Journal of Data Science **10**(2012), 673-691

Nonparametric Inference for Inverse Probability Weighted Estimators with a Randomly Truncated Sample

Xu Zhang
Georgia State University

Abstract: A randomly truncated sample appears when the independent variables T and L are observable if $L < T$. The truncated version Kaplan-Meier estimator is known to be the standard estimation method for the marginal distribution of T or L. The inverse probability weighted (IPW) estimator was suggested as an alternative and its agreement to the truncated version Kaplan-Meier estimator has been proved. This paper centers on the weak convergence of IPW estimators and variance decomposition. The paper shows that the asymptotic variance of an IPW estimator can be decomposed into two sources. The variation for the IPW estimator using known weight functions is the primary source, and the variation due to estimated weights should be included as well. Variance decomposition establishes the connection between a truncated sample and a biased sample with know probabilities of selection. A simulation study was conducted to investigate the practical performance of the proposed variance estimators, as well as the relative magnitude of two sources of variation for various truncation rates. A blood transfusion data set is analyzed to illustrate the nonparametric inference discussed in the paper.

Key words: Inverse probability weighted estimator, Markov processes, randomly truncated sample.

1. Introduction

A truncated sample contains realizations of the random variables (L, T) subject to the constraint $L < T$. Truncation refers to the unobservability of the random variables when $L > T$. Two types of truncation exist in a truncated sample, that is, T is left truncated by L and L is right truncated by T. Independence between T and L needs to be carefully defined. Tsai (1990) clarified that, for a truncated sample, independence between T and L is unidentifiable in the unobserved quadrant $L > T$. Independence in the observed quadrant $L < T$ is defined as quasi-independence. Truncation under quasi-independence is known as random truncation, and the sample is particularly called as a randomly truncated sample.

A fundamental problem with a randomly truncated sample is to estimate the marginal distributions of L and T. Let F and G be the distribution function of T and L, respectively. The standard estimation method for F and G is the truncated version Kaplan-Meier estimators (Kaplan and Meier, 1958; Lynden-Bell, 1971; Woodroofe, 1985). These estimators were shown to be NPMLE (Keiding and Gill, 1990), and their asymptotic properties were studied by Woodroofe (1985), and Wang, Jewell and Tsai (1986) among others.

The inverse probability weighting is a common concept in survey statistics (Horvitz and Thompson, 1952). As a feature of survey design, subjects in the population have unequal probabilities of selection. The standard adjustment is to weight the individual measurement by the reciprocal of the probability of selection. For a sample with known sampling rule, Vardi (1985, Section 8) suggested an inverse probability weighted (IPW) form for estimating the distribution function of a random variable. Vardi pointed out that the proposed estimator is NPMLE and described its limiting distribution. Wang (1989) studied the IPW estimator for a randomly truncated sample, with additional knowledge of the parametric distribution of one variable. The reciprocal of a probability based on the known parametric distribution was used as the weight in the IPW estimator. Wang demonstrated in a simulation study that the IPW estimator using the knowledge of the given parametric distribution is dramatically more efficient than the truncated version Kaplan-Meier estimator. In Wang's paper, the asymptotic variance of the IPW estimator was decomposed into the variation for the IPW estimator using known weight functions, which agrees with the variation of Vardi's IPW estimator (1985), and the variation due to estimation of the weight functions. In another paper, Wang (1991) studied the asymptotic properties of the nonparametric IPW estimator based on a censored and truncated sample. Shen (2003) elucidated the equivalence between the IPW estimators and the truncated version Kaplan-Meier estimators for a randomly truncated sample in which both distributions are completely unspecified.

This paper emphasizes on the weak convergence result of the IPW estimators for a randomly truncated sample, based on Keiding and Gill's Markov process model (1990). The previous study conducted by Wang (1991) has investigated the asymptotic properties of the IPW estimator. There are several differences between the current study and Wang's study. First, Keiding and Gill's framework offers a convenient path for analytically describing the asymptotic variances of the IPW estimators, as well as discovering a connection to Vardi's result (1985). Second, the variation process of a martingale was employed in this study to derive the asymptotic distributions of the IPW estimators. Furthermore, the backwards counting process and martingale process were formally defined so that it is convenient to investigate the inference of the IPW estimator of F.

The Markov process model is presented in this paper, followed by an extension to the nonparametric inference of the IPW estimators. One goal of the paper is to investigate the practical performance of the variance estimators, as well as variance decomposition. The simulation study result shows essentially the same performance between the regular variance estimator derived from a truncated version Kaplan-Meier estimator, and the variance estimator for an IPW estimator. It can be further concluded that, for an IPW estimator, the variation due to estimation of weights increases for a higher rate of truncation.

The remainder of the paper is organized as follows. Section 2 describes Keiding and Gill's representation of the truncation model, as well as the truncated version Kaplan-Meier estimators. Section 3 provides the weak convergence result of the IPW estimators. The result of the simulation study is presented in Section 4. A blood transfusion data set is analyzed in Section 5. The concluding remarks are given in Section 6.

2. Truncation Model and Kaplan-Meier Estimators

Suppose that a truncated sample contains realizations of (L, T) with the constraint $L < T$, and quasi-independence is assumed. Let G and F be the distribution functions of L and T, and let $S = 1 - F$. We assume that F and G are continuous functions on $[0, \infty)$. The truncated sample can be summarized as $(\widetilde{L}_i, \widetilde{T}_i)$, for $i = 1, \cdots, n$, and $\widetilde{L}_i < \widetilde{T}_i$.

Keiding and Gill (1990) used a Markov process to describe random truncation models. Some details of Keiding and Gill's work are provided in this section. The conditional truncation model given $L < T$ can be described by three states of a process $U(x)$,

$$\boxed{0} \xrightarrow{\lambda_1(x)} \boxed{1} \xrightarrow{\lambda_2(x)} \boxed{2}$$

where $U(x) = 0$ when $x < L \wedge T$, $U(x) = 1$ when $L \leq x < T$ and $U(x) = 2$ when $L < T \leq x$. $\lambda_1(x)$ and $\lambda_2(x)$ are respectively the intensities of transition from State 0 to 1, and from State 1 to 2, given $L < T$. It can be further recognized that $\lambda_2(x)$ coincides with the intensity of T, $\phi(x) = dF(x)/\{1-F(x)\}$. The cumulative intensities are given by $\Lambda_1(x) = \int_0^x \lambda_1(u)du$, $\Lambda_2(x) = \Phi(x) = \int_0^x \lambda_2(u)du = \int_0^x \phi(u)du$.

Define the counting processes, $N_{T,i}(x) = I(\widetilde{L}_i \leq \widetilde{T}_i \leq x)$, $N_{L,i}(x) = I(\widetilde{L}_i \leq x)$, $Y_i(x) = I(\widetilde{L}_i \leq x \leq \widetilde{T}_i)$. Let $N_T(x) = \sum_{i=1}^n N_{T,i}(x)$, $N_L(x) = \sum_{i=1}^n N_{L,i}(x)$

and $Y(x) = \sum_{i=1}^n Y_i(x)$. $\widehat{\Lambda}_1(x)$ and $\widehat{\Phi}(x)$ are respectively the Nelson-Aalen estimators of $\Lambda_1(x)$ and $\Phi(x)$, with the explicit expressions

$$\widehat{\Lambda}_1(x) = \int_0^x \frac{dN_L(u)}{\sum_{i=1}^n I(\widetilde{L}_i \geq u)} \quad \text{and} \quad \widehat{\Lambda}_2(x) = \widehat{\Phi}(x) = \int_0^x \frac{dN_T(u)}{Y(u)}.$$

$\widehat{\Lambda}_1(x) - \Lambda_1(x)$ and $\widehat{\Phi}(x) - \Phi(x)$ are square integrable martingales and orthogonal between each other. It is known that $\sqrt{n}(\widehat{\Phi}(x) - \Phi(x))$ converges in distribution to a Gaussian process with variation process $\int_0^x \{\phi(u)du/y(u)\}$, where $y(u) = E[n^{-1}Y(u)]$.

The product-limit estimator of S has the form

$$\widehat{S}(x) = 1 - \widehat{F}(x) = \prod_{u \in [0,x]} \left(1 - d\widehat{\Phi}(u)\right) = \prod_{u \in [0,x]} \left(1 - \frac{dN_T(u)}{Y(u)}\right). \tag{2.1}$$

It is the standard result that $[\widehat{S}(x)/S(x) - 1]$ relates to a square integrable martingale. The variation processes of a martingale lead to the variance estimators of $\widehat{S}(x)$ (Andersen *et al.*, 1993). The predictable variation process yields the famous Greenwood's formula, while the optional variation process yields the common variance estimator,

$$\widehat{\sigma}^2_{L,\mathrm{KM}}(x) = \widehat{S}(x)^2 \int_0^x \frac{dN_T(u)}{Y(u)^2}. \tag{2.2}$$

Regarding estimation of the distribution function of L, the standard method is to define the reverse-time hazard, or retro-hazard, $\gamma^*(x) = dG(x)/G(x)$. The retro-hazard was introduced to reflect the hazard rate of the transformed variable $L_\tau = \tau - L$, which is left truncated by $T_\tau = \tau - T$. More specifically, $\gamma^*(x)$ agrees with the hazard rate of L_τ at $\tau - x$ (Lagakos *et al.*, 1988, Section 3). Here τ is a large constant and, practically, one can choose a value greater than the largest observed value of the truncated sample. Keiding and Gill considered the process using the reversed time,

$$\boxed{0} \xleftarrow{\quad \lambda_1^*(x) \quad} \boxed{1} \xleftarrow{\quad \lambda_2^*(x) \quad} \boxed{2}$$

$\lambda_1^*(x)$ and $\lambda_2^*(x)$ are respectively the intensity processes of transitions from State 1 to 0, and from State 2 to 1, given $L < T$. It should be noted that $\lambda_1^*(x)$ agrees with $\gamma^*(x)$. Let $\Lambda_1^*(x) = \Gamma^*(x) = \int_\infty^x \lambda_1^*(u)du = \int_\infty^x \gamma^*(u)du$ and $\Lambda_2^*(x) = \int_\infty^x \lambda_2^*(u)du$. Both $\Gamma_1^*(x)$ and $\Lambda_2^*(x)$ are defined as decreasing functions so that they can be treated as the cumulative hazard functions of the transformed variables. Define the counting processes $N_{L,i}^*(x) = I(\widetilde{L}_i \geq x)$ and $N_{T,i}^*(x) = I(\widetilde{T}_i \geq x)$. Let

$N_L^*(x) = \sum_{i=1}^n N_{L,i}^*(x)$ and $N_T^*(x) = \sum_{i=1}^n N_{T,i}^*(x)$. The Nelson-Aalen estimators of the cumulative retro-hazards are given by

$$\widehat{\Lambda}_1^*(x) = \widehat{\Gamma}^*(x) = \int_\infty^x \frac{dN_L^*(u)}{Y(u)}, \qquad \widehat{\Lambda}_2^*(x) = \int_\infty^x \frac{dN_T^*(u)}{\sum_{i=1}^n I(\widetilde{T}_i \le u)}.$$

$\widehat{\Gamma}^*(x) - \Gamma^*(x)$ and $\widehat{\Lambda}_2^*(x) - \Lambda_2^*(x)$ are orthogonal square integrable martingales. $\sqrt{n}(\widehat{\Gamma}^*(x) - \Gamma^*(x))$ converges in distribution to a Gaussian process with variation process $\int_\infty^x \{\gamma^*(u)du/y(u)\}$. The product-limit estimator of $G(x)$ is

$$\widehat{G}(x) = \prod_{u \in (x,\infty)} \left(1 - \frac{d\sum_{i=1}^n I(\widetilde{L}_i \le u)}{Y(u)}\right). \tag{2.3}$$

Similarly, $\widehat{G}(x)/G(x) - 1$ can be shown to be a square integrable martingale, and variation processes of a martingale lead to the variance estimators. The common variance estimator for $\widehat{G}(x)$ is given by

$$\widehat{\sigma}_{L,\mathrm{KM}}^2(x) = \widehat{G}(x)^2 \int_\infty^x \frac{dN_L^*(u)}{Y(u)^2}. \tag{2.4}$$

3. IPW Estimators

Let $H(x,y)$ be the joint distribution function of L and T given $L < T$, and let $\beta = P(L < T)$ denote the un-truncated probability. Then

$$H(x,y) = P(T \le x, L \le y | L < T) = \beta \int_0^x G(s \wedge y) dF(s).$$

Let $G^*(x)$ and $F^*(x)$ be respectively the marginal distributions of L and T, given $L < T$. $G^*(x)$ and $F^*(x)$ are given by

$$G^*(x) = P(L^* \le x) = H(\infty, x) = \beta^{-1} \int_0^x [1 - F(s-)] dG(s) \tag{3.1}$$

and

$$F^*(x) = P(T^* \le x) = H(x, \infty) = \beta^{-1} \int_0^x G(s) dF(s). \tag{3.2}$$

The rearrangement of the above equations leads to

$$G(x) = \beta \int_0^x [1 - F(s-)]^{-1} dG^*(s) \tag{3.3}$$

and

$$F(x) = \beta \int_0^x G(s)^{-1} dF^*(s). \tag{3.4}$$

(3.1)-(3.4) suggest several equivalent forms for β,

$$\begin{aligned}\beta &= \int_0^\infty G(s)dF(s) = \int_0^\infty [1 - F(s-)]dG(s) \\ &= \left[\int_0^\infty [1 - F(s-)]^{-1} dG^*(s)\right]^{-1} = \left[\int_0^\infty G(s)^{-1} dF^*(s)\right]^{-1}.\end{aligned}$$

One can use the product-limit estimators, $1 - \widehat{F}(x)$ and $\widehat{G}(x)$, for the functions $1 - F(x)$ and $G(x)$. $F^*(x)$ and $G^*(x)$ can be directly estimated by the empirical estimators, $\widehat{F}^*(x) = n^{-1} \sum_{i=1}^n I(\widetilde{T}_i \le x)$ and $\widehat{G}^*(x) = n^{-1} \sum_{i=1}^n I(\widetilde{L}_i \le x)$. There will be four plug-in estimators for β,

$$\begin{aligned}\widehat{\beta}_1 &= \int_0^\infty \widehat{G}(s)d\widehat{F}(s), \\ \widehat{\beta}_2 &= \int_0^\infty [1 - \widehat{F}(s-)]d\widehat{G}(s), \\ \widehat{\beta}_3 &= \left[\int_0^\infty [1 - \widehat{F}(s-)]^{-1} d\widehat{G}^*(s)\right]^{-1}, \\ \widehat{\beta}_4 &= \left[\int_0^\infty \widehat{G}(s)^{-1} d\widehat{F}^*(s)\right]^{-1}.\end{aligned}$$

Note that all four estimators are evaluated the same. A formal proof of the equivalence between $\widehat{\beta}_3$ and $\widehat{\beta}_4$ can be found in Shen (2005), in a general context in which both truncation and censoring are present. The estimator $\widehat{\beta}_1$ is more frequently used and its asymptotic properties were studied by among others Chao (1987) and He and Yang (1998). The asymptotic distribution of $\widehat{\beta}_3$ was studied by Keiding and Gill (1990).

In (3.3) and (3.4), using the product-limit estimators for G, F and the empirical estimators for G^*, F^*, one will have the IPW estimators of G and F.

$$\begin{aligned}\widehat{G}^{\mathrm{IPW}}(x) &= \widehat{\beta}_3 \int_0^x [1 - \widehat{F}(u-)]^{-1} d\widehat{G}^*(u) = \widehat{\beta}_3 \; n^{-1} \sum_{i=1}^n \frac{1}{1 - \widehat{F}(\widetilde{L}_i-)} I(\widetilde{L}_i \le x), \\ \widehat{\beta}_3 &= \left[\int_0^\infty [1 - \widehat{F}(u-)]^{-1} d\widehat{G}^*(u)\right]^{-1} = \left[n^{-1} \sum_{i=1}^n \frac{1}{1 - \widehat{F}(\widetilde{L}_i-)}\right]^{-1},\end{aligned}$$

$$\begin{aligned}\widehat{F}^{\mathrm{IPW}}(x) &= \widehat{\beta}_4 \int_0^x \widehat{G}(u)^{-1} d\widehat{F}^*(u) = \widehat{\beta}_4 \; n^{-1} \sum_{i=1}^n \frac{1}{\widehat{G}(\widetilde{T}_i)} I(\widetilde{T}_i \le x), \\ \widehat{\beta}_4 &= \left[\int_0^\infty \widehat{G}(u)^{-1} d\widehat{F}^*(u)\right]^{-1} = \left[n^{-1} \sum_{i=1}^n \frac{1}{\widehat{G}(\widetilde{T}_i)}\right]^{-1}.\end{aligned}$$

Regarding the distribution of T, it is more convenient to investigate the IPW estimator of $S = 1 - F$. Define $S^*(x) = 1 - F^*(x)$ and its empirical estimator is given by $\widehat{S}^*(x) = n^{-1}\sum_{i=1}^{n} I(\widetilde{T}_i > x)$. Then the IPW estimator of S is

$$\widehat{S}^{\text{IPW}}(x) = 1 - \widehat{F}^{\text{IPW}}(x) = \widehat{\beta}_4 \int_{\infty}^{x} \widehat{G}(u)^{-1} d\widehat{S}^*(u) = \widehat{\beta}_4 \; n^{-1} \sum_{i=1}^{n} \frac{1}{\widehat{G}(\widetilde{T}_i)} I(\widetilde{T}_i > x),$$

$$\widehat{\beta}_4 = \left[\int_{\infty}^{0} \widehat{G}(u)^{-1} d\widehat{S}^*(u)\right]^{-1} = \left[n^{-1} \sum_{i=1}^{n} \frac{1}{\widehat{G}(\widetilde{T}_i)}\right]^{-1}.$$

Shen (2003) showed that the IPW estimators are essentially alternative expressions of the Kaplan-Meier estimators given in (2.1) and (2.3). Vardi (1985) studied distribution estimation based on s samples subject to s sets of sampling rule. When $s = 1$, the problem reduces to distribution estimation on a biased sample with known probabilities of selection. Vardi suggested the inverse probability weighted estimator and presented the asymptotic properties of the estimator (Section 8). Marginal distribution estimation based on a truncation model has some similarities to the problem studied by Vardi. In a truncated sample, the observed values $(\widetilde{L}_1, \cdots, \widetilde{L}_n)$ (or $(\widetilde{T}_1, \cdots, \widetilde{T}_n)$) can be viewed as a biased sample with unknown, but estimable, probabilities of selection. It will be shown in the remainder of this section and the appendices that the asymptotic variance of each IPW estimator contains two components: one component is identical to Vardi's variance formula for a biased sample with known probabilities of selection, and the other component is the variation due to estimated weights.

The following conditions are required for weak convergence properties of the IPW estimators,

$$\int_0^{\infty} [1 - F(u-)]^{-1} dG(u) < \infty, \quad \int_0^{\infty} G(u)^{-1} dF(u) < \infty.$$

The weak convergence results are presented in this section and the derivations are elaborated in Appendices A and B. The IPW estimator $\sqrt{n}[\widehat{G}^{\text{IPW}}(x) - G(x)]$ converges in distribution to a zero-mean normal random variable with variance

$$\begin{aligned}
&\beta \int_0^{x} [1 - F(u-)]^{-1} dG(u) + \beta G(x)^2 \int_0^{\infty} [1 - F(u-)]^{-1} dG(u) \\
&- 2\beta G(x) \int_0^{x} [1 - F(u-)]^{-1} dG(u) + \int_0^{x} G(u)^2 (1 - G(x))^2 \frac{\phi(u)du}{y(u)} \\
&+ \int_x^{\infty} G(x)^2 (1 - G(u))^2 \frac{\phi(u)du}{y(u)}.
\end{aligned}$$

The expression of the first row agrees with Vardi's result for a biased sample with known probability of selection. The remainder is shown in Appendix A to reflect the variation caused by estimation of weights. The IPW estimator $\sqrt{n}[\widehat{S}^{\mathrm{IPW}}(x) - S(x)]$ converges in distribution to a zero-mean normal random variable with variance

$$\beta \int_{\infty}^{x} G(u)^{-1} dS(u) + \beta S(x)^2 \int_{\infty}^{0} G(u)^{-1} dS(u) - 2\beta S(x) \int_{\infty}^{x} G(u)^{-1} dS(u)$$
$$+ \int_{\infty}^{x} S(u)^2 [1 - S(x)]^2 \frac{\gamma^*(u) du}{y(u)} + \int_{x}^{0} S(x)^2 [1 - S(u)]^2 \frac{\gamma^*(u) du}{y(u)}.$$

The variance of $\widehat{G}^{\mathrm{IPW}}(x)$ can be estimated by

$$\widehat{\sigma}_L^2(x) = \widehat{\sigma}_{L,1}^2(x) + \widehat{\sigma}_{L,2}^2(x), \tag{3.5}$$

where

$$\widehat{\sigma}_{L,1}^2(x) = n^{-1} \left\{ \widehat{\beta}_3 \int_0^x [1 - \widehat{F}(u)]^{-1} d\widehat{G}(u) + \widehat{\beta}_3 \widehat{G}(x)^2 \int_0^{\infty} [1 - \widehat{F}(u)]^{-1} d\widehat{G}(u) \right.$$
$$\left. - 2\widehat{\beta}_3 \widehat{G}(x) \int_0^x [1 - \widehat{F}(u)]^{-1} d\widehat{G}(u) \right\}, \tag{3.6}$$

$$\widehat{\sigma}_{L,2}^2(x) = \int_0^x \widehat{G}(u)^2 (\widehat{G}(x) - 1)^2 \frac{dN_T(u)}{Y(u)^2} + \int_x^{\infty} \widehat{G}(x)^2 (\widehat{G}(u) - 1)^2 \frac{dN_T(u)}{Y(u)^2}. \tag{3.7}$$

Similarly, the variance of $\widehat{S}^{\mathrm{IPW}}(x)$ can be estimated by

$$\widehat{\sigma}_T^2(x) = \widehat{\sigma}_{T,1}^2(x) + \widehat{\sigma}_{T,2}^2(x), \tag{3.8}$$

$$\widehat{\sigma}_{T,1}^2(x) = n^{-1} \left\{ \widehat{\beta}_4 \int_{\infty}^{x} \widehat{G}(u)^{-1} d\widehat{S}(u) + \beta \widehat{S}(x)^2 \int_{\infty}^{0} \widehat{G}(u)^{-1} d\widehat{S}(u) \right.$$
$$\left. - 2\widehat{\beta}_4 \widehat{S}(x) \int_{\infty}^{x} \widehat{G}(u)^{-1} d\widehat{S}(u) \right\}, \tag{3.9}$$

$$\widehat{\sigma}_{T,2}^2(x) = \int_{\infty}^{x} \widehat{S}(u)^2 [1 - \widehat{S}(x)]^2 \frac{dN_L^*(u)}{Y(u)^2} + \int_x^0 \widehat{S}(x)^2 [1 - \widehat{S}(u)]^2 \frac{dN_L^*(u)}{Y(u)^2}. \tag{3.10}$$

4. Simulation Study

The goal of the simulation study is to investigate the practical performances of the variance estimators, as well as to assess the relative proportion of two sources of variation. Two sets of simulation study were conducted separately for the variance estimators given in Formulas (3.5) and (3.8). In both sets of study, the underlying distributions of L and T are respectively uniform and shifted exponential distributions, with explicit formulas given by

$$G(x) = x/c, \quad 0 \le x \le c,$$

and

$$F(x) = 1 - e^{\lambda(x-u)}, \quad x > u, u > 0.$$

It is known that only conditional distribution functions can be estimated based on a truncated sample. Let $a = \min\{\widetilde{L}_1, \cdots, \widetilde{L}_n\}$ and $b = \max\{\widetilde{T}_1, \cdots, \widetilde{T}_n\}$. Practically, the conditional distributions $G_b(x) = P(L \le x | L \le b)$ and $F_a(x) = P(T \le x | T \ge a)$ are estimable. If the interior support of G has an upper limit c and $c < b$, then $G(x) = G_b(x)$. Similarly, if the lower limit of the support of F is greater than a, $F(x) = F_a(x)$. Let $a^{(i)}, b^{(i)}$ be respectively the smallest $\widetilde{L}$ values and the largest $\widetilde{T}$ values for the ith replicate. By choosing the uniform distribution for G, shifted exponential for F, and selecting appropriate parameter values, we can guarantee that $c < b^{(i)}$ and $d > a^{(i)}$, $\forall\, i$. Therefore, the evaluable distribution functions will not vary across replicates and the simulation result can be assessed using the true functions G and F.

The first set of the simulation study focused on estimation of G. Among the settings generated in this category, G was fixed to be the uniform distribution on the interval [0, 1]. F followed shifted exponential, $F(x) = 1 - \exp\{\lambda(x-u)\}, x > u, u > 0$, where the values of λ and u were selected to yield the truncation rates 25% and 50%, respectively. The truncation rate is defined as $(N-n)/N$, where n is the size of the truncated sample and N is the size of the population from which the truncated sample is selected. Two levels of n, 200 and 400, were considered in the simulation study, and a total of 1000 replicates were generated for each setting. The estimation outcome was evaluated at four x values that yield 0.2, 0.4, 0.6, 0.8 in $G(x)$. Let $\widehat{G}^{(i)}(x)$ be the Kaplan-Meier estimate of $G(x)$ for the ith sample, and note that the IPW estimate agrees with the Kaplan-Meier estimate. Bias and sample variance are included in Table 1, and these quantities were computed by

$$\text{Bias} = E[\widehat{G}(x)] - G(x), \quad \text{where } E[\widehat{G}(x)] = \frac{1}{1000}\sum_{i=1}^{1000} \widehat{G}^{(i)}(x),$$

$$\text{var}[\widehat{G}(t)] = \frac{1}{1000-1}\sum_{i=1}^{1000}\left[\widehat{G}^{(i)}(x) - E[\widehat{G}(x)]\right]^2.$$

In Table 1, columns with labels $\widehat{\sigma}^2_{L,\text{KM}}$, $\widehat{\sigma}^2_L, \widehat{\sigma}^2_{L,1}, \widehat{\sigma}^2_{L,2}$ are averages of individual variance estimates across all replicates, using Formulas (2.2), (3.5)-(3.7), respectively. Two components that make up $\widehat{\sigma}^2_L$ are explicitly reported to reveal the relative magnitude between two sources of variation, while individual contribution is further illustrated by the average percentage of $\widehat{\sigma}^2_{L,1}/\widehat{\sigma}^2_L$. The 95% confidence intervals were computed for all the samples, using $\widehat{\sigma}^2_L$, and its major component $\widehat{\sigma}^2_{L,1}$, respectively. The actual coverage rates are included in the last two columns of Table 1.

Table 1: Simulation result for estimating the distribution function of L

n	L%	$G(x)$	Bias	var($\widehat{G}(x)$)	$\widehat{\sigma}^2_{L,\mathrm{KM}}(x)$	$\widehat{\sigma}^2_L(x)$	$\widehat{\sigma}^2_{L,1}(x)$	$\widehat{\sigma}^2_{L,2}(x)$	$\widehat{\sigma}^2_{L,1}(x)/\widehat{\sigma}^2_L(x)$ (%)	$\widehat{\sigma}^2_L(x)$ coverage	$\widehat{\sigma}^2_{L,1}(x)$ coverage
200	25	0.2	0.002	0.77	0.76	0.77	0.68	0.09	88	0.944	0.926
		0.4	0.002	1.32	1.31	1.32	1.16	0.15	88	0.937	0.920
		0.6	0.001	1.50	1.43	1.44	1.32	0.13	91	0.941	0.930
		0.8	0.002	1.07	1.03	1.04	0.99	0.05	95	0.934	0.931
	50	0.2	-0.000	0.75	0.80	0.81	0.60	0.21	74	0.950	0.907
		0.4	-0.002	1.62	1.69	1.70	1.29	0.41	76	0.949	0.912
		0.6	-0.002	2.07	2.16	2.18	1.78	0.40	82	0.953	0.926
		0.8	-0.002	1.72	1.77	1.79	1.61	0.18	90	0.952	0.944
400	25	0.2	0.000	0.36	0.38	0.39	0.34	0.05	88	0.955	0.944
		0.4	-0.000	0.67	0.66	0.66	0.58	0.08	88	0.945	0.927
		0.6	-0.001	0.71	0.72	0.73	0.66	0.07	91	0.946	0.938
		0.8	-0.000	0.50	0.52	0.52	0.50	0.03	95	0.954	0.948
	50	0.2	-0.000	0.39	0.40	0.41	0.30	0.11	73	0.941	0.912
		0.4	-0.000	0.80	0.85	0.85	0.65	0.21	76	0.958	0.915
		0.6	-0.000	1.07	1.09	1.09	0.89	0.20	82	0.951	0.924
		0.8	-0.000	0.82	0.89	0.89	0.80	0.09	90	0.957	0.943

The second set of simulation targeted at estimation of $S = 1 - F$. A total of four settings were generated, in which the shifted exponential, $F(t) = 1 - \exp\{-(x - 0.2)\}, x > 0.2$, was used as the underlying distribution. The underlying distribution of L was uniform $[0, c]$, where value of c varied to meet the predetermined truncation rates, 25% and 50%, respectively. Each setting contained 1000 replicates with two levels of n, 200 and 400. Estimation outcomes were calculated at x values yielding 0.8, 0.6, 0.4, 0.2 in $S(x)$. Simulation results were summarized in Table 2, which contains the similar entries as Table 1.

The following findings can be concluded from Tables 1 and 2. First, the variance estimators directly derived from the IPW estimators are evaluated very close to the regular variance estimators, and both type of variance estimates closely measure the variation contained among distribution probability estimates. Second, when the IPW estimator based variance is decomposed into two components, the variation due to the IPW estimator using known weight function is the major source. It explains 80%-90% of total variation under a light truncation. With a

Table 2: Simulation result for estimating the survival function of T

n	$L\%$	$S(x)$	Bias	$\mathrm{var}(\widehat{S}(x))$	$\widehat{\sigma}^2_{T,\mathrm{KM}}(x)$	$\widehat{\sigma}^2_T(x)$	$\widehat{\sigma}^2_{T,1}(x)$	$\widehat{\sigma}^2_{T,2}(x)$	$\widehat{\sigma}^2_{T,1}(x)/\widehat{\sigma}^2_T(x)$ (%)	$\widehat{\sigma}^2_T(x)$ coverage	$\widehat{\sigma}^2_{T,1}(x)$ coverage
200	25	0.8	-0.001	1.87	1.97	1.95	1.73	0.22	89	0.946	0.928
		0.6	0.000	2.07	2.10	2.12	1.78	0.34	84	0.953	0.928
		0.4	-0.001	1.43	1.48	1.50	1.30	0.20	87	0.950	0.924
		0.2	-0.001	0.74	0.74	0.76	0.71	0.05	93	0.942	0.937
	50	0.8	0.000	2.73	2.72	2.67	2.33	0.34	88	0.936	0.915
		0.6	0.000	3.05	2.92	2.95	2.34	0.61	80	0.951	0.920
		0.4	0.001	2.20	2.06	2.09	1.53	0.56	73	0.935	0.885
		0.2	-0.000	0.95	0.87	0.88	0.63	0.25	71	0.941	0.872
400	25	0.8	0.002	0.97	0.99	0.98	0.88	0.10	90	0.944	0.926
		0.6	0.001	0.99	1.06	1.06	0.89	0.17	84	0.962	0.942
		0.4	0.001	0.72	0.75	0.75	0.65	0.10	87	0.957	0.944
		0.2	0.001	0.38	0.37	0.38	0.35	0.03	93	0.944	0.940
	50	0.8	0.001	1.42	1.37	1.36	1.19	0.16	88	0.938	0.925
		0.6	0.001	1.52	1.47	1.48	1.18	0.30	80	0.947	0.918
		0.4	0.002	1.04	1.04	1.05	0.77	0.28	73	0.953	0.910
		0.2	0.001	0.41	0.44	0.44	0.32	0.13	71	0.955	0.913

50% of truncation, the major component still accounts for at least 70% of the total variation. Finally, although the variation due to estimation of the weights is a minor component of the total variation, it should not be practically omitted because the actual coverage rate becomes obviously worse. On the contrary, the confidence intervals using the variance estimator with both components achieve the nominal coverage level.

A nonparametric IPW estimator agrees with the truncated version Kaplan-Meier estimator, and consequently is the nonparametric maximum likelihood estimator (NPMLE) when the truncation distribution is fully unspecified. When the parametric distribution of the truncation variable can be determined, Wang (1989) showed that the semiparametric IPW estimator becomes the MLE and is more efficient than the truncated version Kaplan-Meier estimator. If additional evidence suggests $G(\bullet;\theta)$ be the parametric form of the distribution function of L, the semiparametric estimator,

$$\widehat{F}(x;\theta) = \left[n^{-1}\sum_{i=1}^{n}\frac{1}{G(\widetilde{T}_i;\theta)}\right]^{-1} \times n^{-1}\sum_{i=1}^{n}\frac{1}{G(\widetilde{T}_i;\theta)}I(\widetilde{T}_i \le x),$$

maximize the marginal likelihood

$$L_m(\widetilde{\mathbf{T}};\theta,F) = \prod_i \frac{G(\widetilde{T}_i;\theta)dF(\widetilde{T}_i)}{\int G(u;\theta)dF(u)}.$$

$\widehat{\theta}$, the MLE of θ, maximizes the conditional likelihood

$$L_c(\widetilde{\mathbf{L}}|\widetilde{\mathbf{T}}, \theta) = \prod_i \frac{dG(\widetilde{L}_i; \theta)}{G(\widetilde{T}_i; \theta)}.$$

The final estimator, $\widehat{F}(x; \widehat{\theta})$, maximizes the full likelihood $L_m \times L_c$. This semiparametric IPW estimator is associated with higher degree of efficiency, and should be advocated whenever the truncation distribution is parameterized.

5. Analysis of Blood Transfusion Data

All AIDS onsets are required to be reported to Center of Disease Control and Prevention (CDC). If patients infected HIV virus from blood transfusion, the exact infection dates can be retrospectively obtained. The AIDS incubation time is the duration from infection of HIV virus to onset of AIDS. Given a cut-off date, an AIDS could not be collected by CDC if the incubation time is greater than the duration between the infection date and the cut-off date. Thus, the observed incubation times are right truncated.

Kalbfleisch and Lawless (1989) provided one blood transfusion data set, including 295 AIDS cases diagnosed by the cut-off date July 1, 1986 and reported to CDC no later than January 1, 1987. The infection dates for those 295 AIDS cases spanned from April 1978 to February 1986. Suppose one is interested in estimating the AIDS incubation time for the individuals in population consisting of all transfusion-related patients infected by July, 1986. The infected subjects were unobservable if their AIDS diagnoses occurred after the cut-off data. As a consequence, the AIDS incubation time is right truncated by the duration between infection and July 1, 1986. This data set has been used as a typical example of right truncation and analyzed by many researchers. Bilker and Wang (1996) carefully noted that, for this particular AIDS data set, left truncation was also present because AIDS was first diagnosed in 1982. The subjects infected before 1982 were not likely to be included in the sample if onset of AIDS happened before the disease was recognized. Therefore, the incubation time was left truncated by the duration between infection date and beginning of 1982. In this study, the blood transfusion data set was used as an illustrative example of right truncation, and the underlying left truncation was not considered.

The AIDS cases were broken down into three age groups, 1-4, 5-59 and 60+, known as "children", "adults" and "elderly patients", respectively. Distribution estimates for three age groups are shown in Figure 1, together with two sets of linear confidence intervals. The confidence intervals using the regular variance estimator are shown as dashed lines. The variance estimator given in Section 3 was evaluated, and the subsequent confidence intervals were plotted in dotted

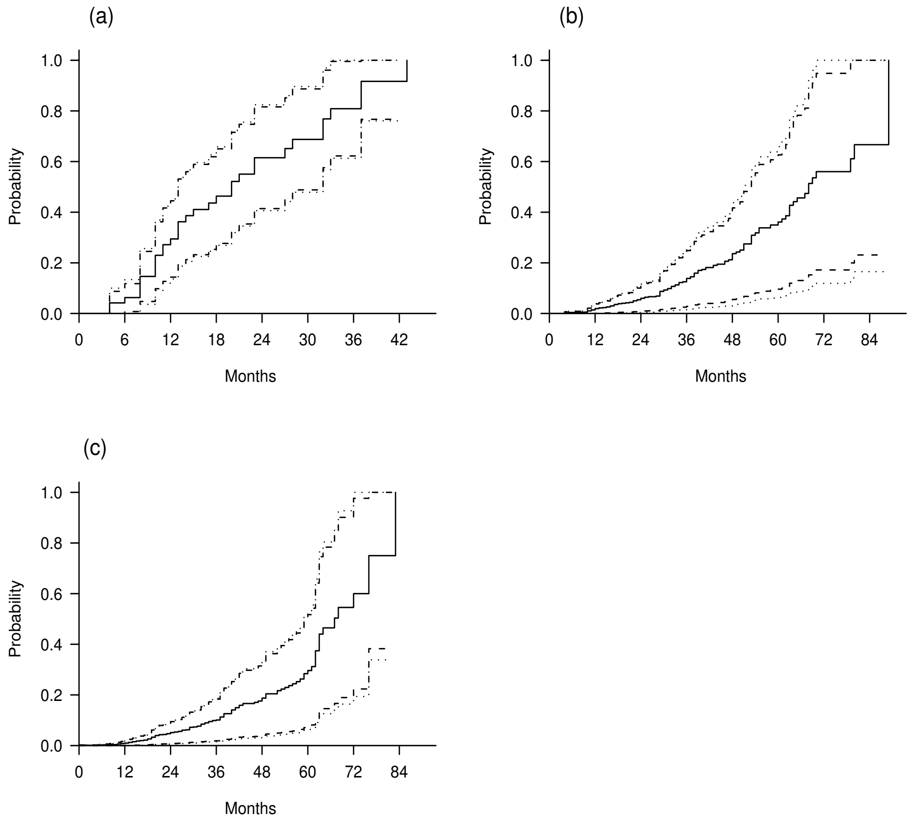

Figure 1: The estimated distribution functions of AIDS incubation time for (a) Children, (b) Adults and (c) Elderly Patients. Each panel contains two sets of 95% confidence intervals, using the regular variance estimator (dashed line) and the variance estimator derived from the IPW estimator (dotted line)

lines. In children and elderly patients, two sets of confidence intervals highly agree with each other. In adults, the confidence interval using the new variance estimator is slight wider.

6. Concluding Remarks

This paper studies nonparametric inference of IPW estimators based on a randomly truncated sample. Development of the inference relied on Keiding and Gill's Markov process model. This conceptual framework leads to the analytical description of asymptotic variances of IPW estimators. Variance decomposition investigated in this paper was aimed at revealing the dissimilarity in precision assessment about an IPW estimator between a truncated sample and a biased sample with known probabilities of selection.

The nonparametric inference developed in this paper can be extended to the context of both censoring and truncation. Let C be the censoring variable, and let $X = \min(T, C)$. The censored and truncated sample can be characterized by (L, X) subject to the condition $L < T$. Under this context, Wang (1991) proposed the IPW estimator of G and established its asymptotic properties, while the inference of the IPW estimator of F has not been studied. The explicit expression of the IPW estimator of F can be found in Shen (2003). This estimator uses the cumulative probability estimates of both L and C. It's asymptotic distribution can be similarly derived using the Keiding and Gill's Markov process model. The total variation of the IPW estimator should contain the additional variation caused by estimation of the probability of the censoring variable. Further research is needed to fully describe the asymptotic distribution of the IPW estimator of F.

The IPW estimators, instead of the truncated version Kaplan-Meier estimators, have to be considered under certain contexts of truncation. The nonparametric inference discussed in this paper will help to elucidate the asymptotic properties of these IPW estimators. One application of the IPW estimator is distribution function estimation with certain type of dependently truncated sample. Jones and Clowley (1992) introduced a Cox-model-based test for examining the association in truncated data, treating L as a covariate of T. The hazard rate of T is specified as $\lambda_T(x; L) = \lambda_{T0}(x)\exp(\alpha L)$. Rejection of the null hypothesis $\alpha = 0$ indicates dependence between T and L. When this Cox model truly describes the dependence in a truncated sample, how to estimate the distribution of L has not been studied yet. The IPW estimator is a solution and its asymptotic distribution needs to be explored. Other application of the IPW estimator includes causal inference with truncated data. The form of weighted average of the IPW estimator can be easily extended to incorporate the propensity scores (Anstrom and Tsiatis, 2001).

Appendix A: Asymptotic Distribution of $\widehat{G}^{\mathbf{IPW}}(x)$

Let (a_F, b_F) be the interior of the convex support of F, where $a_F = \inf\{x : F(x) > 0\}$ and $b_F = \sup\{x : F(x) < 1\}$. a_G and b_G are similarly defined for G. The distribution functions F and G are identifiable if $a_G < b_F$.

We first consider the asymptotic distribution of the IPW estimator $\widehat{G}^{\mathrm{IPW}}(x)$. In the following context, $\approx$ denotes asymptotic equivalence.

$$\begin{aligned}
&\sqrt{n}[\widehat{G}^{\mathrm{IPW}}(x) - G(x)] \\
&= \sqrt{n}\left[\widehat{\beta}_3 \int_0^x [1 - \widehat{F}(u-)]^{-1} d\widehat{G}^*(u) - \beta \int_0^x [1 - F(u-)]^{-1} dG^*(u)\right]
\end{aligned}$$

$$
\begin{aligned}
&\approx \sqrt{n}(\widehat{\beta}_3 - \beta)\int_0^x [1-F(u-)]^{-1} dG^*(u) \\
&\quad + \beta \int_0^x [1-F(u-)]^{-1} d\left[\sqrt{n}\left(\widehat{G}^*(u) - G^*(u)\right)\right] \\
&\quad + \beta \int_0^x \sqrt{n}\left(\frac{1}{1-\widehat{F}(u-)} - \frac{1}{1-F(u-)}\right) dG^*(u).
\end{aligned}
$$

The asymptotic distribution of $\widehat{\beta}_3$ was studied by Keiding and Gill (1990, Section 6). Using the delta method, one can show that

$$
\begin{aligned}
\sqrt{n}(\widehat{\beta}_3 - \beta) \approx -\beta^2 \Bigg[&\int_0^\infty [1-F(u-)]^{-1} d\left[\sqrt{n}\left(\widehat{G}^*(u) - G^*(u)\right)\right] \\
&+ \int_0^\infty \sqrt{n}\left(\frac{1}{1-\widehat{F}(u-)} - \frac{1}{1-F(u-)}\right) dG^*(u) \Bigg].
\end{aligned}
$$

Using the above result and noting that $\beta \int_0^x [1-F(s-)]^{-1} dG^*(s) = G(x)$, one will have

$$
\sqrt{n}[\widehat{G}^{\mathrm{IPW}}(x) - G(x)] \approx W_{L,1}(x) + W_{L,2}(x),
$$

where

$$
\begin{aligned}
W_{L,1}(x) &= \beta \int_0^x [1-F(u-)]^{-1} d\left[\sqrt{n}\left(\widehat{G}^*(u) - G^*(u)\right)\right] \\
&\quad - \beta G(x) \int_0^\infty [1-F(u-)]^{-1} d\left[\sqrt{n}\left(\widehat{G}^*(u) - G^*(u)\right)\right], \\
W_{L,2}(x) &= \beta \int_0^x \sqrt{n}\left(\frac{1}{1-\widehat{F}(u-)} - \frac{1}{1-F(u-)}\right) dG^*(u) \\
&\quad - \beta G(x) \int_0^\infty \sqrt{n}\left(\frac{1}{1-\widehat{F}(u-)} - \frac{1}{1-F(u-)}\right) dG^*(u).
\end{aligned}
$$

In Keiding and Gill's Markov-process-based truncation model, $\widehat{\Lambda}_1 - \Lambda_1$ and $\widehat{\Phi} - \Phi$ are regularly assumed to be orthogonal martingales. As a consequence, $W_{L,1}(x)$ and $W_{L,2}(x)$ are also orthogonal and their covariance is zero. It follows that the asymptotic variance of $\widehat{G}^{\mathrm{IPW}}(x) - G(x)$ is the sum of the variances of $W_{L,1}(x)$ and $W_{L,2}(x)$. Since $\widehat{G}^*(x)$ is the empirical estimator of $G^*(x)$, $W_{L,1}(x)$ has mean zero and the variance is given by

$$
\beta^2 \left[\int_0^x [1-F(u-)]^{-2} dG^*(u) - \left(\int_0^x [1-F(u-)]^{-1} dG^*(u) \right)^2 \right]
$$

$$
\begin{aligned}
&+\beta^2 G(x)^2\left[\int_0^\infty [1-F(u-)]^{-2}dG^*(u)-\left(\int_0^\infty [1-F(u-)]^{-1}dG^*(u)\right)^2\right]\\
&-2\beta^2 G(x)\left[\int_0^x [1-F(u-)]^{-2}dG^*(u)\right.\\
&\left.-\int_0^x [1-F(u-)]^{-1}dG^*(u)\int_0^\infty [1-F(u-)]^{-1}dG^*(u)\right].
\end{aligned}
$$

The above expression can be simplified as

$$
\begin{aligned}
&\beta\int_0^x [1-F(u-)]^{-1}dG(u)+\beta G(x)^2\int_0^\infty [1-F(u-)]^{-1}dG(u)\\
&-2\beta G(x)\int_0^x [1-F(u-)]^{-1}dG(u).
\end{aligned}
$$

The central limit theorem leads to the limiting normal distribution of $W_{L,1}(x)$.

The martingale of the counting process $N_{T,i}(x)$ should be clarified in order to study the limiting distribution of $W_{L,2}(x)$. $N_{T,i}(x)$ has the intensity process $Y_i(x)\phi(x)$. Let $M_{T,i}(x) = N_{T,i}(x) - \int_0^x Y_i(u)\phi(u)du$ and $M_{T,i}(x)$ is the counting process martingale. Let $M_T(u) = \sum_{i=1}^n M_{T,i}(u)$. It is known that $\widehat{\Phi}(x) - \Phi(x) = \int_0^x \{dM_T(u)/Y(u)\}$ with the predictable variation process $\int_0^x \{\phi(u)du/Y(u)\}$. The following steps explain the variance of $W_{L,2}(x)$. First, using the delta method,

$$
\begin{aligned}
W_{L,2}(x) &\approx \beta\int_0^x \frac{1}{1-F(u-)}\sqrt{n}[\widehat{\Phi}(u)-\Phi(u)]dG^*(u)\\
&\quad-\beta G(x)\int_0^\infty \frac{1}{1-F(u-)}\sqrt{n}[\widehat{\Phi}(u)-\Phi(u)]dG^*(u)\\
&=\int_0^x \sqrt{n}[\widehat{\Phi}(u)-\Phi(u)]dG(u)-G(x)\int_0^\infty \sqrt{n}[\widehat{\Phi}(u)-\Phi(u)]dG(u).
\end{aligned}
$$

Integration by parts leads to the expression

$$
\begin{aligned}
&=\sqrt{n}\left\{G(x)(\widehat{\Phi}(x)-\Phi(x))-\int_0^x G(u)d[\widehat{\Phi}(u)-\Phi(u)]\right.\\
&\quad\left.-G(x)(\widehat{\Phi}(\infty)-\Phi(\infty))+G(x)\int_0^\infty G(u)d[\widehat{\Phi}(u)-\Phi(u)]\right\}\\
&=\sqrt{n}\int_0^x G(u)(G(x)-1)\frac{dM_T(u)}{Y(u)}+\sqrt{n}\int_x^\infty G(x)(G(u)-1)\frac{dM_T(u)}{Y(u)}.
\end{aligned}
$$

Based on the martingale central limit theorem, $W_{L,2}(x)$ converges in distribution to a zero-mean normal random variable, and the variance is given by

$$
\int_0^x G(u)^2(1-G(x))^2\frac{\phi(u)du}{y(u)}+\int_x^\infty G(x)^2(1-G(u))^2\frac{\phi(u)du}{y(u)}.
$$

Appendix B: Asymptotic Distribution of $\widehat{S}^{\mathbf{IPW}}(x)$

$$\widehat{S}^{\mathrm{IPW}}(x)=\widehat{\beta}_4\int_{\infty}^{x}\widehat{G}(u)^{-1}d\widehat{S}^*(u),\quad \widehat{\beta}_4=\left(\int_{\infty}^{0}\widehat{G}(u)^{-1}d\widehat{S}^*(u)\right)^{-1}.$$

Similarly, it can be shown that $\sqrt{n}[\widehat{S}^{\mathrm{IPW}}(x)-S(x)]=W_{T,1}(x)+W_{T,2}(x)$, where

$$\begin{aligned}W_{T,1}(x)&=\beta\int_{\infty}^{x}G(u)^{-1}d\{\sqrt{n}[\widehat{S}^*(u)-S^*(u)]\}\\&\quad-\beta S(x)\int_{\infty}^{0}G(u)^{-1}d\{\sqrt{n}[\widehat{S}^*(u)-S^*(u)]\},\\W_{T,2}(x)&=\beta\int_{\infty}^{x}\sqrt{n}[\widehat{G}(u)^{-1}-G(u)^{-1}]dS^*(u)\\&\quad-\beta S(x)\int_{\infty}^{0}\sqrt{n}[\widehat{G}(u)^{-1}-G(u)^{-1}]dS^*(u).\end{aligned}$$

Since $\widehat{\Lambda}_2^*(x)-\Lambda_2^*(x)$ and $\widehat{\Gamma}^*(x)-\Gamma^*(x)$ are orthogonal martingales, $W_{T,1}(x)$ and $W_{T,2}(x)$ have the orthogonal relationship. Therefore, the asymptotic variance of $\widehat{S}^{\mathrm{IPW}}(x)-S(x)$ can be obtained by adding up the variances of $W_{T,1}(x)$ and $W_{T,2}(x)$. Since $\widehat{S}^*(x)$ is the empirical estimator, the variance of $W_{T,1}(x)$ is given by

$$\beta\int_{\infty}^{x}G(u)^{-1}dS(u)+\beta S(x)^2\int_{\infty}^{0}G(u)^{-1}dS(u)-2\beta S(x)\int_{\infty}^{x}G(u)^{-1}dS(u).$$

The counting process $N_{L,i}^*(x)$ has the intensity process $Y_i(x)\gamma^*(x)$. One can define the martingale, $M_{L,i}^*(x)=N_{L,i}^*(x)-\int_{\infty}^{x}Y_i(u)\gamma^*(u)du$ and let $M_L^*(x)=\sum_{i=1}^{n}M_{L,i}^*(x)$. It can be shown that $\widehat{\Gamma}^*(u)-\Gamma^*(u)=\int_{\infty}^{x}\{dM_L^*(u)/Y(u)\}$ with the predictable variation process $\int_{\infty}^{x}\{\gamma^*(u)du/Y(u)\}$. Applying the delta method on $W_{T,2}(x)$ and noting the result that $\beta\int_{\infty}^{x}G(s)^{-1}dS^*(s)=S(x)$, one will have

$$W_{T,2}(x)\approx\int_{\infty}^{x}\sqrt{n}[\widehat{\Gamma}^*(u)-\Gamma^*(u)]dS(u)-S(x)\int_{\infty}^{0}\sqrt{n}[\widehat{\Gamma}^*(u)-\Gamma^*(u)]dS(u).$$

Integration by parts leads to

$$\begin{aligned}&\sqrt{n}\left\{S(x)[\widehat{\Gamma}^*(x)-\Gamma^*(x)]-\int_{\infty}^{x}S(u)d[\widehat{\Gamma}^*(u)-\Gamma^*(u)]-S(x)[\widehat{\Gamma}^*(0)-\Gamma^*(0)]\right.\\&\left.+S(x)\int_{\infty}^{0}S(u)d[\widehat{\Gamma}^*(u)-\Gamma^*(u)]\right\}\\&=\sqrt{n}\int_{\infty}^{x}S(u)[1-S(x)]\frac{dM_L^*(u)}{Y(u)}+\sqrt{n}\int_{x}^{0}S(x)[1-S(u)]\frac{dM_L^*(u)}{Y(u)}.\end{aligned}$$

Utilizing the martingale central limit theorem, $W_{T,2}(x)$ converges in distribution to a zero-mean normal random variable with the variance

$$\int_{\infty}^{x} S(u)^2[1-S(x)]^2\frac{\gamma^*(u)du}{y(u)} + \int_{x}^{0} S(x)^2[1-S(u)]^2\frac{\gamma^*(u)du}{y(u)}.$$

References

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer, New York.

Anstrom, K. J. and Tsiatis, A. A. (2001). Using propensity scores to estimate causal treatment effects with censored time-lagged data. *Biometrics* **57**, 1207-1218.

Bilker, W. B. and Wang, M. C. (1996). A semiparametric extension of the Mann-Whitney test for randomly truncated data. *Biometrics* **52**, 10-20.

Chao, M. T. (1987). Influence curves for randomly truncated data. *Biometrika* **74**, 426-429.

He, S. and Yang, G. L. (1998). Estimation of the truncation probability in the random truncation model. *The Annals of Statistics* **26**, 1011-1027.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663-685.

Jones, M. P. and Crowley, J. (1992). Nonparametric tests of the Markov model for survival data. *Biometrika* **79**, 513-522.

Kalbfleisch, J. D. and Lawless, J. F. (1989). Inference based on retrospective ascertainment: an analysis of the data on transfusion-related AIDS. *Journal of the American Statistical Association* **84**, 360-372.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **84**, 360-372.

Keiding, N. and Gill, R. D. (1990). Random truncation models and Markov processes. *The Annals of Statistics* **18**, 582-602.

Lagakos, S. W., Barraj, L. M. and De Gruttola, V. (1988). Nonparametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, 515-523.

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices of the Royal Astronomical Society* **155**, 95-118.

Shen, P. S. (2003). The product-limit estimate as an inverse-probability-weighted average. *Communications in Statistics - Theory and Methods* **32**, 1119-1133.

Shen, P. S. (2005). Estimation of the truncation probability with left-truncated and right-censored data. *Journal of Nonparametric Statistics* **17**, 957-969.

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

Vardi, Y. (1985). Empirical distributions in selection bias models. *The Annals of Statistics* **13**, 178-203.

Wang, M. C., Jewell, N. P. and Tsai, W. Y. (1986). Asymptotic properties of the product limit estimate under random truncation. *The Annals of Statistics* **14**, 1597-1605.

Wang, M. C. (1989). A semiparametric model for randomly truncated data. *Journal of the American Statistical Association* **84**, 742-748.

Wang, M. C. (1991). Nonparametric estimation from cross-sectional survival data. *Journal of the American Statistical Association* **86**, 130-143.

Woodroofe, M. (1985). Estimating a distribution function with truncated data. *The Annals of Statistics* **13**, 163-177.

Received March 13, 2012; accepted May 10, 2012.

Xu Zhang
Department of Mathematics and Statistics
Georgia State University
Atlanta, GA 30303, USA
xzhang19@gsu.edu