Statistica Sinica Preprint No: SS-2025-0207	
Title	A Model-free Correlation Coefficient for Censored Data
Manuscript ID	SS-2025-0207
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202025.0207
Complete List of Authors	Linlin Dai, Tengfei Li and Kani Chen
Corresponding Authors	Kani Chen
E-mails	makchen@ust.hk
Notice: Accepted author version.	

A Model-Free Correlation Coefficient for Censored Data

Linlin Dai[1], Tengfei Li[2] and Kani Chen[3] *

[1] *Southwestern University of Finance and Economics*

[2] *University of North Carolina at Chapel Hill*

[3] *Hong Kong University of Science and Technology*

Abstract: In clinical studies, assessing statistical associations between covariates and survival outcomes is crucial. To date, there has been no formally defined model-free correlation coefficient for right-censored data that can measure the strength of associations. Traditional methods, such as the Cox proportional hazards model, often struggle with the complexities of non-monotonic or nonlinear relationships. This paper introduces a censored rank-based correlation coefficient (CRC). It consistently estimates a new dependence measure—taking values in $[0, 1]$ and equaling 0 or 1 if and only if the variables are independent or one is a measurable function of the other. The CRC is entirely model-free without depending on the distributions of the variables. It facilitates quick computation with a complexity of $O(n \log n)$ and can effectively detect nonlinear and non-monotonic effects, even under heavy censoring. The p-values for testing independence can be obtained using a power-consistent permutation method. The CRC shows strong consistency and asymptotic normality, outperforming the Cox model and other methods in detecting nonlinear associations in both simulations and real data from the Alzheimer's Disease Neuroimaging Initiative, successfully identifying proteins

*Corresponding author. Email: makchen@ust.hk.

that existing methods fail to detect.

Key words and phrases: Chatterjee's rank correlation; Nonparametric correlation; Right-censoring; Permutation test.

1. Introduction

In clinical studies and biomedical sciences, the evaluation and identification of statistical associations between covariates and survival outcomes is a critical endeavor. For instance, for Alzheimer's disease (AD), the most common form of dementia in older adults, it is crucial to identify factors that affect survival time, such as age, sex, and cognitive function at diagnosis. A challenge is that survival times are unobservable for participants who remain AD-free at follow-up. While the Cox proportional hazards model (CPH) (Cox, 1972) is often used for such analyses (Gambassi et al., 1999), the hazard function may, in practice, exhibit non-monotonic or other complex relationships with covariates, as illustrated in Bhaskaran et al. (2018); Zajacova and Burgard (2012) and Figure 2 in Section 9. Such complexities can violate the linearity assumption of the CPH model.

The aim of this paper is to develop a fully model-free correlation coefficient for right-censored data that can also achieve fast computation. Correlation coefficients are powerful tools to measure the strength of the relationship between the variables. For completely observed data, the most well-known classical coeffi-

cients include Pearson's correlation coefficient, Spearman's ρ and Kendall's τ. While these traditional measures are inadequate for detecting non-monotonic associations (Chatterjee, 2021), this limitation motivates the development of new coefficients, including the maximal correlation coefficient (Breiman and Friedman, 1985), maximal information coefficient (Reshef et al., 2011), distance correlation (Székely et al., 2007; Székely and Rizzo, 2007), and kernel-based methods (Gretton et al., 2005, 2008), among many others. Recently, an innovative rank-based correlation coefficient was introduced by Chatterjee (2021) and Azadkia and Chatterjee (2021), which has quickly caught much attention in statistical community due to its simplicity and ability to detect non-monotonic or nonlinear association; see, e.g., Bickel (2022); Shi et al. (2024); Lin and Han (2023, 2024). Nonetheless, the direct application of existing coefficients to incompletely observed data remains unfeasible, highlighting the need for new approaches.

To our knowledge, there is currently no formally defined model-free correlation coefficient for right-censored data that satisfies both simplicity comparable to Pearson's correlation, and consistency as an estimator of a dependence measure – taking values in $[0, 1]$ and equaling 0 or 1 if and only if the variables are independent or one is a measurable function of the other. In this paper, inspired by Chatterjee (2021) and Azadkia and Chatterjee (2021), we developed a rank-based correlation coefficient tailored for censored data.

While there are many dependence indices for censored data, most are particularly designed for independence testing or variable screening, and thus lack the theoretical properties needed to accurately quantify association strength. Some approaches rely on semiparametric models, which can be affected by model misspecification; see, e.g., Cox (1972); Gray (1992); Zucker and Karr (1990). Alternatively, one can employ nonparametric methods, as illustrated by Le et al. (1994); McKeague et al. (1995); Ma and Mao (2019). Recent advancements have incorporated distance covariance (Székely et al., 2007; Székely and Rizzo, 2007), with Edelmann et al. (2022) developing a distance-covariance-type measure based on inverse-probability-of-censoring weighting (IPCW). Rindt et al. (2021) used the optimal transport to transform the censored dataset into an uncensored form, and then conducted independence testing using a kernel-based criterion. The other kernel-based approach is introduced by Fernández et al. (2023), who proposed a weighted log-rank test of independence between covariates and failure time. For association indices used in high-dimensional variable screening, see He et al. (2013); Song et al. (2014); Li et al. (2016); Pan et al. (2019), among others. However, as these existing dependence indices are specifically tailored to independence testing or variable screening, their population values are not theoretically justified as meaningful correlation measures. Additionally, distance covariance and kernel-based methods generally have a computational cost of $O(n^2)$.

In this paper, we introduce a censored rank-based correlation coefficient (CRC)

to measure the strength of association between survival outcomes and covariates. Building on Chatterjee (2021) and Azadkia and Chatterjee (2021), our contributions are twofold. First, we develop a new dependence measure that replaces the unknown distribution used in Chatterjee (2021) and Azadkia and Chatterjee (2021) with a pre-specified weight function. This substantially simplifies computation for complex data types, such as right-censored data, while retaining a strong ability to capture dependence. The new measure ranges between 0 and 1, and equals 0 or 1 if and only if the two random variables are independent or functionally related, respectively. Second, we construct a CRC for right-censored data. It is entirely model-free, as it does not depend on the distributions of the variables, and thus works robustly for a general class of survival models. Another advantage is that it can be computed in approximately $O(n \log n)$ time. Additionally, the CRC is shown to be strongly consistent and asymptotically normal, with a closed-form expression for its asymptotic variance under independence. The associated permutation test for independence is shown to be size-valid and power-consistent. In the AD data analysis, CRC tests identified two significant proteins, B7H1 and VEGF, associated with survival time. These proteins were not detected by other methods, highlighting the sensitivity of the CRC tests to oscillating signals.

The rest of the paper is organized as follows. Section 2 introduces the population-level dependence measure. Sections 3 and 4 present the proposed CRC and its asymp-

totic properties. Section 5 discusses its connection to Chatterjee's rank correlations. Section 6 studies CRC values under strictly increasing transformations. Section 7 describes the permutation test for independence. Section 8 reports simulation results, and Section 9 presents an application to ADNI data. Supplementary materials present technical proofs and additional simulation and real data results.

2. A New Measure of Dependence

Let T be a random variable in $\mathbb{R}$ with the law μ and X be an $\mathbb{R}^d$-valued random vector. To measure the dependence between T and X, we propose the following measure of dependence in population:

$$\xi(X,T) = \frac{\int \mathrm{var}[E\{\mathbb{1}(T \geq t) \mid X\}]\,\mathrm{d}W(t)}{\int \mathrm{var}\{\mathbb{1}(T \geq t)\}\,\mathrm{d}W(t)}, \tag{2.1}$$

where $\mathbb{1}(\cdot)$ is the indicator function and W is a pre-specified finite measure on the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, with $\mathcal{B}(\mathbb{R})$ representing the Borel σ-algebra on $\mathbb{R}$.

Let $\mathcal{S}_\mu$ and $\mathcal{S}_W$ denote the support of μ and W, respectively. Under the following assumption, we show that the dependence measure (2.1) possesses appealing properties in Theorem 1.

Assumption 1. $\mathcal{S}_\mu \subseteq \mathcal{S}_W$. *In addition, for any* $t \in \mathbb{R}$, *if* $\mu(\{t\}) > 0$, *we must have* $W(\{t\}) > 0$.

Theorem 1. *Suppose Assumption 1 holds and T is not almost surely a constant. Then, we have*

(i) $0 \leq \xi(X,T) \leq 1$;

(ii) (Consistency.) $\xi(X,T) = 0$ *if and only if* X *and* T *are independent;*

(iii) (Perfect dependency.) $\xi(X,T) = 1$ *if and only if* $T = h(X)$ *almost surely, for some measurable function* $h(\cdot)$.

Assumption 1 specifies the relationship between the supports of μ and the measure W, where the former is required to be a subset of the latter. Theorem 1 shows that the dependence measure defined in (2.1) is an effective tool for quantifying the strength of the relationship between X and T (Chatterjee, 2021). Note that the distance correlation does not satisfy the property of perfect dependency for nonlinear functions (Székely et al., 2007). For instance, when $X \sim N(0,1)$ and $T = X^2$, the distance correlation is strictly less than 1, as implied by Theorem 3 (iii) in Székely et al. (2007); see also the simulation results in Section 8.2.

Now, we provide an overview of our motivation and offer some insights into the idea behind the dependence measure (2.1) in the following two remarks.

Remark 1 (*Motivations and insights on (2.1)*)**.** If we take $W(\cdot) = \mu(\cdot)$, then (2.1) is exactly the dependence measure introduced by Dette et al. (2013), defined as:

$$\eta(X,T) = \frac{\int \mathrm{var}[E\{\mathbb{1}(T \geq t) \mid X\}]\mathrm{d}\mu(t)}{\int \mathrm{var}\{\mathbb{1}(T \geq t)\}\mathrm{d}\mu(t)}. \tag{2.2}$$

When T is a continuous random variable, one can write $\mathrm{d}\mu(t) = f_T(t)\mathrm{d}t$, where $f_T(t)$ is the density of T. Therefore, $f_T(t)$ can be regarded as a weight function for all possible values of T. However, $\mu(t)$ is usually unknown in practice and thus is replaced by its empirical counterpart as Chatterjee (2021). In the presence of missing or censored data, approximating $\mu(t)$ becomes much more complex and time-consuming, leading to an estimator that carries a heavy computational burden or yield lower accuracy. To avoid this issue and broaden the applicability of (2.2), we consider to replace μ with a pre-specified measure W that satisfies Assumption 1.

Remark 2 (*Discussion on the measure* W)**.** The form of W may vary across applications as long as Assumption 1 holds. For a continuous T, choosing $\mathrm{d}W(t) = w(t)\,\mathrm{d}t$ with $w(t) > 0$ for all $t \in \mathcal{S}_\mu$ ensures Assumption 1. For complete data, Chatterjee (2021) take $\mathrm{d}W(t) = \mathrm{d}\mu(t)$ and replace μ by its empirical distribution. For right-censored data, let C denote the censoring variable. When T is discrete, we set $\mathrm{d}W(t) = S_c^2(t)\,\mathrm{d}W_0(t)$, where W_0 is a pre-specified finite measure, $S_c(t) = P(C \geq t)$, and $\mathcal{S}_\mu$ is the support of μ. Let $\mathcal{S}_{W_0}$ denote the support of W_0. Assumption 1 is implied by:

Condition (a): $\mathcal{S}_\mu \subseteq \mathcal{S}_{W_0}$. In addition, for any $t \in \mathbb{R}$, if $\mu(\{t\}) > 0$, then $W_0(\{t\}) > 0$ and $S_c(t) > 0$.

For a continuous T, we use $\mathrm{d}W(t) = S_c^2(t)\,w(t)\,\mathrm{d}t$ as in Section 3. Then $w(t) > 0$ and $S_c(t) > 0$ for all $t \in \mathcal{S}_\mu$ guarantee Assumption 1; see Remark 3 for further details.

3. A Rank-Based Correlation Coefficient for Right-Censored Data

Let (X, T, C) be a collection of random variables taking values on $\mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+$, where T is the survival outcome of interest, C is a censoring variable, and X is a d-dimensional covariate vector with $d \geq 1$. In practice, due to the end of study follow-up or withdrawal from the study, we can not directly observe samples from (X, T, C). Instead, only samples from (X, Y, δ) are observable, where $Y = \min(T, C)$ and $\delta = \mathbb{1}(T \leq C)$. Similar to Edelmann et al. (2022), throughout this paper, we assume that the censoring variable C is independent of T and X; see a discussion in Remark 6. Our goal is to construct a simple and model-free correlation coefficient for censored data that consistently estimates (2.1) using all censored information.

Let $\{(X_i, Y_i, \delta_i)\}_{i=1}^n$ be independent and identically distributed (i.i.d.) samples from (X, Y, δ). Here and after, T and C are assumed to be continuous random variables. We take $\mathrm{d}W(t) = S_c^2(t)w(t)\mathrm{d}t$, where $S_c(t) = P(C \geq t)$ and $w(t)$ is a pre-determined weight function satisfying $\int_{\mathbb{R}_+} w(t)\mathrm{d}t < \infty$. This setting of measure W is tailored for the right-censored survival data; see details in Remark 3. Let $N(i)$ index the nearest neighbor of X_i among $\{X_j\}_{j=1}^n$ under the Euclidean metric on $\mathbb{R}^d$, with ties breaking at random. Then, $(X_{N(i)}, Y_{N(i)}, \delta_{N(i)})$ is the observation with $X_{N(i)}$ being the nearest neighbor of X_i. To measure the association between T and X, we

define the following censored rank-based correlation coefficient (CRC):

$$\hat{\xi}_n(X,T) = \frac{\int_{\mathbb{R}_+} Q_n(t)w(t)\mathrm{d}t - \int_{\mathbb{R}_+} G_n(t)^2 w(t)\mathrm{d}t}{\int_{\mathbb{R}_+} \hat{S}_c(t)G_n(t)w(t)\mathrm{d}t - \int_{\mathbb{R}_+} G_n(t)^2 w(t)\mathrm{d}t} \triangleq \frac{T_{n1}-T_{n2}}{T_{n3}-T_{n2}}, \tag{3.1}$$

where $Q_n(t) = n^{-1}\sum_{i=1}^n \mathbb{1}(Y_i \geq t)\mathbb{1}(Y_{N(i)} \geq t)$, $G_n(t) = n^{-1}\sum_{i=1}^n \mathbb{1}(Y_i \geq t)$, and $\hat{S}_c(t)$ is the Kaplan-Meier estimator of $S_c(t)$, defined as

$$\hat{S}_c(t) = \prod_{j=1}^n \left(1 - \frac{1}{\sum_{i=1}^n \mathbb{1}(Y_i \geq Y_j)}\right)^{(1-\delta_j)\mathbb{1}(Y_j \leq t)}.$$

Define "$0/0 = 0$". To better illustrate the idea behind (3.1), we first deduce

$$\begin{aligned} T_{n1} &= \frac{1}{n}\sum_{i=1}^n \int_{\mathbb{R}_+} \frac{\mathbb{1}(Y_i \geq t)}{S_c(t)} \cdot \frac{\mathbb{1}(Y_{N(i)} \geq t)}{S_c(t)} S_c(t)^2 w(t)\mathrm{d}t \\ &= \frac{1}{n}\sum_{i=1}^n \int_{\mathbb{R}_+} \eta_i(t)\eta_i^*(t)\mathrm{d}W(t), \end{aligned}$$

where $\eta_i(t) = \mathbb{1}(Y_i \geq t)/S_c(t)$ and $\eta_i^*(t) = \mathbb{1}(Y_{N(i)} \geq t)/S_c(t)$. Similarly, we can rewrite

$$T_{n2} = \int_{\mathbb{R}_+} \bar{\eta}_n(t)^2 \mathrm{d}W(t), \quad T_{n3} = \int_{\mathbb{R}_+} \bar{\eta}_n(t)\Delta_n(t)\mathrm{d}W(t),$$

where $\bar{\eta}_n(t) = n^{-1}\sum_{i=1}^n \eta_i(t)$ and $\Delta_n(t) = \hat{S}_c(t)/S_c(t)$. Thus, (3.1) becomes

$$\hat{\xi}_n(X,T) = \frac{\int_{\mathbb{R}_+} n^{-1}\sum_{i=1}^n \eta_i(t)\eta_i^*(t)\mathrm{d}W(t) - \int_{\mathbb{R}_+} \bar{\eta}_n(t)^2\mathrm{d}W(t)}{\int_{\mathbb{R}_+} \bar{\eta}_n(t)\Delta_n(t)\mathrm{d}W(t) - \int_{\mathbb{R}_+} \bar{\eta}_n(t)^2\mathrm{d}W(t)}. \tag{3.2}$$

Note that

$$E\{\eta_i(t)\} = \frac{E\{\mathbb{1}(Y_i \geq t)\}}{S_c(t)} = \frac{E\{\mathbb{1}(T_i \geq t)\mathbb{1}(C_i \geq t)\}}{S_c(t)} = E\{\mathbb{1}(T_i \geq t)\}.$$

Therefore, our proposal in (3.1) or (3.2) indeed measures the dependence between X and T and can be viewed as a moment method of the measure (2.1), where we leverage a nearest-neighbor-based estimator to approximate the conditional probability.

Remark 3. The choice $\mathrm{d}W(t) = S_c^2(t)w(t)\,\mathrm{d}t$ avoids having $S_c(t)$ or $\hat{S}_c(t)$ in the denominator, which may become close to zero and lead to unstable terms. The use of pre-specified weight functions has a long tradition in survival analysis. For example, Fleming and Harrington (2013) introduced the $G^{\rho,\gamma}$ family of weights for weighted log-rank tests, where different choices emphasize different parts of the survival curve; see also Fernández et al. (2023) and the IPCW scheme, such as in Edelmann et al. (2022). Our method is similar in spirit: by choosing $w(t)$, one can emphasize early, late, or intermediate survival periods according to the scientific question. In practice, $w(t)$ can be taken, for instance, as the density of a normal or t-distribution; see Section 6. Extensive empirical results in Sections 8 and 9 show that the performance of $\hat{\xi}_n(X,T)$ is stable across a wide range of choices for $w(t)$.

Let $x \wedge y = \min\{x, y\}$. Define $F_W(t) = \int_0^t w(s)ds$, $l_i = \sum_{j=1}^n I(Y_j \geq Y_i)$, and $s_i = \sum_{j=1}^n I(Y_j \leq Y_i)$. Furthermore, let $(Y_{(i)}, \delta_{(i)}, l_{(i)})$ denote the ordered sample such

that $Y_{(1)} \leq \ldots \leq Y_{(n)}$. Through straightforward calculations, the terms in (3.1) can be rewritten as

$$T_{n1} = \frac{1}{n}\sum_{i=1}^{n} F_W(Y_i \wedge Y_{N(i)}), \quad T_{n2} = \frac{1}{n^2}\sum_{i=1}^{n} F_W(Y_i)(n-1+l_i-s_i), \tag{3.3}$$

$$T_{n3} = F_W(Y_{(1)}) + \frac{1}{n}\sum_{k=1}^{n-1}\left[(n-k)\{F_W(Y_{(k+1)}) - F_W(Y_{(k)})\}\prod_{j=1}^{k}\left(1-\frac{1-\delta_{(j)}}{l_{(j)}}\right)\right]. \tag{3.4}$$

Remark 4 (*Computational complexity at* $O(n\log n)$)**.** The proposed $\hat{\xi}_n(X,T)$ avoids estimating complex distribution functions, characteristic functions, or mutual information, eliminating the need for smoothing or tuning parameters. Therefore, as shown in equations (3.3) and (3.4), the computational time of $\hat{\xi}_n(X,T)$ is mainly spent on performing nearest neighbor searches using Euclidean distance or ranking, both of which have a time complexity of $O(n\log n)$; see Sharma et al. (2019). Table 7 provides empirical evidence of CRC's approximately linear computational time, highlighting its potential for fast computation in practice.

4. Asymptotic Properties of $\hat{\xi}_n(X,T)$

We now derive the asymptotic properties of $\hat{\xi}_n(X,T)$. Define $G(t) = P(Y \geq t)$, $G_X(t) = P(Y \geq t \mid X)$, $\widetilde{G}(t) = P(T \geq t)$ and $\widetilde{G}_X(t) = P(T \geq t \mid X)$. Then we have $G_X(t) = P(T \geq t \mid X)P(C \geq t) = \widetilde{G}_X(t)S_c(t)$. For $\mathrm{d}W(t) = S_c^2(t)w(t)\mathrm{d}t$, the

measure $\xi(X,T)$ in (2.1) can be written as

$$\begin{aligned}\xi(X,T) &= \frac{\int_{\mathbb{R}_+} E\{\widetilde{G}_X^2(t)\}\mathrm{d}W(t) - \int_{\mathbb{R}_+} \widetilde{G}(t)^2\mathrm{d}W(t)}{\int_{\mathbb{R}_+} \widetilde{G}(t)\mathrm{d}W(t) - \int_{\mathbb{R}_+} \widetilde{G}(t)^2\mathrm{d}W(t)} \\ &= \frac{\int_{\mathbb{R}_+} E\{G_X^2(t)\}w(t)\mathrm{d}t - \int_{\mathbb{R}_+} G(t)^2w(t)\mathrm{d}t}{\int_{\mathbb{R}_+} G(t)S_c(t)w(t)\mathrm{d}t - \int_{\mathbb{R}_+} G(t)^2w(t)\mathrm{d}t} \triangleq \frac{T_{01}-T_{02}}{T_{03}-T_{02}}.\end{aligned}$$

Let $T_{01}^* \triangleq E\{F_W(Y_1 \wedge Y_{N(1)})\}$ and $\xi_n^*(X,T) \triangleq (T_{01}^* - T_{02})/(T_{03}-T_{02})$. Denote the joint distribution of (X,T,C) by $F_{X,T,C}$. Under the following Assumption 2, we show the asymptotic properties of $\hat{\xi}_n(X,T)$.

Assumption 2. *Suppose T is not almost surely a constant and $\int_{\mathbb{R}_+} w(t)\mathrm{d}t < \infty$ with $w(t) \geq 0$ for all $t \in \mathbb{R}_+$.*

Theorem 2. *For $\mathrm{d}W(t) = S_c^2(t)w(t)\mathrm{d}t$, under Assumption 2, we have that*

(i) (Strong consistency.) $\hat{\xi}_n(X,T)$ converges to $\xi(X,T)$ almost surely, as $n \to \infty$;

(ii) (Asymptotic normality.) If $F_{X,T,C}$ is continuous and T is not a measurable function of X almost surely, then under assumptions (A1)–(A3) in Appendix, $n^{1/2}\{\hat{\xi}_n(X,T) - \xi_n^*(X,T)\}$ converges to a zero-mean normal distribution, where

$$\xi_n^*(X,T) = \xi(X,T) + O\left(\frac{(\log n)^{d+\beta+1+\mathbb{1}(d=1)}}{n^{1/d}}\right),$$

for some fixed constant $\beta > 0$ in Assumption (A2).

Remark 5. We note that the strong consistency of $\hat{\xi}_n(X,T)$ still holds for discrete T when one takes $\mathrm{d}W(t) = S_c^2(t)\,\mathrm{d}W_0(t)$ for some pre-specified finite measure W_0 with $\int_{\mathbb{R}_+} \mathrm{d}W_0(t) < \infty$. To ensure that the properties in Theorem 2.1 hold under this choice, Condition (a) in Remark 2 is required.

Note that $\xi_n^*(X,T)$ is a constant sequence converging to the dependence measure $\xi(X,T)$. In particular, when X and T are independent, we obtain a closed-form expression for the asymptotic variance of $\hat{\xi}_n(X,T)$ in the following proposition.

Proposition 1. *Suppose the conditions of Theorem 2 (ii) hold and X is abosolutely continuous. Under independence between X and T, we have $\xi_n^*(X,T) = 0$ for all $n \geq 1$, and $\sqrt{n}\,\hat{\xi}_n(X,T) \xrightarrow{d} N\left(0, \sigma_0^2\right)$, where*

$$\sigma_0^2 = \frac{(1+\mathfrak{q}_d)\sigma_F^2 + (\mathfrak{o}_d - 2 - 2\mathfrak{q}_d)\sigma_\phi^2}{D^2}. \tag{4.5}$$

Here, $\sigma_F^2 = \mathrm{var}\big(F_W(Y_1 \wedge Y_2)\big)$ and $\sigma_\phi^2 = \mathrm{var}\big(\phi(Y_1)\big)$ with Y_1, Y_2 being i.i.d. random variables with $G(t) = P(Y \geq t)$, $\phi(y) = \int_0^y G(t)\,w(t)\,\mathrm{d}t$, and $D = T_{03} - T_{02}$. Addi-

tionally, for any $d \geq 1$, $\mathfrak{q}_d$ and $\mathfrak{o}_d$ are positive constants depending only on d:

$$\mathfrak{q}_d = \left\{2 - I_{3/4}\left(\frac{d+1}{2}, \frac{1}{2}\right)\right\}^{-1}, \quad I_x(a,b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}\mathrm{d}t}{\int_0^1 t^{a-1}(1-t)^{b-1}\mathrm{d}t},$$

$$\mathfrak{o}_d = \int_{\Gamma_{d;2}} \exp\left[-\lambda\left\{B(\boldsymbol{x}_1, \|\boldsymbol{x}_1\|) \cup B(\boldsymbol{x}_2, \|\boldsymbol{x}_2\|)\right\}\right] \mathrm{d}(\boldsymbol{x}_1, \boldsymbol{x}_2),$$

$$\Gamma_{d;2} = \left\{(\boldsymbol{x}_1, \boldsymbol{x}_2) \in (\mathbb{R}^d)^2 : \max(\|\boldsymbol{x}_1\|, \|\boldsymbol{x}_2\|) < \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|\right\},$$

where $B(\boldsymbol{x}_1, r)$ is the ball of radius r centered at $\boldsymbol{x}_1$, and $\lambda(\cdot)$ is the Lebesgue measure.

In particular, in the absence of censoring ($C = \infty$), we have $S_c(t) = 1$, $Y_i = T_i$, and $\delta_i = 1$. Taking $w(t) = f_T(t)$, it follows that (4.5) reduces to the asymptotic variance given in Theorem 3.1(ii) of Shi et al. (2024) for complete data. Define the plug-in estimator of σ_0^2 as

$$\hat{\sigma}_0^2 = \frac{(1+\mathfrak{q}_d)\hat{\sigma}_F^2 + (\mathfrak{o}_d - 2 - 2\mathfrak{q}_d)\hat{\sigma}_\phi^2}{\hat{D}^2},$$

where $\hat{D} = T_{n3} - T_{n2}$, $\hat{\sigma}_\phi^2 = n^{-1}\sum_{i=1}^n \hat{\phi}(Y_i)^2 - T_{n2}^2$ with $\hat{\phi}(y) = \int_0^y G_n(t)w(t)\,\mathrm{d}t$, and $\hat{\sigma}_F^2 = \binom{n}{2}^{-1}\sum_{i<j} F_W(Y_i \wedge Y_j)^2 - T_{n2}^2$. The next corollary shows the consistency of $\hat{\sigma}_0^2$.

Corollary 1. *Assume all conditions in Proposition 1 hold. Under independence between X and T, $\hat{\sigma}_0^2 \to \sigma_0^2$ almost surely.*

Remark 6. Since the terms in T_{n1} are no longer i.i.d., justifying the asymptotic

normality of Chatterjee-type rank correlations is non-trivial, even for complete data under independent X and T (Azadkia and Chatterjee, 2021; Lin and Han, 2022; Deb et al., 2020; Shi et al., 2024). For theoretical tractability, we assume that C is independent of T and X. This ensures $\hat{\xi}_n(X,T)$—the first fully model-free correlation coefficient for censored data that can measure the strength of associations—retains broad applicability (e.g., in biomedicine) and $O(n\log n)$ computation. While replacing $S_c(t)$ with some $S_c(t|X)$ could relax this assumption to $C \perp T|X$, such an extension introduces theoretical complexities (e.g., nonparametric estimation of $S_c(t|X)$) and additional computational costs, which we leave for future work.

Remark 7. The consistency of $\hat{\sigma}_0^2$ under dependence between X and T remains unknown, and constructing a consistent variance estimator is highly nontrivial even in the complete-data setting (Lin and Han, 2022). Therefore, for testing dependence in practice, we recommend the permutation procedure described in Section 7. This test is finite-sample exact in size (for any n and B) and power-consistent, while avoiding the need for variance estimation. Moreover, it is computationally efficient; see Table 7. Extensive simulation studies in Section 8.3 and in the supplementary material demonstrate its satisfactory finite-sample performance. In particular, the m-out-of-n bootstrap approach by Dette and Kroll (2025) may be adapted to our setting, although extending it to censored outcomes would require additional technical development. We therefore view this as an important direction for future research.

5. Connections to Chatterjee's Rank Correlations

For fully observable data, Azadkia and Chatterjee (2021) and Chatterjee (2021) proposed the following two rank correlation coefficients, respectively:

$$\hat{\eta}_n(X,T) = \frac{\sum_{i=1}^n (n \min\{R_{N(i)}, R_i\} - L_i^2)}{\sum_{i=1}^n L_i(n-L_i)}, \quad \text{for } d \geq 1, \tag{5.6}$$

$$\hat{\eta}_n^*(X,T) = 1 - \frac{n\sum_{i=1}^n |R_{\tilde{N}(i)} - R_i|}{2\sum_{i=1}^n L_i(n-L_i)}, \quad \text{for } d = 1, \tag{5.7}$$

where R_i is the rank of T_i, L_i is the number of j such that $T_j \geq T_i$, and $N(i)$, $\tilde{N}(i)$ index the nearest neighbor of X_i among $\{X_j\}_{j=1}^n$ and the right nearest neighbor of X_i among $\{X_j\}_{j=1}^n$ when $d = 1$.

When the censoring variable $C = \infty$, there are no censored observations. In this scenario, all survival outcomes or failure times are completely observed, namely, $Y_i = T_i$ and $\delta_i = 1$ for all $i = 1, \ldots, n$. It follows that the Kaplan-Meier estimator $\hat{S}_c(t) = 1$. In this case, set $\mathrm{d}W(t) = \mathrm{d}F_n(t)$, where $F_n(t) = n^{-1}\sum_{i=1}^n \mathbb{1}(T_i \leq t)$. By some simple algebra, we have

$$n^2 \sum_{i=1}^n \{F_n(T_i) \wedge F_n(T_{N(i)})\} = n \sum_{i=1}^n \{R_i \wedge R_{N(i)}\}, \quad \text{and}$$

$$n^3 \int_{\mathbb{R}_+} G_n(t)^2 \mathrm{d}F_n(t) = \sum_{i=1}^n L_i^2, \quad n^3 \int_{\mathbb{R}_+} G_n(t) \mathrm{d}F_n(t) = n \sum_{i=1}^n L_i.$$

Therefore, in the case of no censoring, our coefficient $\hat{\xi}_n(X,T)$ reduces to $\hat{\eta}_n(X,T)$ in (5.6).

On the other hand, for $d=1$, when T and X have a continuous joint distribution function, the correlation coefficient $\hat{\eta}_n^*(X,T)$ in (5.7) can be simplified to:

$$\hat{\eta}_n^*(X,T) = 1 - \frac{3}{n^2-1}\sum_{i=1}^{n} |R_{\tilde{N}(i)} - R_i|, \tag{5.8}$$

where we replace the nearest neighbor index $N(i)$ of i in (3.1) with its right nearest neighbor index $\tilde{N}(t)$. Similar to Remark 3 in Lin and Han (2023), one can easily show that $\hat{\xi}_n(X,T)$ reduces to (5.8) with an asymptotically ignorable small term. More precisely,

$$\begin{aligned}\hat{\xi}_n(X,T) &= \hat{\eta}_n^*(X,T) + \frac{3}{n^2-1}\sum_{i=1}^{n}(R_{\tilde{N}(i)} - R_i) \\ &= \hat{\eta}_n^*(X,T) + O(n^{-1}),\end{aligned}$$

where the last equality holds by noting that $\{R_i\}_{i=1}^n$ is a permutation of $\{1,\ldots,n\}$. Then, for any $i,j \in \{1,\ldots,n\}$, the difference between R_i and R_j is at most $n-1$, resulting in $|\sum_{i=1}^n (R_{\tilde{N}(i)} - R_i)| = |R_i - R_j| \le n-1$, for some $i,j \in 1,\ldots,n$.

6. Under Strictly Increasing Transformations

Note that the construction of $\hat{\xi}_n(X,T)$ does not depend on the distributions of X or T. Suppose that $h(\cdot)$ is a strictly increasing mapping from $\mathbb{R}_+$ to $\mathcal{Y} \in \mathcal{B}(\mathbb{R})$. Denote the transformed data as $T^* = h(T)$ and $C^* = h(C)$. Now, we figure out the relationship between the coefficients before and after the transformation.

Suppose $F_W(x) = \int_0^x w(t)\mathrm{d}t$ is a probability measure that satisfies Assumption 1. Write $F_W(x) = P(Z \leq x)$ for all $x \in \mathbb{R}$, where Z is a random variable mapping from some sample space to $\mathbb{R}_+$. Then, we have $F_W(x) = P(Z \leq x) = P(h(Z) \leq h(x)) = F^*_W(h(x))$, where $F^*_W(u)$ is the distribution of the random variable $h(Z)$ satisfying $\int_{\mathcal{Y}} \mathrm{d}F^*_W(u) < \infty$. By equations (3.3) and (3.4), it is easy to see that

$$\hat{\xi}_n(X, T^*, F^*_W) = \hat{\xi}_n(X, T, F_W). \tag{6.9}$$

This implies that the proposed coefficient before and after a strictly increasing transformation of both T and C are equivalent, up to different well-defined probability measures as (2.1). In particular, for linear transformation models, we have $h(t) = \log(t)$. According to (6.9), if we set $F^*_W(t) = F_W(e^t)$, then $\hat{\xi}_n(X, \log(T), F^*_W) = \hat{\xi}_n(X, T, F_W)$. Furthermore, if $F_W(t)$ is taken to be the log-normal distribution with parameters μ and σ^2, then $F^*_W(t)$ is the normal distribution with the same parameters. Similarly, if $F_W(t)$ is the log-t distribution, then $F^*_W(t)$ is the t-distribution

with the same degrees of freedom.

Therefore, in practice, according to (6.9), one can apply a standardized log-transformation to the data to remove the scale effect, while fixing $F_W(t)$ to be, for example, the standard normal distribution or t-distribution. Extensive empirical results in Sections 8 and 9 indicate that the performance of $\hat{\xi}_n(X,T)$ is stable across commonly used choices of $F_W(t)$.

7. Application to Independence Testing

While $\hat{\xi}_n(X,T)$ primarily measures the strength of the relationship between X and T, rather than serving as a test statistic for independence, it can still be used for independence testing if desired. For the null hypothesis $H_0 : \xi(X,T) = 0$, we construct a permutation independence test. Specifically, given the number of permutations B, in each round $b \in \{1,\ldots,B\}$, we draw a permuted sample from the uniform distribution over all possible permutations on $\{1,\ldots,n\}$, denoted by $\{(X_i^{(b)}, Y_i, \delta_i)\}_{i=1}^n$. Then, we evaluate $\hat{\xi}_n^{(b)}(X,T)$ for each permutation as follows:

$$\hat{\xi}_n^{(b)}(X,T) = \frac{\int_{\mathbb{R}_+} Q_n^{(b)}(t)w(t)\mathrm{d}t - \int_{\mathbb{R}_+} G_n(t)^2 w(t)dt}{\int_{\mathbb{R}_+}\{\hat{S}_c(t) - G_n(t)\}G_n(t)w(t)\mathrm{d}t},$$

where $Q_n^{(b)}(t) = n^{-1}\sum_{i=1}^n \mathbb{1}(Y_i \geq t)\mathbb{1}(Y_{N_b(i)} \geq t)$ with $X_{N_b(i)}$ being the nearest neighbor of $X_i^{(b)}$ for pairs $(X_i^{(b)}, Y_i, \delta_i)$ and $(X_{N_b(i)}^{(b)}, Y_{N_b(i)}, \delta_{N_b(i)})$ in the permutation sam-

ples. Under the null, $\hat{\xi}_n^{(b)}(X,T)$ will have the same distribution as $\hat{\xi}_n(X,T)$ because of the independence between X and T and the distribution-freeness of the proposed correlation coefficient. For a given significance level $\alpha \in (0,1)$, the permutation test is defined as

$$T_{\alpha,B}^n = \mathbb{1}\left((B+1)^{-1}\left[1+\sum_{b=1}^{B}\mathbb{1}\{\hat{\xi}_n^{(b)}(X,T) \geq \hat{\xi}_n(X,T)\}\right] \leq \alpha\right).$$

Let P_0 and P_1 denote the product measures of samples with distributions satisfying or violating H_0, respectively. The size validity and power consistency of the test $T_{\alpha,B}^n$ are shown in the following theorem.

Theorem 3. *Suppose that the Assumptions 1 and 2 hold. Then, we have that*

(i) (Size validity.) Under the null, it holds that, for any $n \geq 1$ and $B \geq 1$,

$$P_0(T_{\alpha,B}^n = 1) \leq \alpha.$$

(ii) (Power consistency.) Under the alternative, we have that, for $B = B_n \to \infty$ as $n \to \infty$,

$$\lim_{n\to\infty} P_1(T_{\alpha,B}^n = 1) = 1. \tag{7.10}$$

Simulation studies indicate satisfactory performance of $T_{\alpha,B}^n$ as in Section 8.3.

8. Simulation Studies

8.1 Validity of the Asymptotics

We begin by examining the consistency and asymptotic normality of the proposed CRC method. We model the association between the covariates and the right-censored outcome as both linear and nonlinear. Specifically, we consider the following data generating process (DGP):

$$\text{DGP1}: \quad \log(T) = 2X_1 + 2X_2 + 4X_3 + \epsilon,$$

where the covariates $X = (X_1, \cdots, X_p)^{\mathrm{T}}$ follow the multivariate normal distribution with mean zero and the covariance matrix $\Sigma = (\sigma_{ij})_{p\times p}$ with $p = 3$, $\sigma_{ii} = 1$ and $\sigma_{ij} = \rho,\ i \neq j$, and $\epsilon \sim N(0, 5^2)$ independent of X. In this case, we consider different values of $\rho = 0, 0.4$ and 0.8. The logarithm of censoring variable C is generated from a mixture of the normal and uniform distributions: $N(\mu, 1) + cU(0, 3)$. The censoring rate is around 30%.

Then, we consider the following nonlinear model:

$$\text{DGP2}: \quad T = X^2 + \epsilon,$$

where the covariates X is drawn from standard normal $N(0, 1)$ and ϵ follows from the

chi-squared distribution with the degree of freedom being one and two, respectively. The censoring variable C is assumed to follow the mixture distribution of gamma and uniform distributions, namely, $c_0\text{Gamma}(k,\theta) + c_1 U(0,1)$. The censoring rate is approximately 10%.

Sample sizes of 200, 400, and 800 were used, with 1000 repetitions performed for each size. We utilize various forms of $w(t)$ to calculate the dependence measure. These include the exponential distribution with rate one (Exp(1)), the chi-squared distribution ($\chi^2(4)$), as well as log-type distributions. In Tables 1 and 2, the notation "log-$f(\cdot)$" denotes the log-type distributions characterized by $w(t) = f(\log x)/x$, where $f(\cdot)$ corresponds to the densities of normal distribution, uniform distribution and t distribution, respectively. For comparison, we also set $w(t)$ to be the true density $f_T(t)$ of T, though it is typically unknown in practice.

The estimation results presented in Tables 1 and 2 indicate that the bias and standard deviations exhibit stability across different cases and various forms of $w(t)$. In each scenario, the values of calculated dependence measures based on different $w(t)$ are similar. Furthermore, as observed in Tables 1 and 2, stronger dependence between covariates and T leads to larger values of the dependence measure. In addition, we observe from Figures S.1–S.2 in the supplementary material that the empirical density of $\sqrt{n}\hat{\xi}_n(X,T)$ fits well with a normal distribution.

8.2 Perfect Dependency Estimation

It is worth noting that the CRC is capable of detecting perfect dependency, as established in Theorem 1. To further illustrate this property, we consider the following DGP3:

(a) $T = 4X + 3, \quad X \sim N(0, 1)$ (linear)

(b) $T = X^2 + 1, \quad X \sim N(0, 1)$ (non-monotone)

(c) $T = |\sin(2\pi X)| + 0.5, \quad X \sim U(0, 3)$ (oscillatory)

(d) $T = \exp(\sin(X)), \quad X \sim U(-\pi, \pi)$ (nonlinear and non-monotone)

The censoring variable C is generated from $C \sim \text{Exp}(\lambda_c)$, where λ_c is calibrated so that the censoring rate is approximately 0%, 20%, or 40%. The results are summarized in Tables 3-6. It can be seen that the proposed CRC can effectively detect the correlation when the true value is 1, whereas the distance correlation coefficient with IPCW ($\text{dCor}_{\text{IPCW}}$) is less than 1, as expected. For the linear case, both coefficients are close to 1, while for case (c) in Table 5, the distance correlation is close to 0.08.

8.3 Permutation Tests and Run Times

Here, we evaluate the performance of the proposed permutation tests based on our proposed CRC with various forms of $w(t)$, including the standard normal distribution, the uniform distribution and the t distribution. For comparison, we also

perform the Cox's proportional hazards model based test (CPH), the kernel log-rank test (KLR; Fernández et al. (2023)) and the IPCW-based distance covariance permutation test (IPCW; Edelmann et al. (2022)). For KLR, we choose Gaussian kernels as the kernel functions for time and covariates with the number of bootstrap $M = 2000$ as in Fernández et al. (2023). The number of permutations is set to be $B = 500$ for all other tests.

Let $\tilde{Y}$ be the log-transformed right-censored data. The error ϵ generated from the standard normal is independent of covariates. Then we consider the data generated from the following settings:

DGP4 (Linear): $\tilde{Y} = 0.8(-X_1 - X_2 + X_3) + \epsilon$, where the covariates $X = (X_1, \cdots, X_p)^{\mathrm{T}}$ follow the multivariate normal distribution with mean zero and the covariance matrix $\Sigma = (\sigma_{ij})_{p\times p}$ with $p = 3$, $\sigma_{ii} = 1$ and $\sigma_{ij} = 0.2$. The censoring variable C is generated from the Gaussian mixture model $N(\mu, 0.3^2) + N(1, 1)$.

DGP5 (Step): $\tilde{Y} = f(X) + \epsilon$, where $X \sim U(-2, 2)$ and f takes values $-2, 4, -4, 2, -2$, $2.5, -3$ and 3 in the intervals $[-2, -1.5), [-1.5, -1), [-1, -0.5), [-0.5, 0), [0, 0.5)$, $[0.5, 1), [1, 1.5)$ and $[1.5, 2)$. The censoring variable C is generated from $N(2.3, 0.5^2) - N(\mu, 0.1^2)$.

DGP6 (W-shaped): $\tilde{Y} = 5(|X + 1.5|\mathbb{1}\{X < 0\} + |X - 1.5|\mathbb{1}\{X \geq 0\}) + \epsilon_0$, where $X \sim U(-3, 3)$ and $\epsilon_0 \sim logit(0, 1)$. The censoring variable C is generated

from $2N(1, 0.5^2) - N(\mu, 0.1^2)$.

DGP7 (Modulus): $\tilde{Y} = 5(X - \lfloor X \rfloor) - 1 + \epsilon$, where $X \sim U(-2, 2)$. The censoring variable C is generated from $N(2, 0.5^2) - N(\mu, 0.1^2)$.

DGP8 (Shubert): $\tilde{Y} = 1.5\big[\sum_{k=1}^{5} k\cos\{(k+1)X_1\}\big]\big[\sum_{k=1}^{5} k\cos\{(k+1)X_2\}\big] + \epsilon$, where both X_1 and X_2 follow $U(-2, 2)$ independently. The censoring variable C is generated from the Gaussian mixture model $N(0, 4^2) - N(\mu, 1) + 0.5N(1, 1)$.

DGP9 (Rastrigin): $\tilde{Y} = 10 + \{X_1^2 - 5\cos(2\pi X_1)\} + \{X_2^2 - 5\cos(2\pi X_2)\} + \epsilon$, where both X_1 and X_2 follow $U(-1, 1)$ independently. The censoring variable C is generated from $N(2, 4^2) - N(0, 1) + 5N(\mu, 0.1)$.

To further assess the performance of our CRC test, we also consider commonly used models in survival analysis in Section S2 of the supplementary material. Specifically, we study nonlinear Cox proportional hazards models (DGP10) and accelerated failure time (AFT) models with heterogeneous errors (DGP11).

For DGPs 4–9, Figure 1 presents empirical rejection rates for all tests at a significance level of 0.05, with censoring rates of approximately 45% and 65%, respectively. Tables S.1–S.4 in the supplementary material record the empirical type-I error rates at 45% and 65% censoring rates for DGPs 4–11. The power performance results for DGPs 10 and 11 are presented in Figure S.4 of the supplementary material.

Three main observations can be made from Figures 1 and Tables S.1–S.4: (i) our CRC-based tests demonstrate significant power improvement when the covariates exhibit nonlinear effects or more complex non-additive structures, particularly oscillatory (wavelike) patterns, but perform less effectively in detecting linear structures; (ii) the performance of tests utilizing various forms of $w(t)$ is quite similar; (iii) the size performance of all tests remains stable. Furthermore, the results in Figure S.3 of the supplementary material further reveal the distinct characteristics of each test. Specifically, the KLR test is more sensitive to the period or frequency of the cosine function but remains stable under heavy censoring, while the IPCW method exhibits the opposite behavior. In contrast, the CRC method shows greater power for data with higher oscillatory frequencies but is slightly affected by heavy censoring. In summary, these observations align with the performance of Chatterjee-type tests of independence: these tests excel at handling oscillatory signals, yet may exhibit inferior performance with linear associations.

At last, we compare the computation times in Table 7 for DGP4 with $\rho = 0.4$ and the censoring rate around 10% and 30%. We observe that the computation times of the CRC tests with $B = 500$ are similar to those of the CPH tests and increase approximately linearly with sample size n, while those of IPCW and KLR exhibit growth at $O(n^2)$.

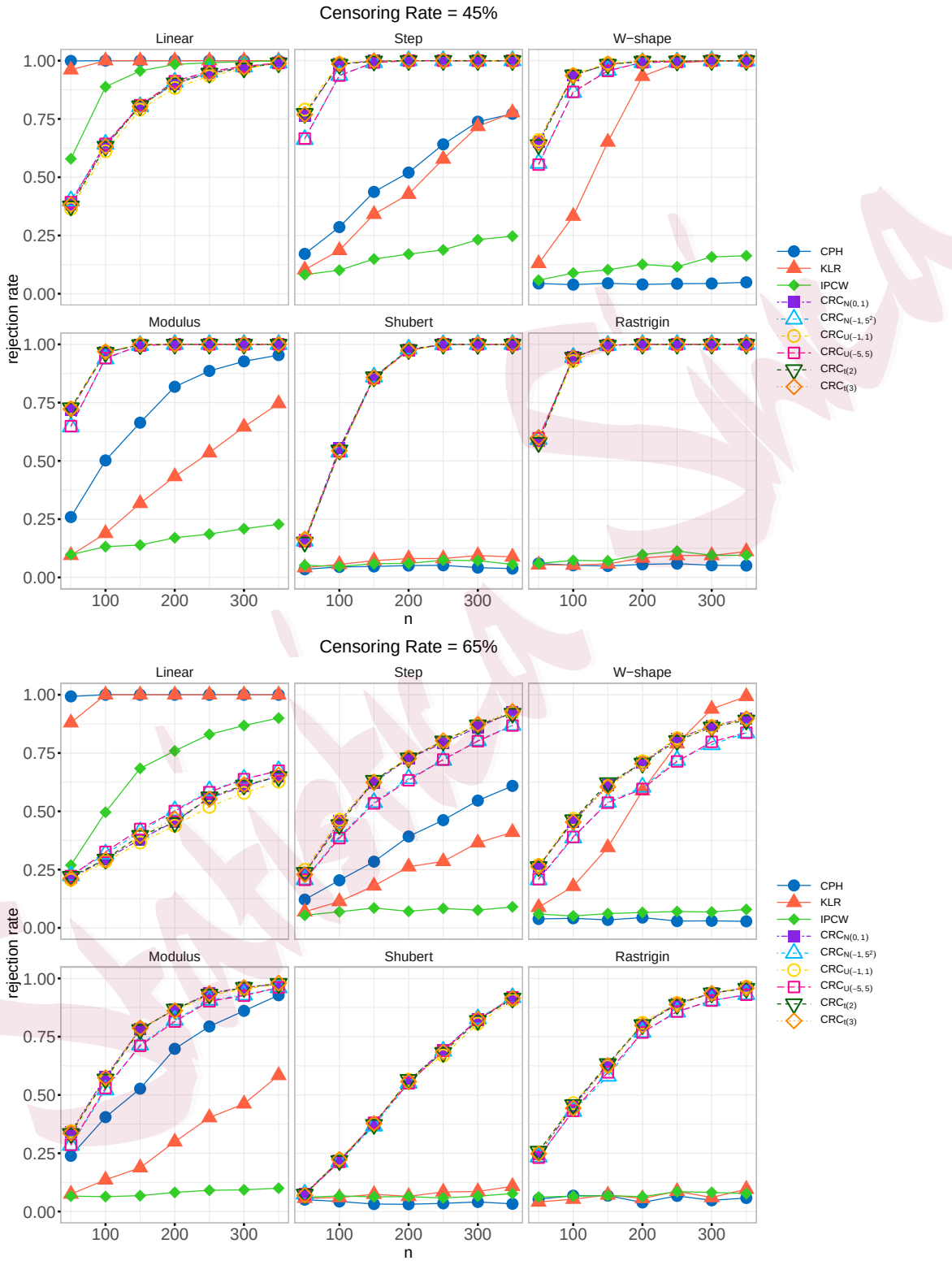

Figure 1: Rejection rates of all tests in DGPs 4–9 with the censoring rates 45% and 65%, respectively. The dashed lines represent our tests with various $w(t)$.

9. Application to the ADNI Data

We applied the proposed method to the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The clinical, demographic and protein data used for this study were obtained through the ADNI data repository (http://adni.loni.ucla.edu/). The survival status was defined as progression from mild cognitive impairment (MCI) to Alzheimer's disease (AD), with survival time calculated as the interval from initial MCI diagnosis to AD conversion. Participants who remained AD-free at the end of follow-up were right-censored. We analyzed associations between survival time and cerebrospinal fluid (CSF) protein levels quantified via SomaLogic Somascan assay (7,008 aptamers/analytes for 6,163 unique proteins that passed quality control) by Shen et al. (2024) and Cruchaga et al. (2024). After aligning ADNI clinical data with proteomic measurements, 109 MCI participants aged between 55 to 89 were included, each with 7029 protein levels. The event was observed for 66 patients (censoring rate around 39.4%).

The proposed CRC methods were benchmarked against the CPH test, the KLR test and the IPCW method. These comparisons were systematically applied to all 7,008 protein analytes to ensure comprehensive evaluation. For the CRC, KLR, and IPCW methods, the number of permutations B was dynamically determined as $\min\{50/p_0, 10^7\}$, where p_0 denotes the raw (unadjusted) p-value. This adaptive approach ensures sufficient permutation replicates to guarantee p-value accuracy while

balancing computational efficiency. To compare the testing methods at a protein-wise significance level of 0.1, we applied the Benjamini-Hochberg procedure to control the false discovery rate at 0.1. The numbers of rejections for an FDR of 0.2, 0.1, and 0.05, respectively, are illustrated in Figure S.5 of the supplementary material.

From the results, the CPH, KLR, and IPCW methods cannot detect proteins that passed FDR < 0.2, while the CRC methods with $w(t)$ as the densities of $N(0,1)$ and $U(-1,1)$ identified 6 and 5 signals using the threshold FDR < 0.2, respectively. Among these, the VEGF and B7-H1 protein analytes pass the FDR < 0.1, and B7-H1 passed the FDR < 0.05 with weight function $U(-1,1)$. This may suggest the higher power of CRC methods. Furthermore, a complete list of the top 100 most significant signals in CRC methods are shown in Figure S.6. It can be observed that protein analytes with higher CRC values generally exhibited greater significance.

The top identified protein analytes associated with the alzheimer's disease (AD) survival status by the CRC method are B7H1 and VEGF. B7-H1, also known as programmed death-ligand 1 (PD-L1), is an immune checkpoint molecule that plays a crucial role in regulating immune responses. Emerging results support that the Alzheimer disease is affected by the ability of the immune system to contain the brain's pathology (Schwartz et al., 2019). The VEGF is a crucial protein involved in angiogenesis and plays a vital role in maintaining brain function and protecting against neurodegeneration. Many evidences suggest that modulation of VEGF ex-

pression is a potential mechanism associated with the risk of developing AD and its clinical deterioration (Echeverria et al., 2017). To further characterize the dependencies detected by the CRC tests, we conducted an exploratory analysis of associations with FDR < 0.1. Cox models with penalized splines (via the `pspline` function in the R package `survival`) revealed nonlinear functional relationships between protein levels and the log-hazard ratio. As shown in Figure 2, B7-H1 and VEGF exhibited oscillatory (wavelike) effects. Such patterns are captured by the CRC methods but missed by conventional models, consistent with our simulation results and underscoring the sensitivity of CRC to oscillatory signals.

Appendix: Additional assumptions for asymptotic normality

Denote $\bar{Y}(t) = \sum_{i=1}^{n} \mathbb{1}(Y_i \geq t)$, i.e., the number of units at risk before t. The following assumptions are required:

(A1) $\liminf_{n\to\infty} \mathrm{var}[W_n] > 0$, where W_n defined in (S1.22) of the Supplementary Material is asymptotically equivalent to $\sqrt{n}\{\hat{\xi}_n(X,T) - \xi_n^*(X,T)\}$.

(A2) There exist fixed constants β, κ, C_1, $C_2 > 0$ such that for any $t \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$, $P(\|X\| \geq t) \leq C_1 e^{-C_2 t}$ and

$$\left|P(T \geq t \mid X = x) - P(T \geq t \mid X = x')\right| \leq \kappa(1 + \|x\|^\beta + \|x'\|^\beta)\|x - x'\|.$$

(A3) The distribution of the censoring variable C is absolutely continuous. As $n \to \infty$, $\sup_{0 \le t < \infty} |n^{-1}\bar{Y}(t) - P(Y \ge t)| \to 0$ in probability.

Assumption (A1) is necessary for the self-normalization central limit theorem. In fact, Lin and Han (2022) (Proposition 1.2) demonstrated that a similar assumption to (A1) holds when T is not a measurable function of X. Assumption (A2) is required to ensure the consistency of nearest-neighbor-based estimators for the conditional probability (Lin and Han, 2022). Assumption (A3) is needed for the asymptotic expansion of Kaplan-Meier estimator.

Supplementary Material

Supplementary materials present technical proofs and additional simulation and real data results.

Acknowledgments

Data used in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this paper. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. Linlin Dai and Tengfei Li contributed equally. Linlin Dai's research was supported by the General Project of

Ministry of Education Foundation on Humanities and Social Sciences (Nos. 24YJC910001) and the Sichuan Provincial Natural Science Foundation Project (Nos. 2025ZNSFSC0816). Kani Chen's research was supported by the grants T32-615-24-R and MOST24SC01.

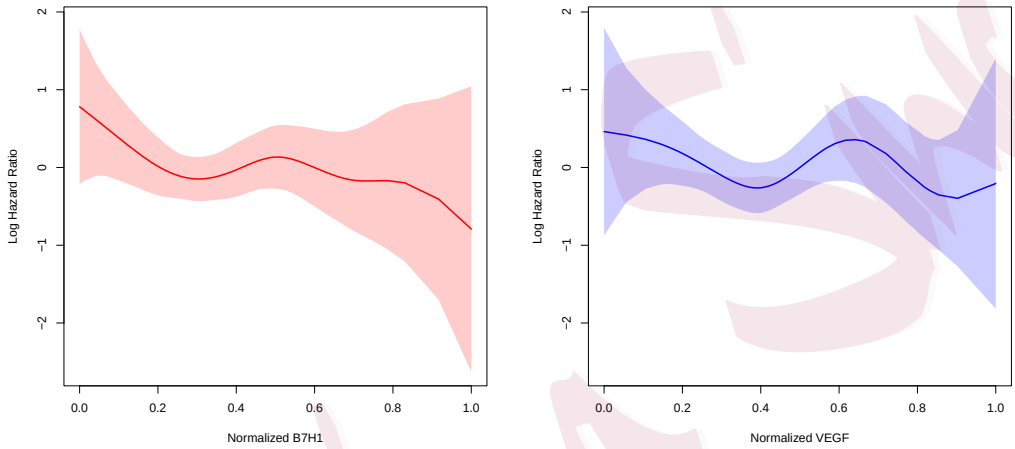

Figure 2: Nonlinear hazard function estimates for associations between survival time and protein levels of B7H1 and VEGF in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. Protein levels were normalized to the interval $[0, 1]$ for visualizations.

References

Azadkia, M. and Chatterjee, S. (2021). A simple measure of conditional dependence. *The Annals of Statistics*, 49, 3070–3102.

Bhaskaran, K., dos-Santos-Silva, I., Leon, D. A., Douglas, I. J. and Smeeth, L. (2018). Association of BMI

with overall and cause-specific mortality: a population-based cohort study of 3.6 million adults in the UK. *The lancet Diabetes & endocrinology*, 6, 944–953.

Biau, G. and Devroye, L. (2015). *Lectures on the Nearest Neighbor Method,* volume 246. Springer.

Bickel, P. J. (2022). Measures of independence and functional dependence. https://arxiv.org/abs/2206.13663.

Bitouzé, D., Laurent, B. and Massart, P. (1999). A Dvoretzky-Kiefer-Wolfowitz type inequality for the Kaplan-Meier estimator. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 35, 735–763.

Breiman, L. and Friedman, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580–598.

Chatterjee, S. (2008). A new method of normal approximation. *The Annals of Probability*, 36, 1584–1610.

Chatterjee, S. (2021). A new coefficient of correlation. *Journal of the American Statistical Association*, 116, 2009–2022.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 34, 187–220.

Deb, N., Ghosal, P. and Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. Available at arXiv:2010.01768.

Dette, H., Siburg, K. F. and Stoimenov, P. A. (2013). A copula-based non-parametric measure of regression dependence. *Scandinavian Journal of Statistics*, 40, 21–41.

Dette, H. and Kroll, M. (2025). A simple bootstrap for Chatterjee's rank correlation. *Biometrika*, 112, asae045.

Echeverria, V., E. Barreto, G., Ávila-Rodriguezc, M., V. Tarasov, V. and Aliev, G. (2017). Is VEGF a key target of cotinine and other potential therapies against Alzheimer disease? *Current Alzheimer Research*, 14, 1155–1163.

Edelmann, D., Welchowski, T. and Benner, A. (2022). A consistent version of distance covariance for right-censored survival data and its application in hypothesis testing. *Biometrics*, 78, 867–879.

Fernández, T., Gretton, A., Rindt, D. and Sejdinovic, D. (2023). A kernel log-rank test of independence for right-censored data. *Journal of the American Statistical Association*, 118, 925–936.

Fleming, T. R. and Harrington, D. P. (2013). *Counting processes and survival analysis* (Vol. 625). John Wiley & Sons.

Gambassi, G., Landi, F., Lapane, K. L., Sgadari, A., Mor, V. and Bernabei, R. (1999). Predictors of mortality in patients with Alzheimer's disease living in nursing homes. *Journal of Neurology, Neurosurgery & Psychiatry*, 67, 59–65.

Gray, R. J. (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, 87, 942–951.

Gretton, A., Bousquet, O., Smola, A. and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. in *Algorithmic Learning Theory*, Berlin: Springer, pp. 63–77.

Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2008). A kernel statistical test

of independence. *Advances in Neural Information Processing Systems*, 585–592.

Cruchaga, C., Ali, M., Shen, Y., Do, A., Wang, L., Western, D. et al. (2024). Multi-cohort cerebrospinal fluid proteomics identifies robust molecular signatures for asymptomatic and symptomatic Alzheimer's disease. *Research Square*, rs-3.

He, X., Wang, L. and Hong, H. G. (2013). Quantile-adaptive model-free variable screening for high-dimensional heterogeneous data. *The Annals of Statistics*, 41, 342–369.

Le, C. T., Grambsch, P. M. and Louis, T. A. (1994). Association between survival time and ordinal covariates. *Biometrics*, 50, 213–219.

Li, J., Zheng, Q., Peng, L. and Huang, Z. (2016). Survival impact index and ultrahigh-dimensional model-free screening with survival outcomes. *Biometrics*, 72, 1145–1154.

Lin, Z. and Han, F. (2022). Limit theorems of Chatterjee's rank correlation. Available at arXiv:2204.08031.

Lin, Z. and Han, F. (2023). On boosting the power of Chatterjee's rank correlation. *Biometrika*, 110, 283–299.

Lin, Z. and Han, F. (2024). On the failure of the bootstrap for Chatterjee's rank correlation. *Biometrika*, 111, 1063–1070.

Ma, L. and Mao, J. (2019). Fisher exact scanning for dependency. *Journal of the American Statistical Association*, 114, 245–258.

McKeague, I. W., Nikabadze, A. M. and Sun, Y. Q. (1995). An omnibus test for independence of a survival time from a covariate. *The Annals of Statistics*, 23, 450–475.

Pan, W., Wang, X., Xiao, W. and Zhu, H. (2019). A generic sure independence screening procedure. *Journal of the American Statistical Association*, 114, 928–937.

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G.,Turnbaugh, P. J., Lander, E. S., Mitzenmacher, M. and Sabeti, P. (2011). Detecting novel associations in large datasets. *Science*, 334, 1518–1524.

Rindt, D., Sejdinovic, D. and Steinsaltz, D. (2021). A kernel-and optimal transport-based test of independence between covariates and right censored lifetimes. *The International Journal of Biostatistics*, 17, 331–348.

Schwartz, M., Arad, M. and Ben-Yehuda, H. (2019). Potential immunotherapy for Alzheimer disease and age-related dementia. *Dialogues in Clinical Neuroscience*, 21, 21–25.

Sharma, N., Chakrabarti, A. and Balas, V. E. (2019). Data management, analytics and innovation. *Proceedings of ICDMAI*, 1.

Shen, Y., Timsina, J., Heo, G., Beric, A., Ali, M., Wang, C. et al. (2024). CSF proteomics identifies early changes in autosomal dominant Alzheimer's disease. *Cell*, 187, 6309–6326.

Shi, H., Drton, M. and Han, F. (2024). On Azadkia-Chatterjee's conditional dependence coefficient. *Bernoulli*, **30**, 851–877.

Song, R., Lu, W., Ma, S. and Jeng, X. J. (2014). Censored rank independence screening for high-dimensional survival data. *Biometrika*, 101, 799–814.

Székely, G. J. and Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics*, 3,

1236–1265.

Székely, G. J., Rizzo, M. L. and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35, 2769–2794.

Zajacova, A. and Burgard, S. A. (2012). Shape of the BMI-mortality association by cause of death, using generalized additive models: NHIS 1986–2006. *Journal of aging and health*, 24, 191–211.

Zucker, D. M. and Karr, A. F. (1990). Nonparametric survival analysis with time-dependent covariate effects: a penalized partial likelihood approach. *The Annals of Statistics*, 18, 329–353.

Linlin Dai

Center of Statistical Research, School of Statistics and Data Science, Southwestern University of Finance and Economics, Chengdu, Sichuan, China.

E-mail: ldaiab@swufe.edu.cn

Tengfei Li

Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.

E-mail: tengfei_li@med.unc.edu

Kani Chen

Department of Mathematics and Industrial Engineering and Decision Analytics, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR.

E-mail: makchen@ust.hk

Table 1: Estimates of $\hat{\xi}_n(X,T)$ for DGP1 with various $w(t)$, as well as the true values of $\xi(X,T)$ (True), the bias (Bias) and the standard deviations (SD)

			$n=200$		$n=400$		$n=800$	
ρ	$w(t)$	True	Bias	SD	Bias	SD	Bias	SD
0	$f_T(t)$	0.3023	−0.0153	0.0789	−0.0055	0.0569	0.0029	0.0374
	Exp(1)	0.3238	−0.0184	0.0915	−0.0132	0.0633	−0.0085	0.0459
	$\chi^2(3)$	0.3240	−0.0235	0.0959	−0.0123	0.0692	−0.0083	0.0474
	log-$U(-1,1)$	0.3257	−0.0204	0.0995	−0.0122	0.0695	−0.0079	0.0486
	log-$U(-5,5)$	0.3172	−0.0205	0.0855	−0.0119	0.0611	−0.0046	0.0401
	log-$N(0,1)$	0.3247	−0.0206	0.0921	−0.0123	0.0635	−0.0064	0.0431
	log-$N(-1,5^2)$	0.3090	−0.0182	0.0846	−0.0086	0.0585	−0.0022	0.0390
	log-$t(2)$	0.3228	−0.0224	0.0892	−0.0134	0.0621	−0.0062	0.0449
	log-$t(3)$	0.3236	−0.0215	0.0918	−0.0134	0.0642	−0.0063	0.0458
0.4	$f_T(t)$	0.3979	−0.0085	0.0737	−0.0050	0.0491	0.0037	0.0365
	Exp(1)	0.4207	−0.0132	0.0854	−0.0061	0.0607	−0.0045	0.0424
	$\chi^2(3)$	0.4207	−0.0147	0.0927	−0.0091	0.0642	−0.0056	0.0457
	log-$U(-1,1)$	0.4220	0.0153	0.0884	−0.0088	0.0658	−0.0050	0.0464
	log-$U(-5,5)$	0.4153	−0.0122	0.0792	−0.0068	0.0572	−0.0035	0.0401
	log-$N(0,1)$	0.4214	−0.0136	0.0916	−0.0063	0.0627	−0.0042	0.0446
	log-$N(-1,5^2)$	0.4074	−0.0100	0.0760	−0.0045	0.0516	−0.0011	0.0361
	log-$t(2)$	0.4197	−0.0154	0.0834	−0.0081	0.0613	−0.0050	0.0418
	log-$t(3)$	0.4202	−0.0141	0.0872	−0.0075	0.0596	−0.0043	0.0430
0.8	$f_T(t)$	0.4624	0.0034	0.0681	0.0027	0.0476	0.0015	0.0349
	Exp(1)	0.4847	−0.0073	0.0825	−0.0033	0.0594	−0.0013	0.0403
	$\chi^2(3)$	0.4850	−0.0085	0.0899	−0.0047	0.0647	−0.0018	0.0478
	log-$U(-1,1)$	0.4857	−0.0080	0.0892	−0.0035	0.0639	−0.0015	0.0434
	log-$U(-5,5)$	0.4805	−0.0080	0.0766	−0.0045	0.0519	−0.0012	0.0371
	log-$N(0,1)$	0.4852	−0.0078	0.0854	−0.0040	0.0576	−0.0011	0.0414
	log-$N(-1,5^2)$	0.4735	−0.0081	0.0704	−0.0032	0.0493	−0.0010	0.0334
	log-$t(2)$	0.4837	−0.0071	0.0794	−0.0046	0.0576	−0.0014	0.0401
	log-$t(3)$	0.4843	−0.0083	0.0803	−0.0039	0.0593	−0.0012	0.0408

Table 2: Estimates of $\hat{\xi}_n(X,T)$ for DGP2 with various $w(t)$, as well as the true values of $\xi(X,T)$ (True), the bias (Bias) and the standard deviations (SD)

				$n=200$		$n=400$		$n=800$	
	Model	n	True	Bias	SD	Bias	SD	Bias	SD
	$\epsilon\sim\chi^2_2$	$f_T(t)$	0.2110	−0.0101	0.0682	−0.0071	0.0506	−0.0038	0.0356
		Exp(1)	0.1784	−0.0091	0.0883	−0.0047	0.0621	−0.0026	0.0406
		$\chi^2(4)$	0.2186	−0.0084	0.0665	−0.0062	0.0490	−0.0029	0.0351
		log-$U(-1,1)$	0.1797	−0.0071	0.0848	−0.0040	0.0608	0.0025	0.0437
		log-$U(-5,5)$	0.1877	−0.0089	0.0769	−0.0052	0.0535	−0.0027	0.0383
		log-$N(0,1)$	0.1861	−0.0081	0.0762	−0.0053	0.0585	−0.0025	0.0383
		log-$N(-1,5^2)$	0.1872	−0.0084	0.0774	−0.0051	0.0538	−0.0031	0.0386
		log-$t(2)$	0.1766	0.0063	0.0793	0.0039	0.0563	0.0029	0.0397
		log-$t(3)$	0.1829	−0.0082	0.0819	0.0034	0.0561	0.0018	0.0392
	$\epsilon\sim\chi^2_1$	$f_T(t)$	0.3506	−0.0088	0.0622	−0.0044	0.0467	−0.0023	0.0334
		Exp(1)	0.3327	−0.0081	0.0698	−0.0042	0.0487	−0.0023	0.0360
		$\chi^2(4)$	0.3717	−0.0087	0.0685	−0.0053	0.0493	−0.0021	0.0341
		log-$U(-1,1)$	0.3468	−0.0067	0.0701	−0.0027	0.0495	0.0012	0.0342
		log-$U(-5,5)$	0.3168	−0.0075	0.0716	−0.0041	0.0477	−0.0018	0.0356
		log-$N(0,1)$	0.3423	−0.0081	0.0701	−0.0032	0.0470	−0.0017	0.0334
		log-$N(-1,5^2)$	0.3170	−0.0094	0.0722	−0.0040	0.0516	−0.0035	0.0363
		log-$t(2)$	0.3374	−0.0067	0.0678	−0.0035	0.0490	0.0015	0.0335
		log-$t(3)$	0.3397	−0.0075	0.0666	−0.0040	0.0486	−0.0017	0.032

Table 3: Bias and standard deviations (in parentheses) of $\hat{\xi}_n(X,T)$ for DGP3 (a)

			The proposed CRC $\hat{\xi}_n(X,T)$ with various $w(t)$					
CR	n	$\mathrm{dCor}_{\mathrm{IPCW}}$	$\mathrm{CRC}_{N(0,1)}$	$\mathrm{CRC}_{N(-1,5^2)}$	$\mathrm{CRC}_{U(-1,1)}$	$\mathrm{CRC}_{U(-5,5)}$	$\mathrm{CRC}_{t(2)}$	$\mathrm{CRC}_{t(3)}$
0%	200	0.0000	−0.0076	−0.0134	−0.0061	−0.0144	−0.0082	−0.0081
		(0.0000)	(0.0008)	(0.0030)	(0.0007)	(0.0037)	(0.0009)	(0.0009)
	400	0.0000	−0.0038	−0.0070	−0.0030	−0.0075	−0.0041	−0.0040
		(0.0000)	(0.0003)	(0.0014)	(0.0003)	(0.0017)	(0.0003)	(0.0003)
	800	0.0000	−0.0019	−0.0037	−0.0015	−0.0040	−0.0021	−0.0020
		(0.0000)	(0.0001)	(0.0006)	(0.0001)	(0.0008)	(0.0001)	(0.0001)
20%	200	−0.0060	−0.0086	−0.0138	−0.0070	−0.0145	−0.0090	−0.0089
		(0.0155)	(0.0556)	(0.0538)	(0.0563)	(0.0559)	(0.0558)	(0.0557)
	400	−0.0030	−0.0049	−0.0077	−0.0041	−0.0081	−0.0052	−0.0051
		(0.0103)	(0.0400)	(0.0377)	(0.0408)	(0.0391)	(0.0401)	(0.0400)
	800	−0.0015	−0.0018	−0.0035	−0.0013	−0.0037	−0.0019	−0.0019
		(0.0073)	(0.0275)	(0.0271)	(0.0276)	(0.0282)	(0.0276)	(0.0276)
40%	200	−0.0347	−0.0118	−0.0172	−0.0099	−0.0179	−0.0123	−0.0122
		(0.0391)	(0.0825)	(0.0739)	(0.0854)	(0.0769)	(0.0829)	(0.0827)
	400	−0.0182	−0.0083	−0.0114	−0.0074	−0.0119	−0.0086	−0.0085
		(0.0250)	(0.0575)	(0.0520)	(0.0595)	(0.0543)	(0.0577)	(0.0576)
	800	−0.0096	−0.0035	−0.0053	−0.0030	−0.0056	−0.0037	−0.0036
		(0.0162)	(0.0408)	(0.0370)	(0.0421)	(0.0386)	(0.0410)	(0.0409)

Table 4: Bias and standard deviations (in parentheses) of $\hat{\xi}_n(X,T)$ for DGP3 (b)

			The proposed CRC $\hat{\xi}_n(X,T)$ with various $w(t)$					
CR	n	$\mathrm{dCor}_{\mathrm{IPCW}}$	$\mathrm{CRC}_{N(0,1)}$	$\mathrm{CRC}_{N(-1,5^2)}$	$\mathrm{CRC}_{U(-1,1)}$	$\mathrm{CRC}_{U(-5,5)}$	$\mathrm{CRC}_{t(2)}$	$\mathrm{CRC}_{t(3)}$
0%	200	−0.4622	−0.0163	−0.0430	−0.0141	−0.0483	−0.0186	−0.0180
		(0.0200)	(0.0026)	(0.0140)	(0.0025)	(0.0153)	(0.0030)	(0.0029)
	400	−0.4619	−0.0078	−0.0209	−0.0069	−0.0218	−0.0088	−0.0085
		(0.0123)	(0.0008)	(0.0057)	(0.0008)	(0.0054)	(0.0009)	(0.0009)
	800	−0.4615	−0.0038	−0.0104	−0.0034	−0.0098	−0.0043	−0.0042
		(0.0079)	(0.0003)	(0.0023)	(0.0003)	(0.0018)	(0.0003)	(0.0003)
20%	200	−0.4687	−0.0177	−0.0342	−0.0156	−0.0371	−0.0190	−0.0187
		(0.0324)	(0.1107)	(0.1217)	(0.1111)	(0.1239)	(0.1096)	(0.1099)
	400	−0.4649	−0.0102	−0.0175	−0.0094	−0.0181	−0.0107	−0.0106
		(0.0182)	(0.0778)	(0.0888)	(0.0777)	(0.0914)	(0.0772)	(0.0774)
	800	−0.4636	−0.0064	−0.0104	−0.0059	−0.0105	−0.0066	−0.0066
		(0.0113)	(0.0539)	(0.0616)	(0.0538)	(0.0636)	(0.0534)	(0.0536)
40%	200	−0.4851	−0.0223	−0.0349	−0.0201	−0.0372	−0.0230	−0.0228
		(0.0564)	(0.2282)	(0.2400)	(0.2369)	(0.2398)	(0.2243)	(0.2254)
	400	−0.4733	−0.0117	−0.0182	−0.0109	−0.0189	−0.0120	−0.0119
		(0.0358)	(0.1573)	(0.1691)	(0.1629)	(0.1700)	(0.1549)	(0.1556)
	800	−0.4686	0.0013	−0.0025	0.0019	−0.0031	0.0010	0.0011
		(0.0221)	(0.1135)	(0.1210)	(0.1173)	(0.1216)	(0.1118)	(0.1122)

Table 5: Bias and standard deviations (in parentheses) of $\hat{\xi}_n(X,T)$ for DGP3 (c)

			The proposed CRC $\hat{\xi}_n(X,T)$ with various $w(t)$					
CR	n	$\mathrm{dCor}_{\mathrm{IPCW}}$	$\mathrm{CRC}_{N(0,1)}$	$\mathrm{CRC}_{N(-1,5^2)}$	$\mathrm{CRC}_{U(-1,1)}$	$\mathrm{CRC}_{U(-5,5)}$	$\mathrm{CRC}_{t(2)}$	$\mathrm{CRC}_{t(3)}$
0%	200	−0.9294	−0.0738	−0.0846	−0.0675	−0.0845	−0.0740	−0.0741
		(0.0452)	(0.0093)	(0.0098)	(0.0099)	(0.0097)	(0.0093)	(0.0093)
	400	−0.9203	−0.0368	−0.0426	−0.0335	−0.0425	−0.0369	−0.0369
		(0.0192)	(0.0032)	(0.0035)	(0.0034)	(0.0035)	(0.0032)	(0.0032)
	800	−0.9183	−0.0184	−0.0215	−0.0168	−0.0215	−0.0185	−0.0185
		(0.0099)	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0011)	(0.0011)
20%	200	−0.9374	−0.0883	−0.0964	−0.0835	−0.0960	−0.0877	−0.0879
		(0.0566)	(0.1322)	(0.1545)	(0.1240)	(0.1528)	(0.1328)	(0.1327)
	400	−0.9207	−0.0432	−0.0465	−0.0406	−0.0462	−0.0428	−0.0429
		(0.0227)	(0.0864)	(0.1015)	(0.0809)	(0.1005)	(0.0868)	(0.0868)
	800	−0.9187	−0.0231	−0.0252	−0.0215	−0.0250	−0.0229	−0.0230
		(0.0121)	(0.0628)	(0.0739)	(0.0588)	(0.0731)	(0.0632)	(0.0632)
40%	200	−0.9536	−0.1063	−0.1089	−0.1061	−0.1078	−0.1062	−0.1063
		(0.0753)	(0.2951)	(0.3189)	(0.2918)	(0.3131)	(0.2945)	(0.2948)
	400	−0.9259	−0.0625	−0.0653	−0.0622	−0.0646	−0.0624	−0.0625
		(0.0318)	(0.2098)	(0.2282)	(0.2062)	(0.2241)	(0.2094)	(0.2096)
	800	−0.9185	−0.0257	−0.0275	−0.0254	−0.0272	−0.0256	−0.0257
		(0.0162)	(0.1410)	(0.1540)	(0.1381)	(0.1513)	(0.1407)	(0.1408)

Table 6: Bias and standard deviations (in parentheses) of $\hat{\xi}_n(X,T)$ for DGP3 (d)

			The proposed CRC $\hat{\xi}_n(X,T)$ with various $w(t)$					
CR	n	$\text{dCor}_{\text{IPCW}}$	$\text{CRC}_{N(0,1)}$	$\text{CRC}_{N(-1,5^2)}$	$\text{CRC}_{U(-1,1)}$	$\text{CRC}_{U(-5,5)}$	$\text{CRC}_{t(2)}$	$\text{CRC}_{t(3)}$
0%	200	−0.1606	−0.0119	−0.0128	−0.0112	−0.0129	−0.0119	−0.0119
		(0.0172)	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0016)	(0.0016)
	400	−0.1609	−0.0059	−0.0064	−0.0056	−0.0064	−0.0059	−0.0059
		(0.0121)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
	800	−0.1611	−0.0029	−0.0032	−0.0028	−0.0032	−0.0029	−0.0029
		(0.0087)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
20%	200	−0.1618	−0.0113	−0.0129	−0.0105	−0.0132	−0.0114	−0.0114
		(0.0207)	(0.0908)	(0.0954)	(0.0881)	(0.0963)	(0.0907)	(0.0907)
	400	−0.1611	−0.0108	−0.0123	−0.0099	−0.0126	−0.0109	−0.0109
		(0.0144)	(0.0635)	(0.0665)	(0.0617)	(0.0671)	(0.0634)	(0.0634)
	800	−0.1614	−0.0009	−0.0013	−0.0008	−0.0013	−0.0009	−0.0009
		(0.0106)	(0.0445)	(0.0466)	(0.0431)	(0.0471)	(0.0445)	(0.0445)
40%	200	−0.1661	−0.0228	−0.0256	−0.0226	−0.0258	−0.0228	−0.0229
		(0.0264)	(0.1769)	(0.1896)	(0.1757)	(0.1919)	(0.1760)	(0.1763)
	400	−0.1630	−0.0153	−0.0176	−0.0145	−0.0179	−0.0153	−0.0153
		(0.0179)	(0.1178)	(0.1273)	(0.1171)	(0.1290)	(0.1173)	(0.1175)
	800	−0.1620	−0.0065	−0.0076	−0.0066	−0.0077	−0.0066	−0.0066
		(0.0127)	(0.0855)	(0.0910)	(0.0853)	(0.0920)	(0.0850)	(0.0851)

Table 7: Run times (in seconds) for all tests with $B = 500$ (and $M = 500$ for KLR) in DGP4 with $\rho = 0.4$, under censoring rates (CR) 10% and 30%, on a machine equipped with an AMD EPYC 7H12 64-Core processor and 768GB of memory

					The proposed CRC test with various $w(t)$					
CR	n	CPH	KLR	IPCW	$\text{CRC}_{N(0,1)}$	$\text{CRC}_{N(-1,5^2)}$	$\text{CRC}_{U(-1,1)}$	$\text{CRC}_{U(-5,5)}$	$\text{CRC}_{t(2)}$	$\text{CRC}_{t(3)}$
10%	100	0.040	0.023	0.154	0.025	0.024	0.024	0.024	0.024	0.024
	500	0.085	0.926	4.737	0.050	0.046	0.046	0.046	0.046	0.046
	1000	0.141	5.765	21.974	0.089	0.087	0.084	0.082	0.082	0.081
	2000	0.206	33.915	89.244	0.174	0.174	0.177	0.175	0.175	0.178
	5000	0.585	263.928	585.934	0.739	0.732	0.728	0.715	0.715	0.714
	10000	1.190	>25mins	>35mins	1.956	2.023	1.805	1.911	1.828	1.883
30%	100	0.036	0.023	0.107	0.025	0.024	0.025	0.024	0.024	0.024
	500	0.074	0.947	2.512	0.051	0.048	0.048	0.048	0.048	0.048
	1000	0.120	5.728	12.231	0.089	0.087	0.084	0.085	0.083	0.081
	2000	0.169	33.062	50.967	0.147	0.150	0.150	0.147	0.152	0.155
	5000	0.477	263.847	363.507	0.745	0.732	0.716	0.739	0.708	0.700
	10000	0.970	>25mins	>20mins	1.924	2.011	1.886	1.839	1.882	1.872