An augmented inverse probability weighted survival function estimator

Sundarraman Subramanian* & Dipankar Bandyopadhyay†

Abstract

We analyze an augmented inverse probability of non-missingness weighted estimator of a survival function for a missing censoring indicator model, in the absence and presence of left truncation. The estimator improves upon its precursor but is still not the best in terms of achieving minimal asymptotic variance.

KEY WORDS: Asymptotic variance; Dikta semiparametric estimator; Empirical estimator; Gaussian process; Gradient vector; Missing at random.

1 Introduction

In this article, we focus on a missing censoring indicator (MCI) model, with observed data $\{(X_i, \xi_i, \sigma_i)_{1 \leq i \leq n}\}$, where $X_i = \min(T_i, C_i)$, T is a lifetime of interest, C is an independent censoring variable, $\xi_i = 1$ when $\delta_i = I(T_i \leq C_i)$ is observed and is 0 otherwise, and $\sigma_i = \xi_i \delta_i$. The censoring indicators δ_i are assumed to be missing at random (MAR), which implies that $P(\xi = 1|X = x, \delta = d) = P(\xi = 1|X = x) = \pi(x)$ (van der Laan and McKeague, 1998). Writing $p(x) = P(\delta = 1|X = x)$, MAR means that ξ and δ are conditionally independent given X: $P(\sigma = 1|X = x) = \pi(x)p(x)$. We investigate semiparametric estimation of $S(t)$, the survival function of T, using an augmented inverse probability weighted (AIPW) scheme.

*Center for Applied Mathematics and Statistics, Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 07102, United States

†Department of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina, Charleston, SC 29425, United States

Inverse probability weighting (IPW) has gained much currency of late. Horvitz and Thompson (1952) introduced IPW in the context of sample survey, while Koul, Susarla, and van Ryzin (1981) first employed the weighting scheme for censored regression. More recently, Robins and Rotnitzky (1992), Satten and Datta (2001, 2004), and Rotnitzky and Robins (2005), among others, have also introduced IPW-based analysis in survival and related settings. However, the IPW estimators are not consistent under misspecification of the weighting probability, leading to the investigation of AIPW estimators (Scharfstein, Rotnitzky and Robins 1999). The key element of such AIPW estimators is that they are doubly robust in the sense that they remain consistent if any one of $p(x)$ and $\pi(x)$ is specified correctly. Robins, Rotnitzky and van der Laan (2000), among others, have studied double robust estimators. See also Tsiatis (2006) and van der Laan and Robins (2003).

2 Estimation

We first introduce some preliminary notation and present a brief review of two semiparametric estimators of $S(t)$ when there are MCIs. These are the Dikta estimator (Subramanian, 2004a) and the IPW estimator. The latter was investigated recently for a left truncated version of the MCI model (Subramanian and Bandyopadhyay, 2008).

2.1 Preliminaries

Let $H(t)$ denote the distribution function of X and $\hat{H}(t)$ denote its empirical estimator. We focus on estimation over $[0, \tau]$, where $H(\tau) < 1$. Let $Q(t) = P(X \leq t, \delta = 1)$ denote the subdistribution function corresponding to uncensored failures. Its empirical estimator, however, is inapplicable when there are MCIs. To address this, specify a model for $p(t)$.

That is, take $p(t) = p(t, \theta)$, where θ is a k dimensional unknown Euclidean parameter. The Dikta semiparametric estimator of $Q(t)$ is given by

$$\hat{Q}_D(t) \;=\; \int_0^t \hat{p}(s) d\hat{H}(s), \tag{1}$$

where $\hat{p}(s) = p(s, \hat{\theta})$, and $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ. The MLE $\hat{\theta}$ can be obtained by maximizing a likelihood based on the complete cases $\xi_i = 1$, see, for example, Subramanian (2004a). Let θ_0 denote the true value of θ and define $p_0(s) = p(s, \theta_0)$. For $r = 1, \ldots, k$, write $p_r(s, \theta_0) = \partial p(s, \theta)/\partial \theta_r|_{\theta=\theta_0}$ and let $P_0(s) = [p_1(s, \theta_0), \ldots, p_k(s, \theta_0)]^T$ denote the gradient vector evaluated at θ_0. Let $I_0 = E[\pi(X) P_0(X) P_0^T(X)/(p_0(X)(1 - p_0(X)))]$. Also, let $\alpha(s, t) = P_0^T(s) I_0^{-1} P_0(t)$, and define

$$\bar{B}(t) \;\equiv\; \frac{1}{n} \sum_{i=1}^n \tilde{B}_i(t) \;=\; \frac{1}{n} \sum_{i=1}^n \frac{\xi_i \left(\delta_i - p_0(X_i)\right)}{p_0(X_i)(1 - p_0(X_i))} \alpha(t, X_i). \tag{2}$$

Let M be such that $\int_0^\infty M(s) dH(s) < \infty$. From Lemma 3.5 of Dikta (1998), we have the following representation, assuming that the model for $p(t)$ is correctly specified:

$$n^{1/2} \left\{\hat{p}(t) - p_0(t)\right\} \;=\; n^{1/2} \bar{B}(t) + R_n(t), \tag{3}$$

where, uniformly for $0 \le t \le \tau$, the term $n^{-1/2}|R_n(t)|$ is bounded above by

$$M(t) \sum_{1 \le r, s \le k} (\hat{\theta}_r - \theta_{0,r})(\hat{\theta}_s - \theta_{0,s}) \;=\; M(t) \left(O_p(n^{-1/2}) + o_p(n^{-1/2})\right)^2 \;= M(t) O_p(n^{-1}).$$

Under the regularity conditions given by Dikta (1998), which we denote henceforth by $\mathbf{R}$, it can be shown that (Subramanian, 2004a) $n^{1/2}(\hat{Q}_D(t) - Q(t))$ is asymptotically linear with influence function given by $A(t) + B(t)$, where

$$A(t) \;=\; p_0(X) I(X \le t) - Q(t), \tag{4}$$

$$B(t) \;=\; \int_0^t \tilde{B}(s) dH(s) \;=\; \frac{\xi \left(\delta - p_0(X)\right)}{p_0(X)(1 - p_0(X))} \int_0^t \alpha(s, X) dH(s). \tag{5}$$

It follows that the process $n^{1/2}(\hat{Q}_D(t) - Q(t))$ converges weakly to a zero-mean Gaussian process whose variance at t [the second moment of $A(t)+B(t)$], the benchmark for comparing competing estimators of $Q(t)$ in this article, is given by

$$V_D(t) \;=\; \int_0^t p_0(s)dQ(s) - Q^2(t) + \int_0^t \int_0^t \alpha(u,v)dH(u)dH(v). \tag{6}$$

Alternatively, specify $\pi(t) = \pi(t,\gamma)$ parametrically, where γ is an unknown m dimensional parameter. Set $\hat{\pi}(s) = \pi(s,\hat{\gamma})$ and $\hat{\gamma}$ is the MLE of γ. Let $\hat{W}(t)$ denote the empirical estimator of $W(t) = P(X \le t, \xi = 1, \sigma = 1)$. From an IPW representation for $Q(t)$ (e.g., Subramanian, 2006), the IPW estimator takes the form

$$\hat{Q}_I(t) \;=\; \int_0^t \frac{1}{\hat{\pi}(s)} d\hat{W}(s). \tag{7}$$

Denote the gradient vector for this scenario by $\breve{P}(s,\theta)$, and by $\breve{P}_0(s)$ when it is evaluated at the true value γ_0. Write $\beta(s,t) = \breve{P}_0^T(s)J_0^{-1}\breve{P}_0(t)$, where the information matrix $J_0 = E[\breve{P}_0(X)\breve{P}_0^T(X)/(\pi_0(X)(1-\pi_0(X)))]$. Define

$$\bar{V}(t) \equiv \frac{1}{n}\sum_{i=1}^n V_i(t) \;=\; \frac{1}{n}\sum_{i=1}^n \frac{(\xi_i - \pi_0(X_i))}{\pi_0(X_i)(1-\pi_0(X_i)}\beta(t,X_i). \tag{8}$$

Analogous to Eq. (3), assuming that the model for $\pi(t)$ is correctly specified, we have that

$$n^{1/2}(\hat{\pi}(t) - \pi_0(t)) \;=\; n^{1/2}\bar{V}(t) + R_n(t). \tag{9}$$

Assuming that $\pi(t)$ is bounded away from 0, uniformly for $t \in [0,\tau]$, Eq. (8) and Eq. (9) and some basic analysis imply that $n^{1/2}(\hat{Q}_I(t) - Q(t))$ is asymptotically linear with influence function given by $\breve{A}(t) + \breve{B}(t)$, where

$$\breve{A}(t) \;=\; \frac{I(X \le t, \xi = 1, \sigma = 1)}{\pi_0(X)} - Q(t), \tag{10}$$

$$\breve{B}(t) \;=\; -\int_0^t \frac{V(s)}{\pi_0(s)}dQ(s) \;=\; \frac{-(\xi - \pi_0(X))}{\pi_0(X)(1-\pi_0(X))}\int_0^t \frac{\beta(s,X)}{\pi_0(s)}dQ(s). \tag{11}$$

It follows that the process $n^{1/2}(\hat{Q}_I(t) - Q(t))$ converges weakly to a zero-mean Gaussian process whose variance at t [the second moment of $\breve{A}(t) + \breve{B}(t)$] is given by

$$V_I(t) = \int_0^t \frac{1}{\pi_0(s)} dQ(s) - Q^2(t) - \int_0^t \int_0^t \frac{\beta(u,v)}{\pi_0(u)\pi_0(v)} dQ(u)dQ(v). \tag{12}$$

Furthermore, it can be shown that $V_I(t) \geq V_D(t)$, see, for example, Subramanian and Bandyopadhyay (2008), where an analogous result is derived for a left truncated version.

Both estimators $\hat{Q}_D(t)$ and $\hat{Q}_I(t)$ work under the tacit assumption that the parametric models for $p(t)$ and $\pi(t)$ respectively are correctly specified, and are inconsistent when the putative models are ill-specified. The AIPW estimator of $Q(t)$ that we analyze in the next subsection is a double robust estimator (Scharfstein et. al., 1999, Tsiatis, 2006), in that it retains consistency when either or both of $p(t)$ and $\pi(t)$ is specified correctly.

2.2 AIPW subdistribution function estimator

The AIPW estimator of $Q(t)$ takes the form (e.g., Tsiatis, 2006)

$$\hat{Q}_{AI}(t) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{\xi_i \delta_i}{\hat{\pi}(X_i)} + \left(1 - \frac{\xi_i}{\hat{\pi}(X_i)}\right) \hat{p}(X_i) \right\} I(X_i \leq t). \tag{13}$$

We analyze $\hat{Q}_{AI}(t)$ for three different scenarios. Scenario 1 represents the case "only $p(t)$ correctly specified", Scenario 2 the case "only $\pi(t)$ correctly specified", and Scenario 3 represents "both $p(t)$ and $\pi(t)$ correctly specified". Clearly, $\hat{Q}_{AI}(t) = \hat{Q}_D(t) + T_n(t)$, where

$$T_n(t) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{\xi_i}{\hat{\pi}(X_i)} (\delta_i - \hat{p}(X_i)) \right\} I(X_i \leq t), \tag{14}$$

and $\hat{Q}_D(t)$ is defined by Eq. (1). We analyze this version of $\hat{Q}_{AI}(t)$ for Scenario 1. On the other hand, we can also write $\hat{Q}_{AI}(t) = \hat{Q}_I(t) + \breve{T}_n(t)$, where $\hat{Q}_I(t)$ is defined by Eq. (7) and

$$\breve{T}_n(t) = \frac{1}{n} \sum_{i=1}^{n} \left(1 - \frac{\xi_i}{\hat{\pi}(X_i)}\right) \hat{p}(X_i) I(X_i \leq t), \tag{15}$$

which is the version that we will analyze for Scenario 2.

We now state some basic assumptions, introduced by Kouassi and Singh (1997). We denote them henceforth by **RP**. When the model for $p(t)$ is misspecified (Scenario 2), we assume that there exists $\theta^* \in I\!R^k$ such that, with $p^*(t) = p(t, \theta^*)$, the following hold:

$$E\left(\hat{\theta} - \theta^*\right) = o\left(n^{-1/2}\right), \tag{16}$$

$$\hat{\theta} - \theta^* = O_p\left(n^{-1/2}\right), \tag{17}$$

$$|p^*(t) - p_0(t)| < \infty. \tag{18}$$

Eq. (16) and Eq. (17) are introduced to ensure that the MLEs $\hat{\theta}$ and $\hat{p}(t)$ converge at the standard $n^{1/2}$ rates even when the parametric model is misspecified. Eq. (18) requires that the bias of $\hat{p}(t)$ is bounded when $p(t, \theta)$ does not hold. Analogous conditions for the MLEs $\hat{\gamma}$, $\hat{\pi}(t)$, and $\pi^*(t) = \pi(t, \gamma^*)$ are also assumed (denoted by **RPI** henceforth). Also, it will be assumed that $\hat{p}(t) \xrightarrow{P} p^*(t)$ and $\hat{\pi}(t) \xrightarrow{P} \pi^*(t)$ uniformly for $t \in [0, \tau]$. The last conditions of uniformity (which are also included in **RP** or **RPI**) can be enforced through appropriate boundedness assumptions, see Eq. (3) and the equation following it.

For Scenario 1, $\hat{p}(t) \xrightarrow{P} p_0(t)$ and $\hat{Q}_D(t) \xrightarrow{P} Q(t)$, uniformly for $t \in [0, \tau]$, see Eq. (3) or Dikta (1998). We can write [cf. Eq. (14)] $T_n(t) = T_{n,1}(t) + T_{n,2}(t) + T_{n,3}(t)$ where

$$T_{n,1}(t) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{\xi_i}{\pi^*(X_i)} \left(\delta_i - p_0(X_i)\right) \right\} I(X_i \le t) \tag{19}$$

$$T_{n,2}(t) = -\frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{\xi_i}{\pi^*(X_i)} \left(\hat{p}(X_i) - p_0(X_i)\right) \right\} I(X_i \le t) \tag{20}$$

$$T_{n,3}(t) = -\frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{\xi_i (\hat{\pi}(X_i) - \pi^*(X_i))}{\hat{\pi}(X_i)\pi^*(X_i)} \left(\delta_i - \hat{p}(X_i)\right) \right\} I(X_i \le t). \tag{21}$$

By the strong law large numbers $T_{n,1}(t) \xrightarrow{\text{a.s.}} 0$. Next, $|T_{n,2}(t)| = O_p\left(\sup_{0 \le t \le \tau} |\hat{p}(t) - p_0(t)|\right)$,

and therefore $T_{n,2}(t) \xrightarrow{P} 0$. Also, $|T_{n,3}(t)| = O_p\left(\sup_{0\le t\le\tau} |\hat{\pi}(t) - \pi^*(t)|\right) \xrightarrow{P} 0$. Therefore, when only $p(t)$ is specified correctly, it is clear that $\hat{Q}_{AI}(t) = \hat{Q}_D(t) + T_n(t)$ retains consistency. Note, however, that $n^{1/2}\left(\hat{Q}_{AI}(t) - Q(t)\right)$ and $n^{1/2}\left(\hat{Q}_D(t) - Q(t)\right)$ do not have the same limiting distribution, since $n^{1/2}T_n(t)$ is not $o_p(1)$, see Theorem 1.

For Scenario 2, we can show that $\hat{\pi}(t) \xrightarrow{P} \pi_0(t)$ and $\hat{Q}_I(t) \xrightarrow{P} Q(t)$, uniformly for $t \in [0, \tau]$. We can write [cf. Eq. (15)] $\breve{T}_n(t) = \breve{T}_{n,1}(t) + \breve{T}_{n,2}(t) + \breve{T}_{n,3}(t)$ where

$$\breve{T}_{n,1}(t) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\xi_i}{\pi_0(X_i)}\right) p^*(X_i) I(X_i \le t) \tag{22}$$

$$\breve{T}_{n,2}(t) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\xi_i(\hat{\pi}(X_i) - \pi_0(X_i))}{\hat{\pi}(X_i)\pi_0(X_i)}\hat{p}(X_i)\right\} I(X_i \le t) \tag{23}$$

$$\breve{T}_{n,3}(t) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\xi_i}{\pi_0(X_i)}\right)\left(\hat{p}(X_i) - p^*(X_i)\right) I(X_i \le t). \tag{24}$$

As before, we can show that $\breve{T}_n(t) \xrightarrow{P} 0$. Therefore, for this case also, $\hat{Q}_{AI}(t) = \hat{Q}_I(t) + \breve{T}_n(t)$ retains consistency. Again, it may be noted that $n^{1/2}\left(\hat{Q}_{AI}(t) - Q(t)\right)$ and $n^{1/2}\left(\hat{Q}_I(t) - Q(t)\right)$ do not have the same limiting distribution, since $n^{1/2}\breve{T}_n(t)$ is not $o_p(1)$, see Theorem 2.

Let $\bar{\pi}(t) = 1 - \pi_0(t)/\pi^*(t)$, which equals 0 when $\pi(t)$ is correctly specified. Along with Eq. (4) and Eq. (5), we also define two other quantities

$$C(t) = \left\{\frac{\xi}{\pi^*(X)}\left(\delta - p_0(X)\right)\right\} I(X \le t), \tag{25}$$

$$D(t) = -\frac{\xi\left(\delta - p_0(X)\right)}{p_0(X)(1 - p_0(X))}\int_0^t \frac{\pi_0(s)}{\pi^*(s)}\alpha(s, X)dH(s), \tag{26}$$

as well as the following variance function $V_1(t)$:

$$V_1(t) = \int_0^t p_0(s)dQ(s) - Q^2(t) + \int_0^t\int_0^t \bar{\pi}(u)\alpha(u,v)dH(u)dH(v) + \int_0^t \frac{\pi_0(u)}{(\pi^*(u))^2}(1 - p_0(u))dQ(u). \tag{27}$$

The following theorem describes the asymptotic distribution of $\hat{Q}_{AI}(t)$ for Scenario 1.

Theorem 1 *Suppose that the regularity conditions* **R** *and* **RPI** *hold and that* $\pi_0(t)$ *is bounded away from 0, uniformly for* $t \in [0, \tau]$. *The process* $n^{1/2}\left(\hat{Q}_{AI}(t) - Q(t)\right)$ *converges weakly in* $D[0, \tau]$ *to a zero-mean Gaussian process* Z_1 *with variance function* $V_1(t)$ *defined by Eq. (27).*

Proof We know that $n^{1/2}(\hat{Q}_D(t) - Q(t))$ is asymptotically linear with influence function $A(t) + B(t)$. From Eq. (19), $n^{1/2}\, T_{n,1}(t)$ is asymptotically linear with influence function $C(t)$. Let $\hat{\tilde{W}}(t)$ denote the empirical estimator of $\tilde{W}(t) = P(X \leq t, \xi = 1)$. For Eq. (20), analogous to calculations in Subramanian and Bean (2008), we can show that, uniformly for $t \in [0, \tau]$,

$$T_{n,2}(t) \;=\; -\int_0^t \frac{\hat{p}(s) - p_0(s)}{\pi^*(s)} d\hat{\tilde{W}}(s) \;=\; -\int_0^t \frac{\hat{p}(s) - p_0(s)}{\pi^*(s)} d\tilde{W}(s) + o_p(n^{-1/2}).$$

It readily follows, from Eq. (3) and then Eq. (2), that $n^{1/2}\, T_{n,2}(t)$ is also asymptotically linear with influence function $D(t)$. Next, note that $dW(s) = p_0(s) d\tilde{W}(s)$. It can be shown that, uniformly for $t \in [0, \tau]$, $T_{n,3}(t)$ defined by Eq. (21) is asymptotically negligible:

$$\begin{aligned} T_{n,3}(t) \;&=\; -\int_0^t \frac{\hat{\pi}(s) - \pi^*(s)}{\hat{\pi}(s)\pi^*(s)} \left\{ d\hat{W}(s) - \hat{p}(s) d\hat{\tilde{W}}(s) \right\} \\ &=\; -\int_0^t \frac{\hat{\pi}(s) - \pi^*(s)}{(\pi^*(s))^2} \left\{ dW(s) - p_0(s) d\tilde{W}(s) \right\} + o_p(n^{-1/2}) \;=\; o_p(n^{-1/2}). \end{aligned}$$

It follows that $n^{1/2}(\hat{Q}_{AI}(t) - Q(t))$ is asymptotically linear with influence function $A(t) + B(t) + C(t) + D(t)$ [cf. Eqs. (4), (5), (25), and (26)]. After some routine variance calculations, it can be shown that the asymptotic variance is given by $V_1(t)$ defined by Eq. (27). □

We next focus on Scenario 2. Let $d_p(t) = p^*(t) - p_0(t)$, which is 0 when $p(t)$ is correctly specified. Along with Eq. (10) and Eq. (11), we also define two other quantities:

$$\breve{C}(t) \;=\; \left(1 - \frac{\xi}{\pi_0(X)}\right) p^*(X) I(X \leq t) \tag{28}$$

$$\breve{D}(t) \;=\; \frac{\xi - \pi_0(X)}{\pi_0(X)(1 - \pi_0(X))} \int_0^t \frac{p^*(s)}{\pi_0(s)} \beta(s, X) dH(s), \tag{29}$$

as well as the following variance function $V_2(t)$:

$$\begin{aligned} V_2(t) &= \int_0^t \frac{1}{\pi_0(u)} dQ(u) - Q^2(t) - \int_0^t \frac{1-\pi_0(u)}{\pi_0(u)} p^*(u) dH(u) \\ &\quad - \int_0^t \int_0^t \frac{d_p(u)}{\pi_0(u)} \frac{d_p(v)}{\pi_0(v)} \beta(u,v) dH(v) dH(u). \end{aligned} \tag{30}$$

The following theorem describes the asymptotic distribution of $\hat{Q}_{AI}(t)$ for Scenario 2.

Theorem 2 *Suppose that the regularity conditions* **R** *and* **RP** *hold and that $\pi_0(t)$ is bounded away from 0, uniformly for $t \in [0, \tau]$. The process $n^{1/2}\left(\hat{Q}_{AI}(t) - Q(t)\right)$ converges weakly in $D[0, \tau]$ to a zero-mean Gaussian process Z_2 with variance function $V_2(t)$ defined by Eq. (30).*

Proof We know that $n^{1/2}(\hat{Q}_I(t) - Q(t))$ is asymptotically linear with influence function $\breve{A}(t) + \breve{B}(t)$. From Eq. (22), $n^{1/2}\breve{T}_{n,1}(t)$ is asymptotically linear with influence function $\breve{C}(t)$. For $\breve{T}_{n,2}(t)$ defined by Eq. (23), we can show that, uniformly for $t \in [0, \tau]$,

$$\breve{T}_{n,2}(t) = \int_0^t \frac{(\hat{\pi}(s) - \pi_0(s))}{\hat{\pi}(s)\pi_0(s)} \hat{p}(s) d\hat{\tilde{W}}(s) = \int_0^t \frac{(\hat{\pi}(s) - \pi_0(s))}{\pi_0^2(s)} p^*(s) d\tilde{W}(s) + o_p(n^{-1/2}).$$

It follows, from Eq. (9) and then Eq. (8), that $n^{1/2}\breve{T}_{n,2}(t)$ is also asymptotically linear with influence function $\breve{D}(t)$. Furthermore, it can be shown that $\breve{T}_{n,3}(t)$, defined by Eq. (24) is $o_p(n^{-1/2})$. It follows that $n^{1/2}(\hat{Q}_{AI}(t) - Q(t))$ is asymptotically linear with influence function $\breve{A}(t) + \breve{B}(t) + \breve{C}(t) + \breve{D}(t)$ [cf. Eqs. (10), (11), (28), and (29)]. After some variance calculations, it can be shown that the asymptotic variance is given by $V_2(t)$ defined by Eq. (30). □

It is difficult to compare the variances $V_1(t)$ and $V_2(t)$ with $V_D(t)$ [Eq. (6)], or for that matter $V_I(t)$ [Eq. (12)]. However, for Scenario 3, a definitive conclusion can be made. Define

$$V(t) = Q(t)(1 - Q(t)) + \int_0^t (1 - p_0(u)) \frac{1 - \pi_0(u)}{\pi_0(u)} dQ(u). \tag{31}$$

The following theorem describes the asymptotic distribution of $\hat{Q}_{AI}(t)$ for Scenario 3.

Theorem 3 *Suppose that the regularity condition* **R** *holds and that* $\pi_0(t)$ *is bounded away from 0, uniformly for* $t \in [0, \tau]$*. The process* $n^{1/2}\left(\hat{Q}_{AI}(t) - Q(t)\right)$ *converges weakly in* $D[0, \tau]$ *to a zero-mean Gaussian process* Z *with variance function* $V(t)$ *defined by Eq. (31).*

Proof Follow proof of Th. 1 with $\pi^*(t) = \pi_0(t)$, or follow proof of Th. 2 with $p^*(t) = p_0(t)$. In each case, the asymptotic variance reduces to $V(t)$, since $\bar{\pi}(t) = 0$ and $d_p(t) = 0$. □

Note that when π and p are both correctly specified (Scenario 3), the variance of the AIPW estimator equals the information bound for estimating $Q(t)$ (Subramanian 2004a), which was in fact derived from the efficient influence function for estimating $S(t)$ (van der Laan and McKeague, 1998). However, $V(t) \geq V_D(t)$, see Subramanian (2004a). In particular, we have the order relation $V_D(t) \leq V(t) \leq V_I(t)$. Thus, although the AIPW estimator $\hat{Q}_{AI}(t)$ improves upon its IPW precursor, it is still inferior to the Dikta estimator $\hat{Q}_D(t)$.

2.3 AIPW survival function estimator

We now briefly describe the AIPW estimator of $S(t)$. This estimator results from employing a series of compactly differentiable mappings of the basic estimators $\hat{Q}_{AI}(t)$ and $1 - \hat{H}(t)$ (Gill and Johansen, 1990). Specifically, the estimator takes the form

$$\hat{S}_{AI}(t) \quad = \quad \prod_{0 \leq s \leq t} \left(1 - \frac{d\hat{Q}(s)}{1 - \hat{H}(s-)}\right). \tag{32}$$

The weak convergence of $\hat{S}_{AI}(t)$ can be derived in a straightforward way through the functional delta method, see p. 1537 of Gill and Johansen (1990). Considering only Scenario 3, $n^{1/2}\{\hat{Q}_{AI}(t) - Q(t), 1 - \hat{H}(t-) - (1 - H(t)\} \xrightarrow{\mathcal{D}} (Z, Z_H)$, in $(D[0, \tau] \times D_-[0, \tau], \| \cdot \|_\infty^\vee)$, as $n \to \infty$, where Z is the limiting Gaussian process defined in the statement of Th. 3. Here, $\| \cdot \|_\infty$ is the supremum norm, and $\| \cdot \|_\infty^\vee$ is the max supremum norm. We then have by

the functional delta method that $n^{1/2}\{\hat{S}_{AI}(t) - S(t)\}$ converges weakly in $(D[0,\tau], \|\cdot\|_\infty)$ to $W^*(t) = -S(t)\int_0^t dW(s)/(1-H(s))$ as $n \to \infty$, where $W(t) = Z(t) - \int_0^t Z_H(s)d\Lambda(s)$. We can then show that the variance function $V^*(t)$ of $W^*(t)$ is given by

$$V^*(t) = S(t)^2\left[\int_0^t \frac{d\Lambda(u)}{1-H(u)} + \int_0^t (1-p_0(u))\frac{1-\pi_0(u)}{\pi_0(u)}\frac{d\Lambda(u)}{1-H(u)}\right], \tag{33}$$

which is the asymptotic variance of the reduced data NPMLE (van der Laan and McKeague, 1998), and of two competing estimators (Subramanian 2004b, 2006) as well. However, $V^*(t) \geq V_D^*(t)$ (Subramanian, 2004a), where

$$V_D^*(t) = S(t)^2\left[\int_0^t \frac{p_0(s)}{1-H(s)}d\Lambda(s) + \int_0^t\int_0^t \frac{\alpha(u,v)}{(1-H(u))(1-H(v))}dH(u)dH(v)\right] \tag{34}$$

is the asymptotic variance of the Dikta semiparametric estimator $\hat{S}_D(t)$. Note that $V_D^*(t)$ can also be derived from Eq. (6) and the functional delta method described above.

2.4 Extension to a left truncated version

The three scenarios investigated above can be analyzed in an analogous way for a left truncated version of the MCI model. For this version, we observe (X, Z, ξ, σ) whenever $X \geq Z$, where Z is a left truncation variable, $\xi = 1$ when δ is observed and 0 otherwise, and $\sigma = \xi\delta$ (Subramanian and Bandyopadhyay, 2008). When $X < Z$, however, nothing is observed. The observed data are $\{X_i, Z_i, \xi_i, \sigma_i\}_{1\leq i\leq n}$, an i.i.d. sample from the distribution of (X, Z, ξ, σ) which is observed i.e., when $(Z_i \leq X_i)$. Let $\tilde{W}(t) = P(Z \leq t \leq X | Z \leq X)$ and $W_1(t) = P(X \leq t, \delta = 1 | Z \leq X)$. The cumulative hazard takes the form $\Lambda(t) = \int_{a_W}^t dW_1(u)/\tilde{W}(u)$, where W denotes the d.f. of X and $a_W = \inf\{x : W(x) > 0\}$. The

asymptotic variance of the AIPW estimator of $S(t)$ for Scenario 3 is [compare with Eq. (33)]

$$V_T(t) \;=\; S(t)^2\left[\int_{a_W}^{t}\frac{d\Lambda(u)}{\tilde{W}(u)}+\int_{a_W}^{t}(1-p_0(u))\frac{1-\pi_0(u)}{\pi_0(u)}\frac{d\Lambda(u)}{\tilde{W}(u)}\right],$$

where $p_0(t) = P(\delta = 1|X = t, Z \leq X)$, and $\pi_0(t) = P(\xi = 1|X = t, Z \leq X)$. The Dikta semiparametric estimator for this version of the MCI model has asymptotic variance

$$\breve{V}_T(t) \;=\; \int_{a_W}^{t}\frac{p_0(u)}{\tilde{W}(u)}d\Lambda(u)+\int_{a_W}^{t}\int_{a_W}^{t}\frac{\alpha(u,v)}{p_0(u)p_0(v)}d\Lambda(u)d\Lambda(v),$$

see Subramanian and Bandyopadhyay (2008). Moreover, we can show that $V_T(t) \geq \breve{V}_T(t)$.

References

Dikta, G., 1998. On semiparametric random censorship models. J. Statist. Plann. Inference 66, 253–279.

Gill, R. D., Johansen, S., 1990. A survey of product-integration with a view toward application in survival analysis. Ann. Statist. 18, 1501–1555.

Horvitz, D.G. , Thompson, D.J., 1952. A generalization of sampling without replacement from a finite universe, J. Ameri. Statist. Assoc. 47, 663–685.

Kouassi, D. A., Singh, J., 1997. A semiparametric approach to hazard estimation with randomly censored observations. J. Ameri. Statist. Assoc. 92, 1351–1355.

Koul, H., Susarla, V., Van Ryzin, J., 1981. Regression analysis of randomly right censored data. Ann. Statist. 9, 1276–1288.

van der Laan, M. J., McKeague, I. W., 1998. Efficient estimation from right-censored data when failure indicators are missing at random. Ann. Statist. 26, 164–182.

van der Laan, M.J., Robins, J.M., 2003. Unified methods for censored longitudinal data and causality. Springer, New York.

Robins, J.M., Rotnitzky, A., 1992. Recovery of information and adjustment for dependent censoring using surrogate markers. AIDS Epidemiology-Methodological Issues, Eds. N.Jewell, K.Dietz, V.Farewell, Boston: Birkhauser, 297–331.

Robins, J.M., Rotnitzky, A., and van der Laan, M.J., 2000. Discussion of "On profile likelihood" by Murphy and van der Vaart. J. Ameri. Statist. Assoc. 95, 477–482.

Rotnitzky, A., Robins, J.M., 2005. Inverse probability weighted estimation in survival analysis, Encyclopedia of Biostatistics, Eds. P. Armitage, T. Colton, New York: Wiley.

Satten, G.A., Datta, S., 2001. The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average, Am. Stat. 55, 207–210.

Satten, G.A., Datta, S., 2004. Marginal analysis of multistage data. Handbook of Statistics 23: Advances in Survival Analysis. Eds. N. Balakrishnan and C.R.Rao. 559–574.

Scharfstein, D.O., Rotnitzky, A., Robins, J.M., 1999. Adjusting for nonignorable drop-out using semiparametric nonresponse models (with discussion). J. Ameri. Statist. Assoc. 94, 1096–1146.

Subramanian, S., 2004a. The missing censoring-indicator model of random censorship. Handbook of Statistics 23: Advances in Survival Analysis. Eds. N. Balakrishnan and C.R.Rao. 123–141.

Subramanian, S., 2004b. Asymptotically efficient estimation of a survival function in the

missing censoring indicator model. J. Nonparametr. Statist. 16, 797–817.

Subramanian, S., 2006. Survival analysis for the missing censoring indicator model using kernel density estimation techniques. Stat. Methodol. 3, 125–136.

Subramanian, S., Bean, D., 2008a. Hazard function estimation from homogeneous right censored data with missing censoring indicators. Stat. Methodol. (in press).

Subramanian, S., Bandyopadhyay, D., 2008b. Semiparametric left truncation and right censorship models with missing censoring indicators. Statist. Probab. Lett. 78, 2572–2577.

Tsiatis, A. A. 2006. Semiparametric theory and missing data. Springer, New York.