Title

Missing Information Principle: A Unified Approach for General Truncated and Censored Survival Data Problems

Journal

Statistical Science, 33(2)

ISSN

0883-4237

Authors

Sun, Yifei
Qin, Jing
Huang, Chiung-Yu

Publication Date

2018-05-01

DOI

10.1214/17-sts638

Peer reviewed

HHS Public Access

Author manuscript

Stat Sci. Author manuscript; available in PMC 2019 July 19.

Published in final edited form as:

Stat Sci. 2018 May ; 33(2): 261–276. doi:10.1214/17-STS638.

Missing Information Principle: A Unified Approach for General Truncated and Censored Survival Data Problems

Yifei Sun [Assistant Professor],

Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY 10032

Jing Qin [Mathematical Statistician], and

Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Chiung-Yu Huang [Professor]

Department of Epidemiology and Biostatistics, School of Medicine, University of California San Francisco, San Francisco, CA 94158

Columbia University and National Institutes of Health and University of California, San Francisco

Abstract

It is well known that truncated survival data are subject to sampling bias, where the sampling weight depends on the underlying truncation time distribution. Recently, there has been a rising interest in developing methods to better exploit the information about the truncation time, thus the sampling weight function, to obtain more efficient estimation. In this paper, we propose to treat truncation and censoring as "missing data mechanism" and apply the missing information principle to develop a unified framework for analyzing left-truncated and right-censored data with unspecified or known truncation time distributions. Our framework is structured in a way that is easy to understand and enjoys a great flexibility for handling different types of models. Moreover, a new test for checking the independence between the underlying truncation time and survival time is derived along the same line. The proposed hypothesis testing procedure utilizes all observed data and hence can yield a much higher power than the conditional Kendall's tau test that only involves comparable pairs of observations under truncation. Simulation studies with practical sample sizes are conducted to compare the performance of the proposed method with its competitors. The proposed methodologies are applied to a dementia study and a nursing house study for illustration.

Keywords

Kendall's tau; Inverse probability weighted estimator; Outcome-dependent sampling; Prevalent sampling; Self-consistency algorithm

1. INTRODUCTION

The prevalent cohort design is frequently used to study the natural history of disease processes. A prevalent cohort consists of individuals with disease at the time of enrollment and is followed for the occurrence of failure events of interest. Compared to the incident

cohort approach, which follows initially undiseased individuals from disease onset to failure, the prevalent cohort approach enjoys the advantage of being more efficient and relatively easy to assemble through existing disease registries. However, this design is known to be subject to sampling bias, because diseased individuals who died before the recruitment period would not be qualified to enter the cohort. As a result, the sampling scheme favors individuals who survive longer and thus is outcome dependent. Statistically speaking, the survival time in a prevalent cohort study is subject to left truncation, where the truncation time is the duration from disease onset to enrollment. A survival time can be observed if and only if it is longer than the truncation time. In the case of a stable disease, the truncation times in the unbiased disease population are uniformly distributed under the stationarity assumption with respect to the disease incidence; moreover, the survival time in the prevalent cohort has a length-biased distribution where the probability of a survival time being sampled is proportional to its length (Lancaster, 1990, Chapter 3).

Statistical analysis of truncated survival time data are usually based on non-parametric and semiparametric conditional likelihood methods, conditioning on the observed truncation time (Lynden-Bell, 1971; Wang, 1991; Tsai, Jewell and Wang, 1987). As a result, the inference procedures do not require information about the underlying truncation time distribution. When such information is available, however, the conditional likelihood approaches are known to be inefficient (Wang, 1991) – this is in contrast with the analysis of right-censored survival data where the knowledge about the independent censoring time distribution is ancillary. For survival data collected in the prevalent cohort study of a stable disease, various authors, including Vardi (1989), Vardi and Zhang (1992), Asgharian, M'Lan and Wolfson (2002), Luo and Tsai (2009), Tsai (2009), Qin et al. (2011), Huang and Qin (2012), and Ning, Qin and Shen (2014), have developed more efficient methods that exploit the properties of uniformly distributed truncation times in the estimation procedure. Readers are referred to Shen, Ning and Qin (2017) for a comprehensive review of recent developments.

In this paper, we present a unified framework for analyzing left-truncated and right-censored data with unspecified or known (but not necessary uniform) truncation time distributions. The proposed framework is structured in a way that is easy to understand and enjoys a great flexibility for handling different types of models. Our idea is to treat truncation and censoring as "missing data mechanism" and apply the missing information principle to develop efficient estimation and hypothesis testing procedures. The missing information principle provides a general paradigm for statistical inference in missing data problems. Its theoretical foundation was formally established by Orchard and Woodbury (1972), whose idea dates back to Yates (1933) and Bartlett (1937). Later, Dempster, Laird and Rubin (1977) provided an extensive generalization and named the procedure EM algorithm. Heuristically, one may replace a complete-data estimating function or an unbiased estimator by their conditional expectations given the observed data to obtain unbiased inference. When applied to the score function, the missing in formation principle reduces to a single iteration of the EM-algorithm (Dempster, Laird and Rubin, 1977). It is worthwhile to point out that, under truncation, the number of individuals being truncated is unknown and thus the sample size needs to be imputed via the missing information principle, adding additional level of complication compared to the usual missing data problems.

We begin by deriving score operators from the nonparametric and semiparametric full likelihood function based on completely observed survival data from a representative sample. The score functions are unbiased if there were no missing data, that is, the survival times are neither truncated nor censored. We then apply the missing information principle to the unbiased estimating function and, based on which, derive iterative self-consistency algorithms to obtain maximum likelihood estimation. Compared to existing likelihood-based methods, a major advantage of our approach is that the proposed algorithm is formulated based on the hazard function, making the extension from nonparametric estimation to semiparametric estimation of the Cox model relatively straightforward. Another important feature of our methodology is that, similar to Vardi (1989) and Qin et al. (2011) for survival data under length-biased sampling, the estimated hazard can have positive support on both censored and uncensored failure time points; this is in contrast with the pseudo partial likelihood-based approaches considered by Luo and Tsai (2009) and Tsai (2009) which only allow jumps at uncensored failure times.

We further demonstrate the use of the missing information principle in hypothesis testing, which receives less attention in the truncation data analysis literature compared to model estimation. Specifically, we consider testing the association between the survival time and the truncation time in the target population. Note that, instead of employing the conditional Kendall's tau statistic based on comparable pairs of survival and truncation times in the prevalent cohort (Tsai, 1990), we evaluate the expected difference between the proportions of concordance and discordance pairs, that is, the unconditional Kendall's tau statistics, in the unbiased population by applying the missing information principle. Extensive simulation studies shows that the new testing procedure outperforms the conditional Kendall's tau test, especially in the case where the proportion of comparable pairs are small.

The rest of the article is organized as follows. We demonstrate the application of the missing information principle with left-truncated and/or right-censored data in the case of one-sample estimation (Section 2) and seimiparametric estimation of the Cox model (Section 3). In both cases, we consider the estimation procedure with and without the knowledge of the truncation time distribution. In particular, self-consistency algorithms which guarantee to yield positive hazard function estimates are proposed to incorporate the information about the truncation time distribution. In Section 4, a nonparametric association test is proposed to illustrate the missing information principle when truncation time distribution is not specified. In Section 5, simulation studies are conducted to evaluate the performance of the proposed algorithms. In Section 6, two data examples are presented to illustrate the proposed approaches. A discussion follows in Section 7 to conclude the paper.

2. NONPARAMETRIC ESTIMATION

2.1 Nonparametric estimation with complete data

Let T^0 denote the survival time in the population of interest. Note that we use superscript 0 for random variables in the target population. Assume that T^0 is absolutely continuous and has a support on $[0, \tau]$, that is, T^0 has a probability density function $f(t)$, $0 \le t \le \tau$. Denote, respectively, by $F(t)$, $S(t)$, $\lambda(t)$, and $\Lambda(t)$, the distribution function, survival function, hazard function, and cumulative hazard function of the survival time T^0. Suppose the data

$\{T_1^0, \ldots, T_n^0\}$ from n subjects are independent and identically distributed (i.i.d.) realizations of T^0. Following Murphy and van der Vaart (2000), we consider the nonparametric likelihood

$$\prod_{i=1}^{n} \Lambda\{T_i^0\} \exp\{-\Lambda(T_i^0)\}$$

with $\Lambda\{t\}$ the jump size of Λ at t so that the likelihood depends smoothly on $\Lambda\{T_1^0\}, \ldots, \Lambda\{T_N^0\}$. It is easy to check that the corresponding score operator (Begun et al., 1983) is given by

$$\Psi(k) = \int_0^T k(u)\{d\tilde{N}^0(u) - I(T^0 \geq u) d\Lambda(u)\},$$

where $\tilde{N}^0(t) = I(T^0 \leq t)$ and $\kappa(u)$ is any bounded, measurable function. Setting $\kappa(u) = I(u \leq t)$ motivates the unbiased estimating equation

$$\sum_{i=1}^{n} \int_0^t d\tilde{M}_i^0(u) = 0$$

with $\tilde{M}_i^0(t) = \tilde{N}_i^0(t) - \int_0^t I(T_i^0 \geq u) d\Lambda(u)$. As a result, solving the complete-data estimating equation $\Sigma_{i=1}^n d\tilde{M}_i^0(t) = 0$ for all $t \in [0, \tau]$ is equivalent to maximizing the nonparametric likelihood function with respect to $\Lambda(t)$.

2.2 Left-truncated data, with an unspecified truncation time distribution

We now consider the scenario where the observation of the survival time T^0 is subject to an independent truncation time A^0, that is, the pair of random variables (T^0, A^0) is observed if and only if $T^0 \geq A^0$. We drop superscript 0 to indicate random variables in the prevalent population. Denote by T and A the survival time and truncation time for individuals in the prevalent population, then (T, A) has the same joint distribution as $(T^0, A^0) \mid T^0 \geq A^0$. For simplicity, we assume that the truncation time A^0 also has support on $[0, \tau]$. In what follows, we consider nonparametric estimation with an unspecified truncation time distribution.

Let $\{(T_i, A_i), i = 1, \ldots n\}$ be i.i.d. copies of (T,A). Following Turnbull's argument of ghost observations (Turnbull, 1976), conditioning on the truncation time A_i, the observation (T_i, A_i) can be considered the remnant of a group of m_i unobserved subjects whose survival times are smaller than A_i. Specifically, let $\mathcal{O}_i^* = \{(T_{ij}^*, A_i), j = 1, \ldots, m_i\}$ be the ghosts corresponding to (T_i, A_i), where $T_{ij}^* < A_i$ and T_{ij}^* is independent of T_i given A_i for all $j = 1, \ldots, m_i$. Note that, given $A_i = a$, the sample size m_i of the group of unobserved subjects follows a negative binomial distribution with parameters 1 and $F(a)$ and thus $E(m_i \mid A_i = a) = F(a)/S(a)$. Moreover, given $A_i = a$, the density function of T_{ij}^* is $f(t)F(a)^{-1}I(t < a)$.

For the ith observed subject, we define the stochastic process

$$\widetilde{M}_i(t) = I(T_i \le t) - \int_0^t I(T_i \ge u) d\Lambda(u).$$

Similarly, for truncated observations (the ghosts) $\mathcal{O}_i^* = \{(T_{ij}^*, A_i), j = 1, \ldots, m_i\}$ we define

$$\widetilde{M}_{ij}^*(t) = I(T_{ij}^* \le t) - \int_0^t I(T_{ij}^* \ge u) d\Lambda(u).$$

Then it follows from the unbiasedness of the score operator with complete data that $d\widetilde{M}_i(t) + \Sigma_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t)$ has a zero-mean. Because T_{ij}^*'s are unobserved, we apply the missing information principle to replace $\Sigma_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t)$ with its conditional expectation

$$E\{\sum_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t) | A_i\} = E(m_i | A_i) \times \frac{S(A_i)}{F(A_i)} I(A_i > t) d\Lambda(t) = I(A_i > t) d\Lambda(t)$$

to obtain the imputed stochastic process

$$d\widetilde{M}_i^A(t) \equiv d\widetilde{M}_i(t) + I(A_i > t) d\Lambda(t) = d\widetilde{N}_i(t) - I(T_i \ge t \ge A_i) d\Lambda(t). \quad (2.1)$$

Solving $\Sigma_{i=1}^n d\widetilde{M}_i^A(t) = 0$ gives $d\Lambda(t) = \Sigma_{i=1}^n I(T_i = t) \Big/ \Sigma_{i=1}^n I(T_i \ge t \ge A_i)$. As expected, the application of the missing information principle to left-truncated data yields the asymptotically efficient nonparametric maximum likelihood estimator (NPMLE), that is, the Lynden-Bell estimator (Lynden-Bell, 1971).

2.3 Left-truncated data, with a known truncation time distribution

In many applications, it is reasonable to assume that the incidence of disease onset follows a specific distribution. As an example, several authors, including Addona and Wolfson (2006) and Huang and Qin (2012), have argued that the incidence of dementia onset in the Canadian Study of Health and Aging, one of the largest epidemiology studies of dementia (McDowell, Hill and Lindsay, 2001), follows a Poisson process; that is, the disease incidence is stable over time. Under this stable disease condition, the underlying truncation time is uniformly distributed and the probability of a survival time being sampled is proportional to its length.

Let H be the known distribution function of the underlying truncation time A^0, and define $\bar{H}(t) = 1 - H(t)$. As an example, under the stable disease condition, A^0 is uniformly distributed and hence $H(t) = t/\tau$ and $\bar{H}(t) = 1 - t/\tau$ for $t \in [0, \tau]$. Applying Turnbull's of ghost observations, the observed data (Ti, A_i) can be viewed as the remnant of a group of m_i

independent subjects $\mathcal{O}_i^* = \{(T_{ij}^*, A_{ij}^*), j = 1, \ldots, m_i\}$ whose survival times satisfy $T_{ij}^* < A_{ij}^*$ are thus not observed. Moreover, (T_{ij}^*, A_{ij}^*) has the same joint distribution as $(T^0, A^0) \mid T^0 < A^0$. Note that, instead of using the conditioning argument as in Section 2.1, the ghosts corresponding to the observation (T_i, A_i) are not constrained to have the same truncation time A_i. Define $\alpha = \text{pr}(T^0 < A^0) = \int_0^\tau \bar{H}(u)dF(u)$ The sample size m_i follows a negative binomial distribution with parameters 1 and α, and hence $E(m_i) = \alpha/(1 - \alpha)$.

Following the spirit of missing information principle, we propose to replace $\Sigma_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t)$ in the unbiased estimating function $d\widetilde{M}_i(t) + \Sigma_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t)$ with its expectation integrating over the given truncation time density function. Specifically, it follows from the result that T_{ij}^* has the density function $\bar{H}(t)f(t) \Big/ \alpha$ that

$$E\{\sum_{j=1}^{m_i} d\widetilde{M}_{ij}^*(t)\} = \left\{\bar{H}(t)dF(t) - d\,\Lambda\,(t)\int_t^\tau \bar{H}(u)dF(u)\right\}/(1-\alpha)$$

and that the imputed stochastic process is

$$d\widetilde{M}_i^H(t) \equiv \left\{d\widetilde{N}_i(t) + \frac{\bar{H}(t)dF(t)}{\int_0^\tau H(u)dF(u)}\right\} - \left\{I(T_i \geq t) + \frac{\int_t^\tau \bar{H}(u)dF(u)}{\int_0^\tau H(u)dF(u)}\right\} d\,\Lambda\,(t)\,. \quad (2.2)$$

Solving $\Sigma_{i=1}^n d\widetilde{M}_i^H(t) = 0$ would yield the nonparametric maximum likelihood estimation (NPMLE) of $\Lambda(\cdot)$ when the underlying truncation time distribution is known. It is easy to see that the estimating equation does not have a closed-form solution. We propose a self-consistency algorithm for deriving the NPMLE.

Define the stochastic processes $d\widetilde{\eta}_i(t) = d\widetilde{N}_i(t) + \bar{H}(t)dF(t) \Big/ \int_0^\tau H(u)dF(u)$ and $\widetilde{\xi}_i(t) = I(T_i \geq t) + \int_t^\tau \bar{H}(u)dF(u) \Big/ \int_0^\tau H(u)dF(u)$, so that $d\widetilde{M}_i^H(t) = d\widetilde{\eta}_i(t) + \widetilde{\xi}_i(t)d\,\Lambda\,(t)$. We consider the class of distributions with jumps at the observed failure times. The self-consistency algorithm is described below:

Step 0. Set initial values for the jumps of $\Lambda^{(0)}(t)$ at observed failure times and obtain $S^{(0)}(t) = \exp\{-\Lambda^{(0)}(t)\}$.

Step k. For the k-th iteration, evaluate $d\widetilde{\eta}_i^{(k)}(t)$ and $\widetilde{\xi}_i^{(k)}(t)$ by replacing $F(t)$ with $1 - S^{(k-1)}(t) = 1 - \exp\{-\Lambda^{(k-1)}(t)\}$ in $d\widetilde{\eta}_i(t)$ and $\widetilde{\xi}_i(t)$. Update $\Lambda(t)$ with

$$\Lambda^{(k)}(t) = \int_0^t \frac{\sum_{i=1}^n d\widetilde{\eta}_i^{(k)}(u)}{\sum_{i=1}^n \widetilde{\xi}_i^{(k)}(u)}.$$

Iterate until a convergence criterion is met.

Interestingly, when the distribution of the truncation time A^0 is known, the construction of imputed stochastic process $d\widetilde{M}_i^H(t)$ does not require A being observed. A closer examination reveals that $d\widetilde{M}_i^H(t)$ can be re-expressed as

$$d\widetilde{M}_i^H(t) = d\widetilde{M}_i^A(t) + \left\{ I(A_i > t) - \int_t^\tau S(u)dH(u) / \int_0^\tau S(u)dH(u) \right\} d\Lambda(t),$$

where $d\widetilde{M}_i^A(t)$ is given by (2.1) for left-truncated data with a unspecified truncation time distribution, and the function in the brackets is simply the empirical estimate of the survival function subtract the conditional survival function of A^0 given $T^0 \geq A^0$.

Recognizing that Ti's can be viewed a biased sample from $F(t)$ with a sampling weight function $H(t)$, an inverse-probability weighted estimator for $\Lambda(t)$ (Wang, 1996) can be given by

$$\int_0^t \frac{\sum_{i=1}^n d\tilde{N}_i(u)}{\sum_{j=1}^n I(T_j \geq u)H(u)/H(T_j)}.$$

The assigned weight is inversely proportional to the probability of a subject being sampled. As a result, the weighted risk set has the same probability structure as that that would be formed by an incidence cohort. In most cases, this simple estimator, though consistent for $\Lambda(t)$, is not identical to the NPMLE obtained by solving $\Sigma_{i=1}^n d\widetilde{M}_i^H(t) = 0$ and hence is not expected to be fully efficient.

2.4 Left-truncated and right-censored data, with an unspecified truncation time distribution

The observation of left-truncated survival time is usually subject to right censoring due to loss to follow-up or end of study. Let C be the censoring time for the residual survival time $V = T - A$, where C is assumed to be independent of V given A. Hence we observe $\{(A_i, Y_i, \Delta_i), i = 1, \ldots, n\}$, where $Y_i = \min(T_i, A_i + C_i)$ and $\Delta_i = I(T_i \leq A_i + C_i)$. For censored individuals, the values of $d\tilde{N}_i(t)$ and $I(T_i \geq t)$ in $d\widetilde{M}_i^A(t)$ given by (2.1) can not be determined completely. It can be verified that $E\{d\tilde{N}_i(t) \mid A_i, Y_i, \Delta_i\} = dN_i(t) + (1 - \Delta_i)I(Y_i < t)dF(t) \Big/ S(Y_i)$ and $E\{I(T_i \geq t) \mid A_i, Y_i, \Delta_i\} = I(Y_i \geq t) + (1 - \Delta_i)I(Y_i < t)S(t)/S(Y_i)$, with $N_i(t) = \Delta_i I(Y_i \leq t)$. We apply the missing information principle to $d\widetilde{M}_i^A(t)$ by replacing $d\tilde{N}_i(t)$ and $I(T_i \geq t)$ with their conditional expectations to yield the imputed stochastic process

$$dM_i^A(t) \equiv dN_i(t) - I(Y_i \geq t \geq A_i)d\Lambda(t). \quad (2.3)$$

It is easy to see that solving $\Sigma_{i=1}^{n} dM_i^A(t) = 0$ yields the Nelson-Aalen estimator, that is, the NPMLE, for left-truncated and right-censored data with unspecified truncation time distribution.

2.5 Left-truncated and right-censored data, with a known truncation time distribution

When the truncation time A^0 follows a known distribution function H, applying the missing information principle to replace $d\widetilde{N}_i(t)$ and $I(T_i \geq t)$ in $d\widetilde{M}_i^H(t)$ in (2.2) with their conditional expectations yields

$$\begin{aligned} dM_i^H(t) \equiv & \left[E\{d\widetilde{N}_i(t)|Y_i,\ \Delta_i\} + \bar{H}(t)dF(t)/\int_0^\tau H(u)dF(u)\right] \\ & - \left[E\{I(T_i \geq t)|Y_i,\ \Delta_i\} + \int_t^\tau \bar{H}(u)dF(u)/\int_0^\tau H(u)dF(u)\right]d\Lambda(t) \\ & = d\eta_i(t) - \xi_i(t)d\Lambda(t), \end{aligned} \tag{2.4}$$

where $d\eta_i(t) = dN_i(t) + (1 - \Delta_i)I(Y_i < t)S(t) \Big/ S(Y_i)d\Lambda(t) + \bar{H}(t)dF(t) \Big/ \int_0^\tau H(u)dF(u)$ and $\xi_i(t) = I(Y_i \geq t) + (1 - \Delta_i)I(Y_i < t)S(t) \Big/ S(Y_i) + \int_t^\tau \bar{H}(u)dF(u) \Big/ \int_0^\tau H(u)dF(u)$. Similar to (2.2), the evaluation of the imputed process (2.4) does not require A_i being observed.

At any time point t^* that no failure event was observed, that is, $\Sigma_{i=1}^{n} dN_i(t^*) = 0$, the equality $\Sigma_{i=1}^{n} dM_i^H(t^*) = 0$ can be implied by either $d\Lambda(t^*) = 0$ or $n^{-1}\Sigma_{i=1}^{n} I(Y_i \geq t^*) = \int_{t^*}^\tau S(u)dH(u) \Big/ \int_0^\tau S(u)dH(u)$. In other words, the NPMLE may have jumps at censored survival times. This is in contrast to right-censored survival data, where the NPMLE has jumps at only uncensored failure times.

Similar to Section 2.3, a self-consistency algorithm can be derived to solve $\Sigma_{i=1}^{n} dM_i^H(t) = 0$. Here in all iterative steps the cumulative hazard function $\Lambda^{(k)}(t)$ for $k \geq 0$ are allowed to have jumps at all censored and uncensored failure times. Specifically, in the kth iteration, we evaluate

$$\Lambda^{(k)}(t) = \int_0^t \frac{\sum_{i=1}^n d\eta_i^{(k)}(u)}{\sum_{i=1}^n \xi_i^{(k)}(u)}, \tag{2.5}$$

where $d\eta_i^{(k)}(t)$ and $\xi_i^{(k)}(t)$ are obtained by substituting $F(t)$ with $1 - S^{(k-1)}(t)$ in $d\eta_i(t)$ and $\xi_i(t)$.

Remark 2.1. It is worthwhile to point out that the proposed method can be applied to analyze right-censored survival data under biased sampling, where the sampling weight is proportional to a known function $H(t)$. Luo and Tsai (2009) considered this setting and proposed a pseudo-partial likelihood approach that allows for jumps only at uncensored

failure times. Their estimation procedure, however, requires estimation of the censoring time distribution. In contrast, the proposed estimator is derived directly from the full likelihood of complete data and thus is expect to be more efficient.

3. SEMIPARAMETRIC ESTIMATION OF THE COX MODEL

In this section, we apply the missing information principle, along the same line as nonparametric estimation, to estimate the Cox proportional hazards model with left-truncated and right-censored data. To proceed, we assume that, given the p-dimensional covariate vector $Z^0 = z$, the conditional hazard function of the survival time T^0 in the target population, $\Lambda(t \mid z)$, follows the proportional hazards model, that is, $\lambda(t \mid z) = \lambda(t) \exp(\beta' z)$, where $\lambda(t)$ is an unspecified baseline hazard function and β is a vector of $p \times 1$ regression parameters.

We begin by deriving the unbiased estimating equations based on complete data. Let the complete data $\{(T_i^0, Z_i^0), i = 1, \ldots, n\}$ be i.i.d. copies of (T^0, Z^0). Define the stochastic process $\widetilde{M}^0(t, \beta) = \widetilde{N}^0(t) - \int_0^t \exp(\beta' Z^0) I(T^0 \geq u) \lambda(u) du$, with $\Lambda(t) = \int_0^t \lambda(u) du$. Denote by $\Lambda\{t\}$ the jump size of Λ at t. The score operators for β and Λ derived from the semiparametric likelihood

$$\prod_{i=1}^{n} \Lambda\{T_i^0\} \exp(\beta' Z_i^0) \exp\{-\Lambda(T_i^0) \exp(\beta' Z_i^0)\}$$

are given by $\Psi_\beta = \int_0^\tau Z^0 d\widetilde{M}^0(u, \beta)$ and $\Psi_\Lambda(\kappa) = \int_0^\tau \kappa(u) d\widetilde{M}^0(u, \beta)$, with κ an arbitrary bounded, measurable function. Setting $\kappa(u) = I(u \leq t)$ and solving the system of estimating equations $\Sigma_{i=1}^n \int_0^\tau Z_i d\widetilde{M}_i^0(u, \beta) = 0$ and $\Sigma_{i=1}^n \int_0^t d\widetilde{M}_i^0(u, \beta) = 0$, for all $t \in [0, \tau]$, yields the semiparametric maximum likelihood estimator.

3.1 Semiparametric estimation with left-truncated data

Under left truncation, we observe (T^0, A^0, Z^0) if and only if $T^0 \geq A^0$, so the observed triplet (T, A, Z) has the same joint distribution as $(T^0, A^0, Z^0) \mid T^0 \geq A^0$. Let $\{(A_i, T_i, Z_i), i = 1, \ldots, n\}$ be i.i.d. copies of (A, T, Z). We impose the usual independent truncation assumption by assuming that A^0 is independent with T^0 conditional on Z^0.

We first consider the case where the distribution of A^0 is left unspecified. Arguing as in Section 2.2, the observation (T_i, A_i, Z_i) corresponds to m_i unobserved ghosts $\{(T_{ij}^*, A_i, Z_i), j = 1, \ldots, m_i\}$, where $T_{ij}^* < A_i$ and T_{ij}^* is independent of T_i given (A_i, Z_i). Conditioning on $A_i = a$ and $Z_i = z$, the sample size m_i follows a negative binomial distribution with parameters 1 and $F(a \mid z) = 1 - \exp\{-\Lambda(a \mid z)\}$, where $\Lambda(a \mid z) = \int_0^a \lambda(u \mid z) du$. Hence we have $E(m_i \mid A_i, Z_i) = F(A_i \mid Z_i)/S(A_i \mid Z_i)$.

Define the stochastic processes $\widetilde{M}_i(u,\beta) = I(T_i \le t) - \int_0^t \exp(\beta' Z_i) I(T_i \ge u) d\,\Lambda\,(u)$ for observed data and $\widetilde{M}^*_{ij}(t,\beta) = I(T^*_{ij} \le t) - \int_0^t \exp(\beta' Z_i) I(T^*_{ij} \ge u) d\,\Lambda\,(u)$ for ghost observations. Following the unbiasedness of the score operators with complete data, we have $E[\int_0^\tau Z_i \{d\widetilde{M}_i(u,\beta) + \Sigma_{j=1}^{m_i} d\widetilde{M}^*_{ij}(u,\beta)\}] = 0$ and $E\int_0^t \{d\widetilde{M}_i(u,\beta) + \Sigma_{j=1}^{m_i} d\widetilde{M}^*_{ij}(u,\beta)\} = 0$. In the spirit of missing information principle, we replace $\Sigma_{j=1}^{m_i} d\widetilde{M}^*_{ij}(t,\beta)$ with its conditional expectation to obtain the imputed stochastic process

$$d\widetilde{M}^A_i(t,\beta) \equiv d\widetilde{M}_i(u,\beta) + E\{\sum_{j=1}^{m_i} d\widetilde{M}^*_{ij}(t,\beta) | A_i, Z_i\}$$

$$= d\widetilde{N}_i(t) - \exp(\beta' Z_i) I(T_i \ge t \ge A_i) d\,\Lambda\,(t)\,.$$

As expected, solving the imputed estimating equations $\Sigma_{i=1}^n \int_0^\tau Z_i d\widetilde{M}^A_i(t,\beta) = 0$ and $\Sigma_{i=1}^n \int_0^t d\widetilde{M}^A_i(u,\beta) = 0$ for all $t \in [0,\ \tau]$ yields the maximum partial likelihood estimator (Wang, Brookmeyer and Jewell, 1993) that is the solution of the partial score equation

$$\sum_{i=1}^n \int_0^\tau \left\{ Z_i - \frac{\sum_{j=1}^n Z_j \exp(\beta' Z_j) I(T_j \ge t \ge A_j)}{\sum_{j=1}^n \exp(\beta' Z_j) I(T_j \ge t \ge A_j)} \right\} d\widetilde{N}_i(t) = 0.$$

Next, we consider the case where A^0 has a known distribution function $H(t)$. Integrating over the given truncation time density function, straightforward algebra gives

$$E\{\sum_{j=1}^{m_i} d\widetilde{M}^*_{ij}(t) | Z_i\} = \frac{\overline{H}(t) dF(t|Z_i)}{\int_0^\tau H(u) dF(u|Z_i)} - \frac{\int_t^\tau \overline{H}(u) dF(u|Z_i)}{\int_0^\tau H(u) dF(u|Z_i)} d\,\Lambda\,(t)\,.$$

Thus, by replacing $\Sigma_{j=1}^{m_i} d\widetilde{M}^*_{ij}(t)$ with its expectation in the unbiased stochastic process, we obtain the imputed stochastic process

$$d\widetilde{M}^H_i(t,\beta) \equiv d\widetilde{M}_i(t,\beta) + E\{\sum_{j=1}^{m_i} d\widetilde{M}^*_{ij}(t,\beta) | Z_i\}$$

$$= \left\{ d\widetilde{N}_i(t) + \frac{\overline{H}(t) dF(t|Z_i)}{\int_0^\tau H(u) dF(u|Z_i)} \right\} - \left\{ I(T_i \ge t) + \frac{\int_t^\tau \overline{H}(u) dF(u|Z_i)}{\int_0^\tau H(u) dF(u|Z_i)} \right\} d\,\Lambda\,(t)\,.$$

Solving $\Sigma_{i=1}^n \int_0^\tau Z_i d\widetilde{M}^H_i(t,\beta) = 0$ and $\Sigma_{i=1}^n \int_0^t d\widetilde{M}^H_i(u,\beta) = 0$ for all $t \in [0,\ \tau]$ would yield the semiparametric maximum likelihood estimator when the distribution of A^0 is known.

Define the stochastic processes $d\tilde{\eta}_i(t,\beta) = d\tilde{N}_i(t) + \bar{H}(t)dF(t \mid Z_i) \Big/ \int_0^\tau H(u)dF(u \mid Z_i)$ and $\tilde{\xi}_i(t,\beta) = I(T_i \geq t) + \int_t^\tau \bar{H}(u)dF(u \mid Z_i) \Big/ \int_0^\tau H(u)dF(u \mid Z_i)$. The solution to $\Sigma_{i=1}^n \int_0^t d\tilde{M}_i^H(u,\beta) = 0$ and $\Sigma_{i=1}^n \int_0^\tau Z_i d\tilde{M}_i^H(t,\beta) = 0$ satisfies

$$d\,\Lambda\,(t) = \frac{\sum_{i=1}^n d\tilde{\eta}_i(t,\beta)}{\sum_{i=1}^n \exp(\beta' Z_i)\tilde{\xi}_i(t,\beta)}, \quad (3.1)$$

and

$$\sum_{i=1}^n \int_0^\tau \left\{ Z_i - \frac{\sum_{j=1}^n Z_j \exp(\beta' Z_j)\tilde{\xi}_j(t,\beta)}{\sum_{j=1}^n \exp(\beta' Z_j)\tilde{\xi}_j(t,\beta)} \right\} d\tilde{\eta}_i(t,\beta) = 0. \quad (3.2)$$

Based on (3.1) and (3.2), we propose the following iterative algorithm to obtain estimates of β and $\Lambda(t)$. As before, we consider the family of Λ that only jumps at the unique failure times in the following algorithm.

Step 0. Set initial values for $\beta^{(0)}$ and the jumps of $\Lambda^{(0)}(t)$ at observed failure times. Compute $S^{(0)}(t \mid Z_i) = \exp\{-\Lambda^{(0)}(t)\exp(\beta^{(0)'} Z_i)\}$.

Step k. For the k-th iteration, evaluate $d\tilde{\eta}_i^{(k)}$ and $\tilde{\xi}_i^{(k)}$ by replacing $F(t \mid Z_i)$ with $1 - S^{(k-1)}(t \mid Z_i) = 1 - \exp\{-\Lambda^{(k-1)}(t)\exp(\beta^{(k-1)'})Z_i\}$ in $d\tilde{\eta}_i$ and $\tilde{\xi}_i$. Solve

$$\sum_{i=1}^n \int_0^\tau \left\{ Z_i - \frac{\sum_{j=1}^n Z_j \exp(\beta' Z_j)\tilde{\xi}_j^{(k)}(t,\beta^{(k-1)})}{\sum_{j=1}^n \exp(\beta' Z_j)\tilde{\xi}_j^{(k)}(t,\beta^{(k-1)})} \right\} d\tilde{\eta}_i^{(k)}(t,\beta^{(k-1)}) = 0.$$

for β to obtain $\beta^{(k)}$, and update $\Lambda(t)$ with

$$\Lambda^{(k)}(t) = \int_0^t \frac{\sum_{i=1}^n d\tilde{\eta}_i^{(k)}(u,\beta^{(k-1)})}{\sum_{i=1}^n \exp\{\beta^{(k)'} Z_i\}\tilde{\xi}_i^{(k)}(u,\beta^{(k-1)})}.$$

Iterate until a convergence criterion is met.

3.2. Semiparametric estimation with left-truncated and right-censored data

In the presence of right censoring C in addition to left truncation, the observed data $\{(A_i, Y_i, \Delta_i, Z_i)\}$, $i = 1, \ldots, n\}$ are i.i.d copies of (A, Y, Δ, Z), where $Y = \min(T, A + C)$ and $\Delta = I(A + C \geq T)$. We assume that the censoring time C is independent of (A, T) given Z. As pointed out by one reviewer, one may also assume that C is independent of T given (A, Z). We adopt the former assumption to be consistent with the existing literature.

We now consider the case where the distribution of the underlying truncation time A^0 is not specified. Similarly as before, it can be verified that $E\{d\widetilde{N}_i(t) \mid A_i, Y_i, \Delta_i, Z_i\} = dN_i(t) + (1 - \Delta_i)I(Y_i < t)dF(t \mid Z_i) \Big/ S(Y_i \mid Z_i)$ and $E(I(T_i \geq t) \mid A_i, Y_i, \Delta_i, Z_i\} = I(Y_i \geq t) + (1 - \Delta_i)I(Y_i < t)S(t \mid Z_i)/S(Y_i \mid Z_i)$. Applying the missing information principle to $d\widetilde{M}_i^A(t, \beta)$ by replacing $d\widetilde{N}_i(t)$ and $I(T_i \geq t)$ with their conditional expectations yields

$$dM_i^A(t, \beta) = dN_i(t) - \exp(\beta' Z_i)I(Y_i \geq t \geq A_i)d\Lambda(t).$$

As expected, solving the system of imputed estimating equations $\Sigma_{i=1}^n \int_0^\tau Z_i dM_i^A(t, \beta) = 0$ and $\Sigma_{i=1}^n \int_0^t dM_i^A(u, \beta) = 0$ for all $t \in [0, \tau]$ yields the maximum partial likelihood estimator, which is the solution of the partial score equation

$$\sum_{i=1}^n \int_0^\tau \left\{ Z_i - \frac{\sum_{k=1}^n Z_k \exp(\beta' Z_k)I(Y_k \geq t \geq A_k)}{\sum_{k=1}^n \exp(\beta' Z_k)I(Y_k \geq t \geq A_k)} \right\} dN_i(t) = 0.$$

Finally, we consider the case where the distribution of A^0 is known to be H. Replacing $d\widetilde{N}_i(t)$ and $I(T_i \geq t)$ with their conditional expectations in $d\widetilde{M}_i^H(t, \beta)$ yields

$$dM_i^H(t, \beta) = d\eta_i(t, \beta) - \xi_i(t, \beta)\exp(\beta' Z_i)d\Lambda(t),$$

where

$$d\eta_i(t, \beta) = dN_i(t) + (1 - \Delta_i)I(Y_i < t)\frac{dF(t|Z_i)}{S(Y_i|Z_i)} + \frac{\overline{H}(t)dF(t|Z_i)}{\int_0^\tau H(u)dF(u|Z_i)}$$

$$\xi_i(t, \beta) = I(Y_i \geq t) + (1 - \Delta_i)I(Y_i < t)\frac{S(t|Z_i)}{S(Y_i|Z_i)} + \frac{\int_t^\tau \overline{H}(u)dF(u|Z_i)}{\int_0^\tau H(u)dF(u|Z_i)}.$$

Solving the imputed estimating equations $\Sigma_{i=1}^n \int_0^\tau Z_i dM_i^H(t, \beta) = 0$ and $\Sigma_{i=1}^n \int_0^t dM_i^H(u, \beta) = 0$ for all $t \in [0, \tau]$ gives estimates of β and $\Lambda(t)$. Because a closed solution does not exist, we develop a self-consistency algorithm for model estimation.

Arguing as in Section 2.5, the estimated baseline cumulative hazard function obtained by solving the imputed estimating equations may have jumps at censored survival times. Hence in all iterative steps the baseline cumulative hazard function $\Lambda^{(k)}(t)$ for $k \geq 0$ are allowed to have jumps at all censored and uncensored failure times. At the kth iteration, we compute $d\eta_i^{(k)}(t, \beta^{(k-1)})$ and $\xi_i^{(k)}(t, \beta^{(k-1)})$ by substituting, $\{\beta, F(t \mid z)\}$ in $d\eta_i(t,\beta)$ and $\xi_i(t,\beta)$ by the estimates from the $(k-1)$th iteration, and solve the equations along the same line as those in

Step k of the algorithm in Section 3.1, where $d\tilde{\eta}_i^{(k)}(t, \beta^{(k-1)})$ and $\tilde{\xi}_i^{(k)}(t, \beta^{(k-1)})$ are replaced by $\xi_i^{(k)}(t, \beta^{(k-1)})$ and $d\eta_i^{(k)}(t, \beta^{(k-1)})$. Interestingly, in the special case where H is the distribution function of an uniform random variable, the proposed self-consistency algorithm will converge to the semiparametric MLE described in Qin et al. (2011).

Denote by $(\hat{\beta}, \hat{\Lambda})$ the estimators obtained by the proposed self-consistency algorithm and by (β_0, Λ_0) the true parameter values. The large-sample properties of $(\hat{\beta}, \hat{\Lambda})$ is summarized in Theorem 3.1, with regularity conditions and asymptotic distribution given in the Appendix. The proof closely follows Theorems 1 and 2 in Qin et al. (2011) and thus is omitted in this article.

THEOREM 3.1. *Under the regularity conditions (A1)~ (A6),* $\sqrt{n}\{\hat{\beta} - \beta_0, \hat{\Lambda}(t) - \Lambda_0(t)\}(t \in (0, \tau])$ *converges weakly to a zero-mean Gaussian process defined in the* Appendix *as* $n \to \infty$.

REMARK 3.1. The self-consistency algorithm described in this section can also be applied to right-censored survival data under biased sampling. For this problem, Tsai (2009) proposed a pseudo-partial likelihood approach to incorporate the knowledge about the sampling weght function $H(t)$ in the estimation procedure. This approach, however, involves estimating the random censoring time distribution and can be inefficient when the censroing proportion is high. Our estimator, on the other hand, naturally accounts for covariate-dependent censoring and is in general more efficient as it is derived from the full likelihood of complete data.

When the underlying truncation time distribution depends on the covariates and is left unspecified, application of the MIP results in the conditional likelihood approach. When the underlying truncation time distribution depends on the covariates and is specified, the self-consistency algorithm can be easily extended to gain efficiency. For example, suppose the cumulative distribution function of A^0 conditional on $Z^0 = z$ is $H(\cdot \mid z)$, we can replace $H(\cdot)$ in $\eta_i(t, \beta)$ and $\xi_i(t, \beta)$ with $H(\cdot \mid Z_i)$ to estimate β and Λ. However, this approach is not practically interesting, since $H(\cdot \mid z)$ is usually treated as a nuisance function.

4. NONPARAMETRIC ASSOCIATION TEST FOR INDEPENDENT TRUNCATION

The preceding sections illustrate the application of the missing information principle in model estimation. In this section, we consider nonparametric test for the association between the underlying truncation time and the failure time. Under left truncation, the validity of most statistical methods for left-truncated survival time data requires the assumption of quasi-independence to hold, that is, the failure time and the truncation time are independent in the observable region. In the literature, Kendall's tau (Kendall and Gibbons, 1990) is a popular nonparametric measure of association between two failure time random variables because of its rank-invariance property. To measure the association between the underlying truncation time A^0 and the underlying survival time T^0, Kendall's tau can be defined as $K = E[\text{sgn}\{(A_1^0 - A_2^0)(T_1^0 - T_2^0)\}]$, where $\text{sgn}(u)$ is the sign of u. Clearly, K does not depend on

the marginal distribution; moreover, $-1 \le K \le 1$ and $K = 0$ when A^0 and T^0 are independent. For completely observed data, K can be consistently estimated by

$$\hat{K} = \frac{1}{\binom{n}{2}} \sum_{i<j} \operatorname{sgn}\{(A_i^0 - A_j^0)(T_i^0 - T_j^0)\}.$$

A pair of subjects (i,j) is said to be concordant if $(A_i^0 - A_j^0)(T_i^0 - T_j^0) > 0$, and discordant if $(A_i^0 - A_j^0)(T_i^0 - T_j^0) < 0$. As pointed out by many authors, including Tsai (1990), Martin and Betensky (2005), and Oakes (2008), Kendall's tau is not directly applicable to left truncated data, as the observed data (A, T) are a biased sample of (A^0, T^0); moreover, association in the observed bivariate random variable (A, T) arises naturally due to sampling constraint. Failing to account for sampling bias in the construction of test statistics usually leads to incorrect conclusions.

In the absence of right censoring, Tsai (1990) considered conditional Kendall's tau

$$K_c = E\{\operatorname{sgn}(A_1 - A_2)\operatorname{sgn}(T_1 - T_2) | \max(A_1, A_2) \le \min(T_1, T_2)\},$$

for testing the association between A^0 and T^0 under left truncation. It is easy to see that independence of A^0 and T^0 in the observable region $\{(a,t) : 0 \le a \le t \le \tau\}$ implies $K_c = 0$ (but not vice versa). Estimation of the conditional Kendall's tau is based on comparable pairs $\{(A_i, T_i), (A_j, T_j)\}$ that satisfy $\max(A_i, A_j) \le \min(T_i, T_j)$ (Bhattacharya, Chernoff and Yang, 1983), and thus can be very inefficient when the number of comparable pairs is small. Specifically, with a negative correlation between the underlying truncation time and survival time, $A_i \ge A_j$ implies that T_i is likely to be smaller than T_j. As a result, the comparability condition is likely to be satisfied and the conditional Kendall's tau is likely to utilize most available information. On the other hand, with a positive correlation, fewer pairs are expected to satisfy the comparability condition, as the condition further requires $A_i \le T_j$ when $A_i \ge A_j$. In what follows, we consider alternative tests that can better utilize the observed data.

Instead of employing the conditional Kendall's tau for testing association, we propose to apply the missing information principle to construct new test statistics. Arguing as before, we begin by deriving the test statistic using complete data from enrolled individuals and their corresponding (unobserved) ghosts, that is, $\{(A_i, T_i)(A_i, T_{ip}^*); p = 1, \ldots, n\}$. If the compete data were observed, the contribution of any pair of subjects (i,j) to the construction of Kendall's tau statistic is given by $u_{ij} = u_{ij}^{(0)} + u_{ij}^{(1)} + u_{ji}^{(1)} + u_{ij}^{(2)}$, where

$$u_{ij}^{(0)} = \mathrm{sgn}(A_i - A_j)\mathrm{sgn}(T_i - T_j),$$
$$u_{ij}^{(1)} = \sum_{p=1}^{m_i} \mathrm{sgn}(A_i - A_j)\mathrm{sgn}(T_{ip}^* - T_j),$$
$$u_{ij}^{(2)} = \sum_{p=1}^{m_i}\sum_{q=1}^{m_j} \mathrm{sgn}(A_i - A_j)\mathrm{sgn}(T_{ip}^* - T_{jq}^*).$$

Under the null hypothesis that A^0 and T^0 are independent, it is easy to see that u_{ij} has mean zero. Thus a Kendall's tau type test statistic based on the observed data and "ghost" data is given by

$$K^0 = \frac{\sum_{i=1}^n \sum_{j=i+1}^n u_{ij}}{\sum_{i=1}^n \sum_{j=i+1}^n (m_i+1)(m_j+1)},$$

and the denominator $\sum_{i=1}^n \sum_{j=i+1}^n (m_i+1)(m_j+1)$ is the number of comparable pairs and normalize K^0 to be in $[-1, 1]$.

In the absence of right censoring, we apply the missing information principle and replace the unknown quantities in u_{ij} with their expectations conditioning on the observed data. Under the null hypothesis, following the arguments in Section 2.2, it can be verified that the conditional expectations of $u_{ij}^{(k)}$, $k = 0, 1, 2$, given the observed pair (A_i, T_i) and (A_j, T_j) is

$$v_{ij}^{(0)} = E\{u_{ij}^{(0)} | A_i, T_i, A_j, T_j\} = \mathrm{sgn}(A_i - A_j)\mathrm{sgn}(T_i - T_j) = u_{ij}^{(0)},$$

$$v_{ij}^{(1)} = E\{u_{ij}^{(1)} | A_i, T_i, A_j, T_j\}$$
$$= \mathrm{sgn}(A_i - A_j)\left\{\mathrm{sgn}(A_i - T_j)\frac{F(A_i)}{S(A_i)} - 2I(A_i > T_j)\frac{F(T_j)}{S(A_i)} - I(A_i = T_j)\frac{F(A_i)}{S(A_i)}\right\},$$

$$v_{ij}^{(2)} = E\{u_{ij}^{(2)} | A_i, T_i, A_j, T_j\}$$
$$= \mathrm{sgn}(A_i - A_j)\frac{\int_0^{A_j}\int_0^{A_j} \mathrm{sgn}(u - v)dF(u)dF(v)}{S(A_i)S(A_j)}.$$

Moreover, the quantity $(m_i + 1)(m_j + 1)$ can be imputed as

$$m_{ij} = E\{(m_i+1)(m_j+1) | A_i, A_j\} = 1 + \frac{F(A_i)}{S(A_i)} + \frac{F(A_j)}{S(A_j)} + \frac{F(A_i)F(A_j)}{S(A_i)S(A_j)}.$$

Define $v_{ij} = v_{ij}^{(0)} + v_{ij}^{(1)} + v_{ji}^{(1)} + v_{ij}^{(2)}$, then the imputed test statistic given the observed left truncated data is given by $\tilde{K}^A = \sum_{i=1}^{n} \sum_{j=i+1}^{n} v_{ij} \Big/ \sum_{i=1}^{n} \sum_{j=i+1}^{n} m_{ij}$.

When the observation of left-truncated survival times is further subject to independent right censoring, we apply the missing information principle to obtain the imputed test statistic $K^A = \sum_{i=1}^{n} \sum_{j=i+1}^{n} k_{ij} \Big/ \sum_{i=1}^{n} \sum_{j=i+1}^{n} m_{ij}$, where $k_{ij} = k_{ij}^{(0)} + k_{ij}^{(1)} + k_{ji}^{(1)} + k_{ij}^{(2)}$ with

$$k_{ij}^{(0)} = E\{v_{ij}^{(0)} | A_i, Y_i, \Delta_i, A_j, Y_j, \Delta_J\}$$

$$= \operatorname{sgn}(A_i - A_j)\Big[\Delta_i \Delta_j \operatorname{sgn}(Y_i - Y_j)$$

$$+ \Delta_i (1 - \Delta_j)\left\{ I(Y_i \geq Y_j) \frac{S(Y_j) - S(Y_i)}{S(Y_j)} - \frac{S(Y_i \vee Y_j)}{S(Y_j)} \right\}$$

$$+ (1 - \Delta_i) \Delta_j \left\{ \frac{S(Y_i \vee Y_j)}{S(Y_i)} - I(Y_i \leq Y_j) \frac{S(Y_i) - S(Y_j)}{S(Y_i)} \right\}$$

$$+ (1 - \Delta_i)(1 - \Delta_j) \frac{\int_{Y_i}^{\tau} \int_{Y_j}^{\tau} \operatorname{sgn}(u - v) dF(u) dF(v)}{S(Y_i) S(Y_j)} \Big],$$

$$k_{ij}^{(1)} = E\{v_{ij}^{(1)} | A_i, Y_i, \Delta_i, A_j, Y_j, \Delta_j\}$$

$$= \operatorname{sgn}(A_i - A_j)\Big[\Delta_j \left\{ \operatorname{sgn}(A_i - Y_j) \frac{F(A_i)}{S(A_i)} - 2I(A_i > Y_j) \frac{F(Y_j)}{S(A_i)} \right\}$$

$$+ (1 - \Delta_j)\left\{ I(Y_j \leq A_i) \frac{S(Y_j) - S(A_i)}{S(Y_j)} - \frac{S(A_i \vee Y_j)}{S(Y_j)} \right\} \frac{F(A_i)}{S(A_i)}$$

$$- 2(1 - \Delta_j) I(A_i > Y_j) \frac{\int_{Y_j}^{A_i} F(u) dF(u)}{S(Y_j) S(A_i)} \Big],$$

$$k_{ij}^{(2)} = E\{v_{ij}^{(2)} | A_i, Y_i, \Delta_i, A_j, Y_j, \Delta_j\} = v_{ij}^{(2)}$$

The test statistic involves the unknown functions S and $F = 1 - S$. Intuitively, one may replace the survival function S by the product-limit estimator for left-truncated and right-

censored data. Our simulation shows that the type I error rate of the test is close to the pre-specified nominal level. Denote by $\hat{K}^A$ the test statistic with S replaced by the product limit estimator. Following a standard argument and applying the functional delta method, we can show that $n^{-1/2}\hat{K}^A$ converges to a zero-mean normal distribution under the null hypothesis. The formula of the asymptotic variance is very complicated, hence we recommend using nonparametric bootstrap method to obtain the confidence interval of the test statistic and reject the null hypothesis at a significance level of 0.05 if the 95% confidence interval does not cover 0.

Note that because a small value of $S(A)$ in the denominator can result in a very large $k_{ij}^{(m)}(m = 1, 2)$, in practice, to stabilize the test statistic, we only evaluate Kendall's tau on the region $\{(a, t) : 0 \le a \le \tau_0\}$, where τ_0 is an arbitrary constant smaller than τ. Specifically, define $K_{\tau_0}^A = \Sigma_{i=1}^n \Sigma_{j=i+1}^n k_{ij} I(A_i \le \tau_0, A_j \le \tau_0) \Big/ \Sigma_{i=1}^n \Sigma_{j=i+1}^n m_{ij} I(A_i \le \tau_0, A_j \le \tau_0)$, and we use $\hat{K}_{\tau_0}^A$ with estimated survival function S as the testing statistics.

5. SIMULATION STUDY

Numerical simulations were carried out to evaluate the performance of the nonparametric and semiparametric estimators from the iterative algorithms in Section 2.5 and Section 3.2. We considered the following two scenarios for the underlying truncation time random variable: (I) A^0 follows an exponential distribution with with survival function $\bar{H}(t) = \exp(-t)$; (II) A^0 follows a Weibull distribution with survival function $\bar{H}(t) = \exp(-t^2/4)$. The time from enrollment to loss to follow-up was generated from a uniform distribution so that the censoring rate was approximately 25% and 50%. We generated 1000 datasets, each with a sample size of $n = 100$ and 400.

We first evaluated the nonparametric estimation procedure for left-truncated and right-censored data given in Section 2.5. To simulate left-truncated data, we generated the unbiased survival time T^0 from a Weibull distribution with hazard function $0.5t$ repeatedly until there are n subjects satisfying the sampling constraint $A^0 < T^0$, where A^0 were simulated under Scenarios (I) and (II). Table 1 reports the summary statistics for the proposed nonparametric estimator. We compared the proposed estimator $\hat{S}(t)$ with the product-limit estimator $\hat{S}_{\mathrm{PL}}(t)$ for left-truncated and right-censored data, and the nonparametric pseudo partial likelihood estimator $\hat{S}_{\mathrm{PPL}}(t)$ proposed by Luo and Tsai (2009). It can be seen that both $\hat{S}$ and $\hat{S}_{PPL}$ have smaller mean square error than $\hat{S}_{PL}$ in all the scenarios, thus improvement is gained by using information from the underlying truncation time distribution. In Scenario I, the proposed estimator $\hat{S}$ has similar performance as $\hat{S}_{PPL}$; in Scenario II, $\hat{S}$ has smaller variance and larger bias compared to $\hat{S}_{PPL}$ and $\hat{S}_{PL}$, and $\hat{S}$ performs best in terms of mean square error. For all the three estimators, the bias decreases as sample size increase.

In the second set of simulation studies, we evaluated the semiparametric estimation procedure presented in Section 3.2. We generated the unbiased failure time T^0 from the proportional hazards model with two covariates, where the continuous covariate Z_1 follows a uniform distribution on [0, 1], the binary covariate Z_2 follows Bernoulli distribution with success probability 0.5. The coefficients are set to be $\beta = (1, 1)$ and the baseline hazard function is set to be $2t$. The variance estimation follows a perturbation procedure in Qin et al. (2011). We compared the proposed estimator $\hat{\beta}$ in Section 3.2 with the pseudo partial likelihood estimator $\hat{\beta}_{\mathrm{PPL}}$ proposed in Tsai (2009) and the partial likelihood estimator $\hat{\beta}_{\mathrm{PL}}$ for left-truncated and right-censored data. Table 2 reports the summary statistics for the three estimators. It can be observed that our proposed method has negligible bias, and, as expected, the bias decreases with sample size. The variance of $\hat{\beta}$ is smaller than that of both $\hat{\beta}_{\mathrm{PL}}$ and $\hat{\beta}_{\mathrm{PPL}}$, and the coverage probability is close to the nominal level with moderate sample sizes.

We also evaluated the nonparametric test in Section 4 via a series of simulations. We compared the power of our proposed test with the conditional Kendall's tau test in Tsai (1990). We generated (A^0, T^0) from bivariate log-normal distribution truncated at τ, and the associated normal distribution has mean (μ_1, μ_2) and variance-covariance matrix $(\sigma_{ij})_{2\times 2}$. The censoring time C was generated from uniform distribution to produce different censoring rate. In all the scenarios, we set $\sigma_{11} = \sigma_{22} = 0.5$, and the other parameters are set to produce different associations and truncation proportions $\alpha = P(A^0 > T^0)$. We set $\tau = 4$ when $\mu_2 = 0$ and 6 when $\mu_2 = 0.5$. The results are presented in Table 3, with a sample size of 100 and 1000 iterations. Both tests maintains the nominal level under the null hypothesis. As expected, the proposed test substantially outperforms the conditional Kendall's tau test when the rate of truncation is high and the correlation is positive, while the two tests have similar performance when the correlation is negative.

6. DATA ANALYSIS

6.1 Analysis of Canadian Study of Health and Aging

In this section, we illustrate the proposed methods by analyzing data from the Canadian Study of Health and Aging (CSHA), one of the largest epidemiology studies of dementia (Wolfson et al., 2001; McDowell, Hill and Lindsay, 2001). CSHA recruited a prevalent cohort of individuals aged 65 and older with dementia during the period between February 1991 and May 1992. In our data analysis, the survival time of interest is the time from onset to death and the truncation time in the prevalent cohort is the duration from the onset of dementia to study enrollment. A total of 807 subjects were analyzed; among them, 249 were diagnosed with possible Alzheimer's disease, 388 had probable Alzheimer's disease, and 170 had vascular dementia.

To assess the effect of dementia subtypes on mortality, we fit a Cox proportional hazards model with indicators of probable Alzheimer (X_1) and vascular dementia (X_2) as covariates. Several authors, including Addona and Wolfson (2006) and Huang and Qin (2012), have examined the stationarity assumption with respect to disease incidence and found that the stable disease condition holds approximately. Instead of imposing an uniform distribution

for the underlying truncation time distribution, however, for illustrative purpose we apply the result obtained in Huang, Ning and Qin (2015) to employ the density function $h(t) \propto \exp(0.183t - 0.028t^2 + 0.001t^3)I(0 \le t \le \tau)$, with $\tau = 19.85$ years. Note that $h(t)$ is a member of Neyman's smooth alternative which includes uniform distribution as a special case.

We compare the proposed method to the pseudo partial likelihood approach (Tsai, 2009) by using the same truncation time density function $h(t)$ in both estimation procedures. Applying the proposed self-consistency algorithm described in Section 3.2, the estimated regression coefficients are 0.151 (asymptotic standard error [ASE], 0.065; 95% confidence interval [CI], 0.023 to 0.278) for probable Alzheimer and 0.229 (ASE, 0.079; 95% CI, 0.074 to 0.384) for vascular dementia. The estimated covariate effects are similar to the maximum likelihood estimator reported in Qin et al. (2011) obtained under the stable disease assumption. Thus our analysis suggests that probable Alzheimer and vascular dementia are associated with significantly worse survival compared to possible Alzheimer. On the other hand, the pseudo partial likelihood method gives regression coefficient estimates 0.064 (ASE, 0.082) for probable Alzheimer and 0.161 (ASE, 0.106) for vascular dementia. It is easy to see that both regression coefficients are not significantly different from 0 using Tsai's method.

6.2 Testing independent truncation for nursing home data

We next illustrate the proposed test of independence by analyzing the well-known Channing House data (Hyde, 1977). The study recorded age at entry and age at death for 462 residents of a retirement center, Channing House, from 1964 to 1975. The survival time is left-truncated by study entry and right-censored by end of study or loss to follow-up. We apply the testing statistic $\hat{\kappa}_A$ in Section 4 to test the null hypothesis that the underlying survival time and underlying the truncation time are independent of each other within each gender group.

Because the variance of the proposed testing procedure is quite complicated, we adopt the nonparametric bootstrap method with 1000 replicates to construct the 95% bootstrap CI for the test statistic. The value of the proposed test statistic is 0.005 (95% bootstrap CI, 0.003 to 0.039) for males and is −0.004 (95% bootstrap CI, −0.025 to 0.014) for females. Thus we conclude that the association between the underlying survival time and the underlying truncation time was significantly different than 0 in the male group, whereas the association was not significant in the female group. For comparison, we also apply the conditional Kendall's tau test statistic developed by Tsai (1990). The conditional Kendall's tau is 0.198 (95% bootstrap CI, 0.003 to 0.362) for males and 0.051 (95% bootstrap CI, −0.046 to 0.164) females. Hence the results of our proposed test are consistent with that based on conditional Kendall's tau test.

7. REMARKS

The main goal of this paper is to develop a unified framework for analyzing left-truncated and right-censored data with an unspecified or known truncation time distribution. Our methodologies are developed based on the idea of treating truncation and censoring as

"missing data mechanisms" and applying the missing information principle to unbiased estimating equations obtained in the absence of left truncation and right censoring. Specifically, we derived imputed estimating function from the score function derived from the full nonparametric likelihood (Section 2) and semiparametric likelihood (Section 3) with complete data. This is in contrast with the estimation procedure developed in Luo and Tsai (2009) and Tsai (2009), where the authors derived a pseudo-partial likelihood by integrating the partial likelihood over the given truncation time distribution. As a result, their estimators are not expected to be more efficient than the proposed estimators which are based on the full likelihood of complete data. Moreover, the evaluation of pseudo-partial likelihood requires estimation of the censoring time distribution and is thus less desirable.

In addition to model estimation, we also demonstrate the use of the missing information principle to hypothesis testing problem. In particular, in Section 4 we derive a new nonparametric test for checking the independence between the underlying survival time and the underlying truncation time based on Kendall's tau statistic. Unlike the conditional Kendall's tau test that are constructed based on comparable pairs subject to truncation and censoring, our new testing procedure utilizes data from all individuals and hence is expected to be more efficient. Results of simulation studies show that the proposed test enjoys a substantial gain in power, compared to the conditional Kendall's tau test, when the underlying truncation time and survival time random variables are positively correlated.

Finally, with minor modifications, the missing information principle can be applied to handle more complicated data structures, such as double truncation and competing risk models, as well as non-Cox models, such as accelerated failure time models and additive hazards models. Further research is warranted.

ACKNOWLEDGMENT

The CSHA was supported by the Seniors Independence Research Program, through the National Health Research and Development Program (NHRDP) of Health Canada (project 6606-3954-MC[S]). The progression of dementia project within the CSHA was supported by Pfizer Canada through the Health Activity Program of the Medical Research Council of Canada and the Pharmaceutical Manufacturers Association of Canada; by the NHRDP (project 6603-1417-302[R]); by Bayer; and by the British Columbia Health Research Foundation (projects 38 [93-2] and 34 [96-1]).

APPENDIX

Define $\hat{\theta} = (\hat{\beta}, \hat{\Lambda})$, $\theta_0 = (\beta_0, \Lambda_0)$, $S(\cdot \mid Z) = \exp\{-\Lambda(\cdot)\exp\beta' Z\}$ and $S_0(\cdot \mid Z) = \exp\{-\Lambda_0(\cdot)\exp\beta_0' Z\}$. The log-likelihood function based on the observed data is

$$l(\theta) = \sum_{i=1}^{n}\left[\int_0^\tau \{\beta' Z_i + \log d\Lambda(u)\}dN_i(u) - \int_0^\tau I(Y_i \geq t)\exp\{\beta' Z_i\}d\Lambda(u) - \log\int_0^\tau S(u|Z_i)dH(u)\right].$$

The score function of (β, Λ) is $U_n(\beta, \Lambda) = (U_{1n}(\beta, \Lambda), U_{2n}(\cdot, \beta, \Lambda))$, where

$$U_{1n}(\beta, \Lambda) = \frac{1}{n}\sum_{i=1}^{n}\left[\int_0^{\tau} Z_i dN_i(t) - \int_0^{\tau} Z_i\left\{I(Y_i \geq t) - \frac{\int_t^{\tau} S(u|Z_i)dH(u)}{\int_0^{\tau} S(u|Z_i)dH(u)}\right\}\exp\{\beta' Z_i\} d\Lambda(t)\right],$$

$$U_{2n}(t, \beta, \Lambda) = \frac{1}{n}\sum_{i=1}^{n}\left[\int_0^{t} dN_i(u) - \int_0^{t}\left\{I(Y_i \geq u) - \frac{\int_u^{\tau} S(v|Z_i)dH(v)}{\int_0^{\tau} S(v|Z_i)dH(v)}\right\}\exp\{\beta' Z_i\} d\Lambda(u)\right].$$

We assume the following regularity conditions for Theorem 3.1.

(A1) The true value of λ_0 is continuously differentiable. In addition, the upper bound τ of the support is finite. The parameter space of Λ contains all the nondecreasing functions Λ satisfying $\Lambda(0) = 0$ and $\Lambda(\tau) < \infty$.

(A2) The true value of β_0 is in a compact parameter space $\mathcal{B}$.

(A3) The truncation time distribution H has a density h on $[0, \tau]$.

(A4) The residual censoring time C has a continuous survival function S_C.

(A5) The covariate Z is bounded.

(A6) The matrix $-\partial E U_{1n}(\beta, \hat{\Lambda}(\cdot, \beta)) \Big/ \partial\beta$ evaluated at β_0 is positive definite.

Condition (A6) implies that the information matrix of the profile likelihood evaluated at the true value β_0 is positive definite, which is a classical condition that appears in the study of the Cox model for traditional survival data (Andersen et al. (1993), page 497). (A6) guarantees the existence and uniqueness of the solution $\hat{\beta}$ in large samples. (A6) also implies that J_0, the fisher information matrix of β for known Λ_0 is positive definite and thus the map σ_{11} defined below is invertible.

Following Qin et al. (2011), it can be shown that $\sqrt{n}U_n(\theta_0)$ converges weakly to $W = (W_1, W_2)$, where W_1 is a zero mean Gaussian random vector and W_2 is a zero mean Gaussian process. Define $\mu_0(Z) = \int_0^{\tau} S_0(t \mid Z)dH(t)$, and

$$K_1^{(l)}(t) = E\left\{Z^{\otimes l}\exp\{\beta_0' Z\}S_0(t|Z)\mu_0(Z)^{-1}\int_0^{t} S_C(u|Z)dH(u)\right\},$$

$$K_2^{(l)}(t, u) = \int_u^{\tau} E\left[Z^{\otimes l}\exp\{2\beta_0' Z\}S_0(v|Z)\mu_0(Z)^{-1}\cdot\left\{\Lambda_0(t \wedge v) - \int_0^{t}\left(\int_S^{\tau} S_0(w|Z)dH(w)\right)d\Lambda_0(s)\mu_0(Z)^{-1}\right\}\right]dv.$$

Then the Frechet derivative $\dot{U}_{\psi_0}$ is

$$\dot{U}_{\psi_0}(\beta, \Lambda) = -(\sigma_{11}(\beta) + \sigma_{12}(\Lambda), \sigma_{21}(\beta)(\cdot) + \sigma_{22}(\Lambda)(\cdot))$$

where

$$\sigma_{11}(\beta) = J_0\beta, J_0 = \int_0^\tau K_1^{(2)}(u)d\,\Lambda_0\,(u) + \int_0^\tau K_2^{(2)}(\tau,u)d\,\Lambda_0\,(u),$$
$$\sigma_{12}(\,\Lambda\,) = \int_0^\tau K_1^{(1)}(u)d\,\Lambda\,(u) + \int_0^\tau K_2^{(1)}(\tau,u)d\,\Lambda\,(u),$$
$$\sigma_{21}(\beta)(t) = \left\{\int_0^t K_1^{(1)}(u)d\,\Lambda_0\,(u) + \int_0^\tau K_2^{(1)}(t,u)d\,\Lambda_0\,(u)\right\}'\beta,$$
$$\sigma_{22}(\,\Lambda\,)(t) = \int_0^t K_1^{(0)}(u)d\,\Lambda\,(u) + \int_0^\tau K_2^{(0)}(t,u)d\,\Lambda\,(u)\,.$$

The inverse of Frechet derivative is

$$\dot{U}_{\psi_0}^{-1}(\beta,\,\Lambda) = -\begin{pmatrix}\sigma_{11}^{-1} + \sigma_{11}^{-1}\sigma_{12}\,\Phi^{-1}\,\sigma_{21}\sigma_{11}^{-1} & -\sigma_{11}^{-1}\sigma_{12}\,\Phi^{-1} \\ -\,\Phi^{-1}\,\sigma_{21}\sigma_{11}^{-1} & \Phi^{-1}\end{pmatrix}\begin{pmatrix}\beta \\ \Lambda\end{pmatrix},$$

where $\sigma_{11}^{-1}(\beta) = J_0^{-1}\beta$, the functional $\Phi = \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12}$ and Φ^{-1} exists by applying the Fredholm integral equations of the second kind. Thus $\sqrt{n}(\hat{\theta} - \theta_0)$ converges weakly to a tight mean zero Gaussian process $-\dot{U}_{\psi_0}^{-1}(W_1, W_2)$.

REFERENCES

Addona V and Wolfson DB (2006). A formal test for the stationarity of the incidence rate using data from a prevalent cohort study with follow-up. Lifetime Data Anal 12 267–284. MR2328577 [PubMed: 16917734]

Andersen PK, Borgan OR, Gill RD and Keiding N (1993). Statistical models based on counting processes. Springer Series in Statistics. Springer-Verlag, New York MR1198884

Asgharian M, M'Lan CE and Wolfson DB (2002). Length-biased sampling with right censoring: an unconditional approach. J. Amer. Statist. Assoc 97 201–209. MR1947280

Bartlett MS (1937). Some Examples of Statistical Methods of Research in Agriculture and Applied Biology. J. R. Statist. Soc. B 4 137–183.

Begun JM, Hall WJ, Huang W-M and Wellner JA (1983). Information and asymptotic efficiency in parametric–nonparametric models. Ann. Statist 11 432–452. MR696057

Bhattacharya PK, Chernoff H and Yang SS (1983). Nonparametric estimation of the slope of a truncated regression. Ann. Statist 11 505–514. MR696063

Dempster AP, Laird NM and Rubin DB (1977). Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc. Ser. B 39 1–38. With discussion. MR0501537

Huang C-Y, Ning J and Qin J (2015). Semiparametric likelihood inference for left-truncated and right-censored data. Biostatistics 16 785–798. MR3449843 [PubMed: 25796430]

Huang C-Y and Qin J (2012). Composite partial likelihood estimation under length-biased sampling, with application to a prevalent cohort study of dementia. J. Amer. Statist. Assoc 107 946–957. MR3010882

Hyde J (1977). Testing survival under right censoring and left truncation. Biometrika 64 225–230. MR0494775

Kendall M and Gibbons JD (1990). Rank correlation methods, fifth ed. A Charles Griffin Title. Edward Arnold, London MR1079065

Lancaster T (1990). The Econometric Analysis of Transition Data. Cambridge: Cambridge University Press.

Luo X and Tsai WY (2009). Nonparametric estimation for right-censored length-biased data: a pseudo-partial likelihood approach. Biometrika 96 873–886. MR2767276

Lynden-Bell D (1971). Article Navigation A Method of Allowing for Known Observational Selection in Small Samples Applied to 3CR Quasars. Mon. Not. R. Astron. Soc 155 95–118.

Martin EC and Betensky RA (2005). Testing quasi-independence of failure and truncation times via conditional Kendall's tau. J. Amer. Statist. Assoc 100 484–492. MR2160552

McDowell I, Hill G and Lindsay J (2001). An Overview of the Canadian Study of Health and Aging. International Psychogeriatirics 13 1–18.

Murphy SA and van der Vaart AW (2000). On profile likelihood. J. Amer. Statist. Assoc 95 449–485. With comments and a rejoinder by the authors. MR1803168

Ning J, Qin J and Shen Y (2014). Score estimating equations from embedded likelihood functions under accelerated failure time model. J. Amer. Statist. Assoc 109 1625–1635. MR3293615

Oakes D (2008). On consistency of Kendall's tau under censoring. Biometrika 95 997–1001. MR2461227

Orchard T and Woodbury MA (1972). A missing information principle: theory and applications. 697–715. MR0400516

Qin J, Ning J, Liu H and Shen Y (2011). Maximum likelihood estimations and EM algorithms with length-biased data. J. Amer. Statist. Assoc 106 1434–1449. MR2896847

Shen Y, Ning J and Qin J (2017). Nonparametric and semiparametric regression estimation for length-biased survival data. Lifetime Data Anal 23 3–24. MR3601682 [PubMed: 27086362]

Tsai W-Y (1990). Testing the assumption of independence of truncation time and failure time. Biometrika 77 169–177. MR1049418

Tsai WY (2009). Pseudo-partial likelihood for proportional hazards models with biased-sampling data. Biometrika 96 601–615. MR2538760 [PubMed: 22422175]

Tsai W-Y, Jewell NP and Wang M-C (1987). A note on the product-limit estimator under right censoring and left truncation. Biometrika 74 883–886.

Turnbull BW (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. J. Roy. Statist. Soc. Ser. B 38 290–295. MR0652727

Vardi Y (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. Biometrika 76 751–761. MR1041420

Vardi Y and Zhang C-H (1992). Large sample study of empirical distributions in a random-multiplicative censoring model. Ann. Statist 20 1022–1039. MR1165604

Wang M-C (1991). Nonparametric estimation from cross-sectional survival data. J. Amer. Statist. Assoc 86 130–143. MR1137104

Wang M-C (1996). Hazards regression analysis for length-biased data. Biometrika 83 343–354. MR1439788

Wang M-C, Brookmeyer R and Jewell NP (1993). Statistical models for prevalent cohort data. Biometrics 49 1–11. MR1221402 [PubMed: 8513095]

Wolfson C, Wolfson DB, Asgharian M, M'lan CE, Ostbye T, Rockwood K and Hogan DB (2001). A reevaluation of the duration of survival after the onset of dementia. New England Journal of Medicine 344 1111–1116. [PubMed: 11297701]

Yates F (1933). The Analysis of Replicated Experiments When the Field results Are Incomplete. Emporium Journal of Experimental Agriculture 1 129–142.

Table 1

Simulation summary statistics of $\hat{S}(t)$,$\hat{S}_{PPL}(t)$ and $\hat{S}_{PL}(t)$

			$\hat{S}(t)$			$\hat{S}_{\mathrm{PPL}}(t)$			$\hat{S}_{\mathrm{PL}}(t)$		
n	**Cen**	***S*(*t*)**	**Bias**	**SE**	**ESMSE**	**Bias**	**SE**	**ESMSE**	**Bias**	**SE**	**ESMSE**
						Scenario I					
100	25%	0.75	6	55	55	5	56	56	1	60	60
		0.5	5	56	56	4	56	56	2	60	60
		0.25	3	47	47	3	46	47	4	49	49
	50%	0.75	5	57	58	6	58	58	0.4	62	62
		0.5	−1	61	60	3	61	61	1	64	64
		0.25	−10	55	56	5	56	56	7	59	60
400	25%	0.75	1	28	28	1	28	28	−0.3	30	30
		0.5	1	28	28	0.3	28	28	0.1	29	29
		0.25	0.4	23	23	1	23	23	2	24	24
	50%	0.75	0.4	29	29	1	29	29	−1	31	31
		0.5	−2	30	30	0.3	30	30	0.2	32	32
		0.25	−6	27	28	2	28	28	4	29	30
						Scenario II					
100	25%	0.75	61	70	93	33	92	97	17	108	109
		0.5	48	69	84	27	79	84	16	91	93
		0.25	25	47	53	10	50	51	10	56	57
	50%	0.75	59	74	94	45	91	101	17	111	112
		0.5	45	72	85	37	83	91	16	95	96
		0.25	20	54	58	13	59	60	12	65	66
400	25%	0.75	30	42	51	12	55	56	2	64	64
		0.5	22	36	42	10	43	44	3	50	50
		0.25	10	23	25	2	25	25	1	30	30
	50%	0.75	30	43	52	19	56	59	2	65	65
		0.5	22	38	44	15	45	48	4	52	52
		0.25	8	25	27	1	28	28	1	33	33

Note: Cen is the censoring rate; Bias is the empirical bias (×1000); SE is the empirical standard error (×1000); ESMSE is the square root of empirical mean square error (×1000). $\hat{S}(t)$ is the proposed estimator; $\hat{S}_{PPL}(t)$ is the pseudo partial likelihood estimator; $\hat{S}_{PL}(t)$ is the partial likelihood estimator.

Table 2

Simulation summary statistics of $\hat{\beta}$, $\hat{\beta}_{PPL}$ and $\hat{\beta}_{PL}$

		$\hat{\beta}$				$\hat{\beta}_{\text{PPL}}$		$\hat{\beta}_{\text{PL}}$	
n	**Cen**	**Bias**	**SE**	**SEE**	**CP**	**Bias**	**SE**	**Bias**	**SE**
					Scenario I				
100	25%	(2,1)	(38,25)	(37,23)	(95,93)	(3,3)	(40,26)	(3,3)	(45,28)
	50%	(11,4)	(41,28)	(40,26)	(94,92)	(3,4)	(50,31)	(2,4)	(55,33)
400	25%	(2,0.3)	(18,12)	(18,11)	(95,94)	(1,1)	(19,12)	(0.5,1)	(21,13)
	50%	(6,2)	(21,14)	(21,13)	(93,93)	(1,1)	(23,14)	(0.3,1)	(25,16)
					Scenario II				
100	25%	(0.3,2)	(35,23)	(33,22)	(94,94)	(5,4)	(39,25)	(5,3)	(46,29)
	50%	(6,0.4)	(37,26)	(35,23)	(92,93)	(9,8)	(53,32)	(5,3)	(55,35)
400	25%	(0.3,0.1)	(16,11)	(16,11)	(95,95)	(3,2)	(18,11)	(1,1)	(21,14)
	50%	(3,1)	(18,12)	(18,11)	(94,95)	(5,4)	(24,15)	(1,0.4)	(26,16)

Note: Cen is the censoring rate; Bias is the empirical bias (×100); SE is the empirical standard error (×100); SEE is the empirical mean of the standard error estimates; CP is the empirical coverage probability (×100) of the 95% confidence interval. $\hat{\beta}$ is the proposed estimator; $\hat{\beta}_{PPL}$ is the pseudo partial likelihood estimator; $\hat{\beta}_{PL}$ is the partial likelihood estimator.

Table 3

Simulated power of the proposed test and conditional Kendall's tau test

			Proposed test			Tsai' test		
(μ_1, μ_2)	σ_{12}	α	0%	25%	50%	0%	25%	50%
(0,0)	0.3	0.51	98	87	32	44	27	16
(0,0)	0	0.51	6	7	4	5	6	4
(0,0)	−0.3	0.51	84	74	74	85	71	57
(0,0.5)	0.3	0.22	91	80	40	76	64	47
(0,0.5)	0	0.32	6	5	7	5	5	4
(0,0.5)	−0.3	0.36	93	87	82	93	86	73

Note: Tsai's test is based on conditional Kendall's tau (Tsai, 1990). 0%, 25% and 50% are the censoring rates. α is the proportion of truncation. The table presents power (×100) in each scenario.