)}\right]\right\}.$$

Using Equations (16) and (30)-(33), we get

$$\sqrt{nh_n}\Big(\bar{\xi}^{\mathscr{W}}{}_{n}^{*}(x)-\bar{\xi}^{\mathscr{W}}_{G_1G_2}(x)\Big)$$

$$\begin{aligned} =\ & \sqrt{nh_n}\frac{\ln G_2(x)}{G_1(x)}\left\{\int_0^x t\int_0^t\Big(g_{1n}^{*}(u)-g_1(u)\Big)du\,dt\right. \\ & \left.+\bar{\delta}^{\mathscr{W}}_{T_1}(x)\int_0^x\Big(g_{1n}^{*}(t)-g_1(t)\Big)dt\right\}+\sqrt{nh_n}\frac{\bar{\delta}^{\mathscr{W}}_{T_1}(x)}{G_2(x)}\int_0^x\Big(g_{2n}^{*}(t)-g_2(t)\Big)dt \\ & -\frac{\sqrt{nh_n}}{G_1(x)}\left\{\int_0^x t\ln G_2(t)\int_0^t\Big(g_{1n}^{*}(u)-g_1(u)\Big)du\,dt\right. \\ & \left.+\int_0^x\frac{tG_1(t)}{G_2(t)}\int_0^t\Big(g_{2n}^{*}(u)-g_2(u)\Big)du\,dt-\frac{\mathscr{V}(x)}{G_1(x)}\int_0^x\Big(g_{1n}^{*}(t)-g_1(t)\Big)dt\right\}. \end{aligned} \tag{34}$$

Using the asymptotic normality property of the kernel density estimator from Masry (1986) and Cai (1998), we get

$$\sqrt{nh_n}\big(g_{in}^{*}(t)-g_i(t)\big)\to\mathscr{N}\big(0,\sigma_{i+}^2\big),\quad \sigma_{i+}^2=C_{2+}\eta_1\frac{g_i(t)dt}{1-R_i(t)};\ i=1,2.$$

Thus, the proof is established by Eq. (34). □

In the following Section, we present a Monte-Carlo simulation study to evaluate the estimator's performance, recommend an empirical competitor for the proposed estimator, and provide comparisons based on bias and MSE. We also examine a real-world data application.

4. SIMULATION STUDY

To generate dependent data, we used the AR(1) model to produce random variables. We simulated 350 simulated samples of the form $T_{ir}=\sqrt{(1-\rho^2)}|X_{ir}|$, where i =1 and 2,

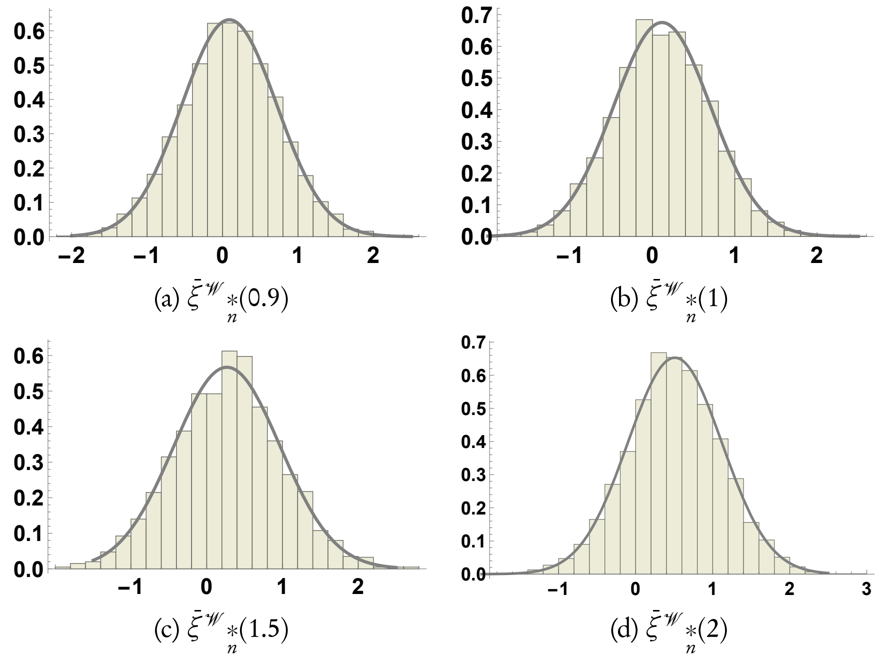


Figure 2 – Histogram of $\bar{\xi}^{\mathscr{W}}_{n}{}^{*}(x)$ with normal density curve from sample of size 200.

$r = \{1,2,\ldots,350\}$. The variables $\{X_{ir}\}$ were generated using an AR(1) with parameter $\rho = 0.5$ and white noise parameters (0, 1) and (0, 2), respectively. The resulting T_{ir} values are stationary and α—mixing, following half normal densities with parameters 1 and 2 respectively. Exponential distributions with parameters 1 and 2 were used for censoring observations. We used the Epanechnikov kernel function and selected the optimal bandwidth for the recursive kernel estimator, as discussed in Yousri (2014).

We repeated the method 200 times to verify that the proposed estimator exhibits acceptable asymptotic normality. The resulting histograms of the estimator in Eq. (8) are shown in Figure 2. We tested AIC and BIC in every situation and passed each time.

The simulation study and the data in Table 1 validate the estimator's performance in terms of MISE and confidence interval (CI) estimates.

In the next sub-Section, we propose an empirical estimator to numerically compare the behaviour of the estimators based on their bias and MSE.

4.0.1. Comparative study

Let $\{T_{ir}\}_{1\leq r\leq n}$, for i =1 and 2, be a sequence of n identical lifetime variables and be censored on the right by independently and identically distributed random variables

TABLE 1

Comparison of MISE of $\bar{\xi}^{\mathcal{W}*}_{n}(x)$ and 95% confidence interval of $\bar{\xi}^{\mathcal{W}*}_{n}(1.7)$.

n	50		100	
ρ	MISE	95% CI	MISE	95% CI
-0.9	0.089	(0.284, 0.381)	0.036	(0.309, 0.378)
-0.6	0.010	(0.299, 0.378)	0.007	(0.330, 0.371)
-0.3	0.084	(0.306, 0.461)	0.018	(0.311, 0.431)
0	0.028	(0.303, 0.395)	0.021	(0.325, 0.386)
0.3	0.059	(0.293, 0.406)	0.011	(0.300, 0.385)
0.6	0.059	(0.319, 0.402)	0.018	(0.322, 0.381)
0.9	0.113	(0.281, 0.411)	0.014	(0.298, 0.378)

n	200		300	
ρ	MISE	95% CI	MISE	95% CI
-0.9	0.011	(0.317, 0.370)	0.007	(0.324, 0.369)
-0.6	0.006	(0.333, 0.361)	0.003	(0.345, 0.359)
-0.3	0.006	(0.317, 0.410)	0.001	(0.323, 0.408)
0	0.005	(0.337, 0.376)	0.003	(0.349, 0.374)
0.3	0.008	(0.325, 0.392)	0.005	(0.331, 0.386)
0.6	0.006	(0.329, 0.367)	0.005	(0.335, 0.362)
0.9	0.008	(0.307, 0.364)	0.005	(0.314, 0.361)

$\{X_{ir}\}_{1\le r\le n}$, so that X_{ir} and T_{ir} are independent. In this censoring scheme one can observe (Y_{ir},Δ_{ir}), where $Y_{ir} = \min(T_{ir},X_{ir})$ and $\Delta_{ir} = I(T_{ir} \le X_{ir})$. Denote $\{Y_{ir:n}\}_{1\le r\le n}$, the sample order statistics and let

$$M_j = \sum_{r=1}^{n} I(Y_{1r} \le Y_{2j:n}), \quad r = 1,2,\dots,n, \tag{35}$$

the number of random variables of the first censored sample that are less than or equal to j^{th} order statistics of the second censored sample. Moreover, we rename by $Y_{1(j,1)} < Y_{1(j,2)} < \dots$ the random sample of the first censored sample belonging to $(Y_{2j:n}, Y_{2(j+1):n}]$, if any. Then, in the context of right-censoring, we get the empirical estimator of WDCPI measure as

$$\begin{aligned}\xi\bar{\mathcal{W}}_n^{\text{cen}}(x) &= -\frac{1}{2n}\sum_{j=1}^{n-1}\left[\frac{M_{j+1}Y^2_{2(j+1):n} - M_jY^2_{2j:n} - \sum_{k=1}^{M_{j+1}-M_j} Y^2_{1(j,k)}}{G_{1+}(x)\sum_{r=1}^{n} I(X_{1r} > Y_{2j:n})}\right] \\ &\quad\times \ln\left(\frac{j}{G_{2+}(x)\sum_{r=1}^{n} I(X_{2r} > Y_{2j:n}}\right) I(Y_{1j:n} \le x),\end{aligned} \tag{36}$$

where,

$$G_{i+}(x) = 1 - \prod_{1\le r\le n}\left(1 - \frac{\Delta_{ir:n}}{n-r+1}\right)^{I(Y_{ir:n}\le x)} \qquad i \in \{1,2\}.$$

On the basis of bias and MSE, we compared the performance of the proposed estimator in Eq. (8) with its corresponding competitor in Eq. (36). The bias and MSE of the estimators are shown in Tables 2 and 3, respectively.

In the following sub-Section, we conduct a real data analysis to study the use of estimators in real situations.

4.0.2. Real data analysis

In this sub-Section, we took a numerical example based on real life data set to illustrate the performance of the estimators in Equations (8) and (36).

EXAMPLE 8. *Consider the data from King* et al. *(1979), which discussed the tumour free time (in days) of 30 rats in the unsaturated fat diet group, in order to evaluate the performance of the estimators for estimating the WDCPI measure.*

TABLE 2
Comparison of bias of the estimators $\bar{\xi}^{W}_{n^*}(x)$ and $\bar{\xi}^{W}_{n}{}^{cen}(x)$

n	100		200		300	
x	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$
1.1	0.113	0.298	0.047	0.206	0.029	0.164
1.3	0.245	0.424	0.150	0.231	0.069	0.171
1.5	0.338	0.456	0.229	0.277	0.193	0.229
1.7	0.355	0.596	0.259	0.425	0.251	0.339
1.9	0.406	0.812	0.283	0.594	0.276	0.508

TABLE 3
Comparison of MSE of the estimators $\bar{\xi}^{W}_{n^*}(x)$ and $\bar{\xi}^{W}_{n}{}^{cen}(x)$

n	100		200		300	
x	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$	$\bar{\xi}^{W}_{n^*}(x)$	$\bar{\xi}^{W}_{n}{}^{cen}(x)$
1.1	0.013	0.129	0.004	0.052	0.001	0.029
1.3	0.060	0.185	0.027	0.054	0.005	0.029
1.5	0.115	0.213	0.053	0.084	0.037	0.054
1.7	0.126	0.355	0.067	0.193	0.064	0.115
1.9	0.165	0.659	0.080	0.371	0.077	0.259

TABLE 4
Fitting details of real data

	Distribution	Parameters	AIC	BIC
Example 8:	Uniform	(59.97, 178.39)	-9.69	-9.65
	Log-normal	(4.54, 0.34)	**-9.79**	**-9.75**
	Gamma	(8.472, 11.769)	**-9.85**	**-9.80**
	Extreme value	(83.42, 26.18)	-9.79	-9.75

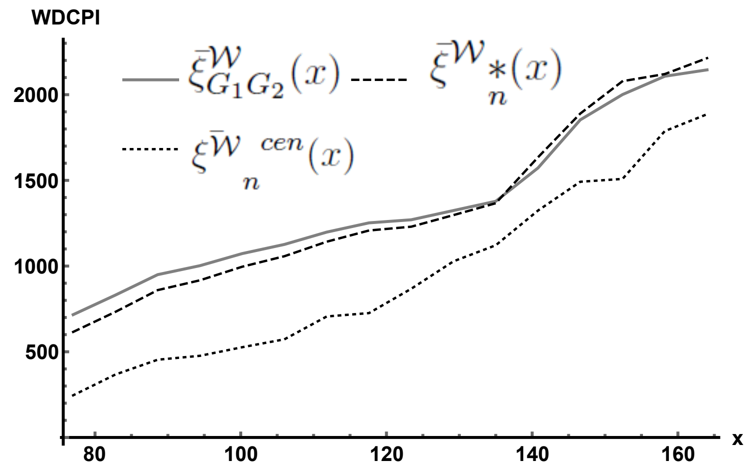


Figure 3 – Plot for the WDCPI measure and estimators of the tumour free time (in days) of 30 rats in the unsaturated fat diet group.

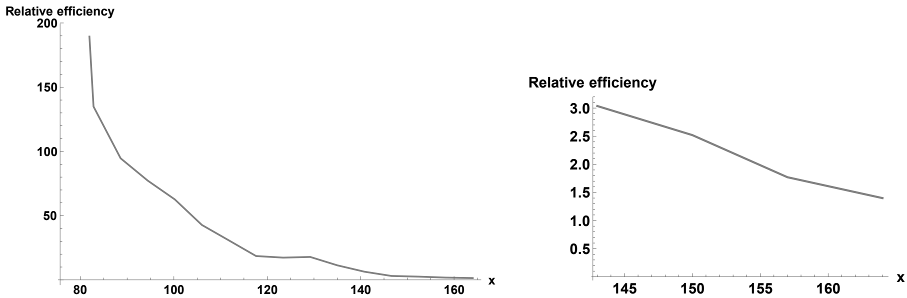
We fitted the real example to four distributions based on AIC and BIC, as shown in Table 4. The bold figures in the table represent the better fit of two distributions. we use the bootstrapping method to analyse the performance of estimators in real-world situations. In Example 8, we use the gamma distribution as the original, the log-normal distribution as the suggested, and randomly selected censoring random variables as exponential with parameters of 0.0009 and 0.0003, respectively. Figure 3 shows the theoretical and estimators values of the WDCPI measure and Figure 4 depicts the relative efficiency of the kernel estimator compared to its corresponding empirical estimator.

4.0.3. Conclusions

In this sub-Section, the conclusion of the numerical study is presented. We examined the asymptotic properties of the kernel estimator and numerically compared it to the empirical estimator, leading to the following conclusions.

The asymptotic normality of the kernel estimator is validated by Figure 2. Table 1 shows that as sample size increases, MISE and the size of the confidence interval both decrease, indicating improved estimator accuracy. Tables 2 and 3 show that as sample size increases, bias and MSE decrease. The table shows that the proposed estimator outperforms its competitor in terms of bias and MSE. Additionally, the x values have a direct influence on biases and MSEs. The simulation results demonstrate the effectiveness of the proposed estimator under right-censoring schemes, which is essential for real-world applications involving incomplete data.

The observations obtained from Figure 3 and Figure 4 are as follows: In Figure 3,



(c)

(d) Enlargement of (c) for $143 \leq x \leq 164$.

Figure 4 – Relative efficiency of the kernel estimator with respect to the empirical estimator for the WDCPI measure of the tumour free time (in days) of 30 rats in the unsaturated fat diet group.

as x increases, the theoretical value $\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x)$ and the two estimators generally show an increasing trend, indicating that the WDCPIM increases over x when the log-normal distribution is chosen for fitting the data set instead of the gamma distribution under a right censoring scheme. The kernel estimator $\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)$ tends to closely follow the general shape of the theoretical value $\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x)$, although with some deviations. This indicates that $\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)$ slightly diverges in its estimation compared to the empirical estimator $\bar{\xi}^{W}{}_{n}^{\text{cen}}(x)$. However, in this example, $\bar{\xi}^{\mathcal{W}}{}_{n}^{*}(x)$ is observed to be the better estimator of $\bar{\xi}^{\mathcal{W}}_{G_1G_2}(x)$ due to its closer alignment with the theoretical value.

Additionally, the relative efficiency of the kernel estimator with respect to the empirical estimator, as depicted in Figure 4, shows that relative efficiency decreases as x increases. However, the relative efficiency remains greater than 1, indicating that the kernel estimator consistently outperforms the empirical estimator across the range of x. This highlights the superiority of the kernel estimator in this context.

Acknowledgements

We sincerely thank the Editor and anonymous reviewer(s) for their valuable comments and constructive feedback, which have significantly enhanced the quality of this work.

References

P. K. Bhatia, H. C. Taneja (1993). *On quantitative-qualitative measure of inaccuracy reversible symmetry*. Information Sciences, 67, pp. 277–282.

Z. CAI (1998). *Kernel density and hazard rate estimation for censored dependent data*. Journal of Multivariate Analysis, 67, no. 1, pp. 23–34.

Z. CAI, G. G. ROUSSAS (1992). *Uniform strong estimation under α-mixing, with rates*. Statistics and Probability Letters, 15, no. 1, pp. 47–55.

S. DANESHI, A. NEZAKATI, S. TAHMASEBI (2019). *On weighted cumulative past (residual) inaccuracy for record values*. Journal of Inequalities and Applications, 2019, no. 134.

G. DIAL (1987). *On non-additive measures of inaccuracy and a coding theorem*. Journal of Information and Optimization Sciences, 8, no. 1, pp. 113–118.

A. GHOSH, C. KUNDU (2020). *On generalized conditional cumulative residual inaccuracy measure*. Communications in Statistics - Theory and Methods, 49, no. 6, pp. 1402–1421.

R. GOEL, H. TANEJA, V. KUMAR (2018). *Kerridge measure of inaccuracy for record statistics*. Journal of Information and Optimization Sciences, 39, no. 5, pp. 1149–1161.

D. S. HOODA, R. K. TUTEJA (1985). *On characterization of non-additive measures of relative information and inaccuracy*. Bulletin Calcutta Mathematical Society, 77, pp. 363–369.

S. JALAYERI, M. KHORASHADIZADEH (2017). *An extension of weighted cumulative inaccuracy measure based on doubly truncation*. Pakistan Journal of Statistics and Operation Research, 13, no. 1, pp. 103–113.

L. S. J. JAMES, R. ANITA (2006). *Modeling international consumption patterns*. Review of Income and Wealth, 52, pp. 603–624.

S. KAYAL (2018). *Quantile-based cumulative inaccuracy measures*. Physica A: Statistical Mechanics and its Applications, 510, pp. 329–344.

D. F. KERRIDGE (1961). *Inaccuracy and inference*. Journal of the Royal Statistical Society, Series B, 23, pp. 184–194.

M. M. KING, D. M. BAILEY, D. G. GIBSON, J. V. PITHA, P. B. MCCAY (1979). *Incidence and growth of mammary tumors induced by 7, 12-dimethylbenz[a]anthracene as related to the dietary content of fat and antioxidant*. Journal of the National Cancer Institute, 63, no. 3, pp. 657–663.

V. KUMAR, R. SRIVASTAVA, H. C. TANEJA (2010). *Length biased weighted residual inaccuracy measure*. METRON, 68, p. 153–160.

V. KUMAR, H. C. TANEJA (2015). *Dynamic cumulative residual and past inaccuracy measures*. Journal of Statistical Theory and Applications, 14, no. 4, pp. 399–412.

C. Kundu, A. D. Crescenzo, M. Longobardi (2016). *On cumulative residual (past) inaccuracy for truncated random variables*. Metrika, 79, no. 3, pp. 335–356.

E. Masry (1986). *Recursive probability density estimation for weakly dependent stationary processes*. IEEE Transactions on Information Theory, 32, no. 2, pp. 254–267.

É. A. Nadaraya (1965). *On non-parametric estimates of density functions and regression curves*. Theory of Probability and Its Applications, 10, no. 1, pp. 186–190.

P. Nath (1968). *Inaccuracy and coding theory*. Metrika, 13, pp. 123–135.

G. Rajesh, R. Richu, S. M. Sunoj (2021). *Kernel estimation of entropy function for length-biased censored data*. Journal of the Indian Society for Probability and Statistics, 22, pp. 251–263.

G. Rajesh, E. I. A. Sathar, K. V. Viswakala (2017). *Estimation of inaccuracy measure for censored dependent data*. Communications in Statistics - Theory and Methods, 46, no. 20, p. 10058–10070.

R. Richu, G. Rajesh, S. M. Sunoj (2022). *Kernel estimation of extropy function under length-biased sampling*. Statistics and Probability Letters, 181, no. 109290.

M. Rosenblatt (1956). *A central limit theorem and a strong mixing condition*. Proceedings of the National Academy of Sciences of the United States of America, 42, no. 1, pp. 43–47.

E. I. A. Sathar, K. V. Viswakala, G. Rajesh (2019). *Estimation of past inaccuracy measure for the right censored dependent data*. Communications in Statistics - Theory and Methods, 50, no. 6, p. 1446–1455.

C. E. Shannon (1948). *A mathematical theory of communication*. The Bell System Technical Journal, 27, pp. 379–423.

K. V. Viswakala, E. I. A. Sathar (2021). *Estimation of weighted residual inaccuracy measure for right censored dependent data*. Journal of the Indian Society for Probability and Statistics, 22, pp. 343–356.

K. V. Viswakala, E. I. A. Sathar (2024). *Kernel estimation of the dynamic cumulative past inaccuracy measure for right censored dependent data*. REVSTAT-Statistical Journal, 22, no. 4, p. 527–544.

C. Wolverton, T. J. Wagner (1969). *Asymptotically optimal discriminant functions for pattern classification*. IEEE Transactions on Information Theory, 15, pp. 258–265.

S. Yousri (2014). *Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method*. Journal of Probability and Statistics, 2014, no. 739640.