

# Efron–Petrosian integrals for doubly truncated data with covariates: An asymptotic analysis

JACOBO DE UÑA-ÁLVAREZ<sup>1</sup> and INGRID VAN KEILEGOM<sup>2</sup>

<sup>1</sup>*Department of Statistics and OR, SiDOR research group & CINBIO, Universidade de Vigo, Spain.*  
*E-mail: jacobou@uvigo.es*

<sup>2</sup>*Research Centre for Operations Research and Statistics (ORSTAT), KU Leuven, Belgium.*  
*E-mail: ingrid.vankeilegom@kuleuven.be*

In survival analysis, epidemiology and related fields there exists an increasing interest in statistical methods for doubly truncated data. Double truncation appears with interval sampling and other sampling schemes, and refers to situations in which the target variable is subject to two (left and right) random observation limits. Doubly truncated data require specific corrections for the observational bias, and this affects a variety of settings including the estimation of marginal and multivariate distributions, regression problems, and multi-state models. In this work multivariate Efron–Petrosian integrals for doubly truncated data are introduced. These integrals naturally arise when the goal is the estimation of the mean of a general transformation which involves the doubly truncated variable and covariates. An asymptotic representation of the Efron–Petrosian integrals as a sum of i.i.d. terms is derived and, from this, consistency and distributional convergence are established. As a by-product, uniform i.i.d. representations for the marginal nonparametric maximum likelihood estimator and its corresponding weighting process are provided. Applications to correlation analysis, regression, and competing risks models are presented. A simulation study is reported too.

*Keywords:* Donsker class; double truncation; interval sampling; survival analysis; weak convergence

## 1. Introduction

Doubly truncated data appear in epidemiology, survival analysis and other fields such as astronomy, reliability engineering or econometrics. Under double truncation, the variable of interest  $X$  is observed only when the condition  $U \leq X \leq V$  is satisfied, where  $(U, V)$  with  $U \leq V$  is the couple of left and right truncating variables independent of  $X$ . In this setting, the distribution function (df) of  $X$  is not identifiable unless  $a_U \leq a_X$  and  $b_X \leq b_V$ , where  $a_\xi$  and  $b_\xi$  denote the lower and upper limits of the support of a random variable  $\xi$  (Woodrooffe [33]). However, even under such identifiability conditions, the sampling information on  $X$  is biased due to the double truncation phenomenon, so specific estimation and inference procedures are needed.

In the seminal paper by Efron and Petrosian [7], the nonparametric (conditional) maximum likelihood estimator (NPMLE) of the df of a doubly truncated variable was introduced; two different iterative algorithms to compute the NPMLE were proposed too. Shen [24] proved that the Efron–Petrosian NPMLE maximizes indeed the full likelihood, and provided a new iterative algorithm to compute both the NPMLE of the df of  $X$  and that pertaining to the truncating couple  $(U, V)$ . He also stated some asymptotic properties of the NPMLE. Moreira and de Uña-Álvarez [15] investigated the performance of the bootstrap, while Moreira et al. [18] presented an R package implementing the iterative algorithms of Efron and Petrosian [7] and Shen [24]. Moreira and de Uña-Álvarez [16], see also Shen [26], proposed a semiparametric estimator alternative to the NPMLE when a parametric model for the truncating variables is available; and Moreira et al. [19] developed goodness-of-fit tests for such semiparametric model. Zhu and Wang [36,37] considered estimation and inference for bivariate data

when the first component is doubly truncated by interval sampling. Austin et al. [2] introduced an estimator with improved efficiency with respect to the Efron–Petrosian NPMLE for special cases of double truncation. Also, Emura et al. [8] investigated the parametric MLE under double truncation for the special exponential family. More recently, Shen and Liu [25], Mandel et al. [14] and Rennert and Xie [21] independently investigated Cox regression with a doubly truncated response; see also Ying et al. [35] for rank-based regression. In this paper additional theoretical developments for the random double truncation model are provided.

Let  $(X_i, U_i, V_i)$ ,  $1 \leq i \leq n$ , be the sampling information, that is, independent and identically distributed (i.i.d.) observations with the conditional distribution of  $(X, U, V)$  given  $U \leq X \leq V$ . Let  $\alpha = P(U \leq X \leq V) > 0$  be the probability of no truncation, and let  $G(x) = P(U \leq X \leq V | X = x) = P(U \leq x \leq V)$  be the sampling probability for a particular value  $X = x$ . The df’s of  $X$  and  $X_1$ , denoted here by  $F$  and  $F^*$  respectively, are linked through

$$F(x) = \alpha \int_{a_X}^x G(t)^{-1} dF^*(t) \tag{1.1}$$

with

$$\alpha = \left[ \int_{a_X}^{b_X} G(t)^{-1} dF^*(t) \right]^{-1},$$

where we assume that  $G(t) > 0$  on the support of  $X$ . Then, an estimator of  $F$  is given by

$$F_n(x) = \alpha_n \int_{a_X}^x G_n(t)^{-1} dF_n^*(t) \tag{1.2}$$

with

$$\alpha_n = \left[ \int_{a_X}^{b_X} G_n(t)^{-1} dF_n^*(t) \right]^{-1},$$

where  $F_n^*$  is the ordinary empirical df of the  $X_i$ ’s and  $G_n$  is an estimator for  $G$ . When  $G_n$  is constructed by maximum-likelihood principles,  $F_n$  equals the Efron–Petrosian NPMLE. See Section 2 for details.

Equations (1.1) and (1.2) indicate that, in order to estimate consistently the df of  $X$ , the observed value  $X_i$  must be inversely weighted by its sampling probability  $G(X_i)$ , which is in practice replaced by  $G_n(X_i)$ . This strategy applies to the more general setting in which a covariate vector  $Z$  is attached to  $X$  too. To be explicit, assume that one is interested in the estimation of

$$S^\varphi = E[\varphi(X, Z)] = \int \varphi(x, z) dF_{XZ}(x, z), \tag{1.3}$$

where  $\varphi$  is an arbitrary integrable transformation and  $F_{XZ}$  stands for the df of  $(X, Z)$ . Special cases of (1.3) naturally appear in the regression setting and in the estimation of cumulative incidence functions, see Section 3 for applications. The joint df  $F_{XZ}(x, z)$  itself is a special case of (1.3) too, which is obtained by considering the indicator function  $\varphi(\tilde{x}, \tilde{z}) = I(\tilde{x} \leq x, \tilde{z} \leq z)$ .

Assume that the truncating couple  $(U, V)$  is independent of  $(X, Z)$ . This is stronger than the conditional independence between  $(U, V)$  and  $X$  given  $Z$ , often used in regression. Note, however, that the unconditional independence is common in the estimation of multivariate distributions under random truncation; see, for example, Gürler [10]. It holds that

$$S^\varphi = \alpha \int \varphi(x, z) G(x)^{-1} dF_{XZ}^*(x, z) = E \left[ \frac{\varphi(X_1, Z_1)}{a(X_1)} \right], \tag{1.4}$$

where  $a(x) = \alpha^{-1}G(x)$  and  $F_{XZ}^*$  is the conditional df of  $(X, Z)$  given  $U \leq X \leq V$ . This suggests the estimator

$$S_n^\varphi = \alpha_n \int \varphi(x, z)G_n(x)^{-1} dF_{XZ,n}^*(x, z) = \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a_n(X_i)}, \tag{1.5}$$

where  $Z_i$  is the  $Z$ -vector attached to  $X_i$ ,  $1 \leq i \leq n$ ,  $F_{XZ,n}^*$  is the ordinary empirical df of the  $(X_i, Z_i)$ 's, and  $a_n(x) = \alpha_n^{-1}G_n(x)$ . When there are no ties among the  $X_i$ 's, (1.5) can be written as  $\sum_{i=1}^n \varphi(X_i, Z_i)F_n\{X_i\}$ , where  $F_n\{X_i\}$  denotes the jump of the Efron–Petrosian NPMLE (1.2) at  $X_i$ , see Section 2.

In this paper, we obtain an asymptotic representation of the Efron–Petrosian integral with covariates,  $S_n^\varphi$ , as a sum of i.i.d. terms plus a remainder which converges to zero in probability at a suitable rate. To this end, a uniform representation for  $a_n(x) - a(x)$  will be derived. The key result will be the analogous representation for the Efron–Petrosian NPMLE itself, since the weighting process  $a_n(\cdot)$  can be written as a functional involving  $F_n(\cdot)$ . Compared to the related literature, the results on the multivariate Efron–Petrosian integral  $S_n^\varphi$  are novel. On the other hand, the uniform i.i.d. representations established for  $S_n^\varphi - S^\varphi$ ,  $F_n(x) - F(x)$  and  $a_n(x) - a(x)$  are of independent interest, since they allow for new applications in the setting of doubly truncated data. These applications go beyond the distributional convergence of  $S_n^\varphi$ ,  $F_n(\cdot)$  and  $a_n(\cdot)$ , see Section 3. Asymptotic i.i.d. representations as those provided in Section 2 open a door for the investigation of testing problems too. This comprises for example, goodness-of-fit for regression models under double truncation (similarly as in Sánchez–Sellero et al. [22], for left-truncated and right-censored data) as well as testing for semiparametric truncation models (Moreira et al. [19]). Apart from this, our proofs repair several flaws in previous papers which investigate the asymptotic properties of the NPMLE  $F_n$ ; see Remark A.1 for details.

The rest of the paper is organized as follows. In Section 2, our main results are stated. Section 3 reports a number of applications of the main result in regression and to the estimation of important targets in the scope of multi-state models. An illustrative simulation study is reported in Section 4. Section 5 gives a final discussion. Some technical lemmas are collected and separately proved in the Appendix.

For the main results in Section 2, besides the independence between  $(U, V)$  and  $(X, Z)$  and the identifiability conditions  $a_U \leq a_X$  and  $b_X \leq b_V$  (which are implied by assumption (A3) below), we assume that  $U$  and  $V$  are absolutely continuous. This does not exclude the frequent situation in which the joint density of  $(U, V)$  does not exist. Such a situation occurs for example with interval sampling, since  $V = U + \tau$  for some constant  $\tau$  (the interval width) in that case. Continuity of  $F$  will not be imposed.

## 2. Main results

In this section, we introduce the main results of this piece of work. Following Efron and Petrosian [7], to introduce the NPMLE of the df of  $X$ , we consider the class of discrete distributions supported on  $\{X_i, 1 \leq i \leq n\}$ . Let  $\phi_i$  be the mass attached to  $X_i$  by  $F$ . Similarly, consider a discrete bivariate df  $K$  for the couple  $(U, V)$ , which gives mass  $\psi_i$  to  $(U_i, V_i)$ . Then, the likelihood of  $(X_i, U_i, V_i)$ ,  $1 \leq i \leq n$ , is given by

$$\mathcal{L}(\phi, \psi) = \prod_{i=1}^n \alpha^{-1} \phi_i \psi_i = \prod_{i=1}^n \frac{\phi_i \psi_i}{\sum_{j=1}^n \Phi_j \psi_j}, \tag{2.1}$$

where  $\phi = (\phi_1, \dots, \phi_n)$ ,  $\psi = (\psi_1, \dots, \psi_n)$ ,  $\Phi_j = \sum_{i=1}^n \phi_i J_{ji}$ , and  $J_{ji} = I(U_j \leq X_i \leq V_j)$ . The NPMLE of  $F$  and  $K$  are given by  $F_n(x) = \sum_{i=1}^n \hat{\phi}_i I(X_i \leq x)$  and  $K_n(u, v) = \sum_{i=1}^n \hat{\psi}_i I(U_i \leq u, V_i \leq v)$ .

$v$ ) respectively, where  $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_n)$  and  $\hat{\psi} = (\hat{\psi}_1, \dots, \hat{\psi}_n)$  are the maximizers of (2.1) under the constraints  $\sum_{i=1}^n \hat{\phi}_i = 1, \sum_{i=1}^n \hat{\psi}_i = 1$ . The likelihood (2.1) is self-adapted to ties in the sense that, if a value  $X_i = x_k$  has multiplicity  $n_k$ , the mass attached to  $x_k$  will be  $n_k \hat{\phi}_i$ . Existence and uniqueness of the NPMLE is discussed in Xiao and Hudgens [34].

The NPMLE  $G_n(x)$  of  $G(x)$  is given by

$$G_n(x) = \sum_{i=1}^n \hat{\psi}_i I(U_i \leq x \leq V_i),$$

while the NPMLE of  $\alpha$  is  $\alpha_n = \sum_{j=1}^n \hat{\Phi}_j \hat{\psi}_j$ , where  $\hat{\Phi}_j = \sum_{i=1}^n \hat{\phi}_i J_{ji}$ . It is easily seen that  $\hat{\phi}_i = \alpha_n (nG_n(X_i))^{-1}, 1 \leq i \leq n$ , with  $\alpha_n = 1 / \sum_{i=1}^n (nG_n(X_i))^{-1}$ . Then, the NPMLE  $F_n$  equals the estimator introduced in (1.2).

Another property of the NPMLE is that  $\hat{\psi}_i = \alpha_n n^{-1} [F_n(V_i) - F_n(U_i-)]^{-1}, 1 \leq i \leq n$ , with  $\alpha_n = 1 / \sum_{i=1}^n n^{-1} [F_n(V_i) - F_n(U_i-)]^{-1}$ . Here,  $F_n(x-)$  stands for the left-continuous version of  $F_n(x)$ . This property shows that the NPMLE  $F_n(x)$  is the solution of

$$\mathcal{U}(F, F_n^*, K_n^*)(x) = 0, \tag{2.2}$$

with

$$\mathcal{U}(F, F_n^*, K_n^*)(x) = F(x) - \int_{a_X}^x \frac{dF_n^*(t)}{\tilde{a}_n(t)},$$

where

$$\tilde{a}_n(t) = \int_{u \leq t \leq v} \frac{dK_n^*(u, v)}{F(v) - F(u-)},$$

and where  $K_n^*$  denotes the ordinary empirical df of the  $(U_i, V_i)$ 's. In the following result, we establish the in-probability uniform consistency and an i.i.d. representation for the solution  $F_n$  of equation (2.2). Put  $K^*$  for the joint df of  $(U_1, V_1)$ , and introduce

$$a(t) = \alpha^{-1} G(t) = \int_{u \leq t \leq v} \frac{dK^*(u, v)}{F(v) - F(u-)},$$

$$\bar{a}_n(t) = \int_{u \leq t \leq v} \frac{dK^*(u, v)}{F_n(v) - F_n(u-)}$$

and

$$a_n(t) = \alpha_n^{-1} G_n(t) = \int_{u \leq t \leq v} \frac{dK_n^*(u, v)}{F_n(v) - F_n(u-)}.$$

Throughout this section we will refer to the following assumptions:

- (A1)  $F$  has only a finite number of discontinuity points and is continuous everywhere else;  $a_X < a_{V_1}, b_{U_1} < b_X; b_X - a_X < \infty$ ; and  $P(V_1 - U_1 \geq \gamma) = 1$  for some  $\gamma > 0$
- (A2)  $\inf_{(u,v) \in \text{supp}(K^*)} (F(v) - F(u-)) > \delta_1 > 0$
- (A3)  $\inf_{a_X \leq x \leq b_X} G(x) > \delta_2 > 0$
- (A4) The marginal densities of  $U_1$  and  $V_1$  are bounded on  $[a_X, b_X]$ , and  $b_X - a_X < \infty$

Assumption (A1), together with (A2)–(A4), is needed for the uniform consistency of  $F_n$ . Indeed, (A1) is used to guarantee that, given  $F^*$  and  $K^*$ , the solution to  $\mathcal{U}(F, F^*, K^*)(x) = 0$  is unique; see the

Appendix for details. Condition (A1) allows for continuous, discrete and mixed  $F$ -distributions, only limiting the amount of point masses to be finite. On the other hand, conditions  $a_X < a_{V_1}$ ,  $b_{U_1} < b_X$  and  $P(V_1 - U_1 \geq \gamma) = 1$  in (A1) are only slightly stronger versions of the obvious identifiability conditions  $a_X \leq a_V$ ,  $b_U \leq b_X$  and  $P(V - U \geq 0) = 1$ . We mention that, with interval sampling,  $P(V_1 - U_1 \geq \gamma) = 1$  holds for any  $\gamma$  smaller than or equal to the interval width. Condition  $P(V_1 - U_1 \geq \gamma) = 1$  was needed by Xiao and Hudgens [34] to prove that there exists a unique NPMLE of  $F$ , so it precludes problems in the very definition of  $F_n$ ; it was also used by Frank et al. [9] to investigate the particular setting in which  $V - U$  is independent of  $U$ . Assumptions (A2) and (A3) state that the sampling probability for the truncating couple and the variable of interest remain bounded away from zero. Interestingly, (A2) holds under (A1), but we keep the former separately for ease of reference; note that (A1) is not required for Lemmas A.2–A.5 in the Appendix. On its turn, by using arguments similar to those in the proof to Lemma A.6, it is easy to see that (A3) is satisfied under  $P(V_1 - U_1 \geq \gamma) = 1$  when  $a_{U_1} < a_X$ ,  $b_X < b_{V_1}$  hold. Finally, condition (A4) is needed for the boundedness of the operator  $\mathcal{A}: \ell^\infty[a_X, b_X] \rightarrow \ell^\infty[a_X, b_X]$  appearing in Theorem 2.1 below; here,  $\ell^\infty[a_X, b_X]$  denotes the space of bounded functions defined on the support of  $X$ , with the norm  $\|h\|_\infty = \sup_{a_X \leq x \leq b_X} |h(x)|$ . The proof of Theorem 2.1 is postponed to the Appendix.

**Theorem 2.1.** *Let  $F_n$  be the solution of (2.2). Under the assumptions (A1)–(A4), (a) it holds  $\sup_{a_X \leq x \leq b_X} |F_n(x) - F(x)| = o_P(1)$ ; and (b) there exists a bounded linear operator  $\mathcal{A} = \mathcal{A}(F, F^*, K^*)$  such that*

$$F_n(x) - F(x) = \frac{1}{n} \sum_{i=1}^n \mathcal{A}[h_{X_i, U_i, V_i}](x) + o_P(n^{-1/2}),$$

uniformly in  $x \in [a_X, b_X]$ , where

$$h_{X_i, U_i, V_i}(x) = \frac{I(X_i \leq x)}{a(X_i)} - \frac{I(U_i \leq x)}{F(V_i) - F(U_i-)} \int_{a_X \vee U_i}^{x \wedge V_i} \frac{dF^*(t)}{a(t)^2}.$$

**Remark 2.1.** Our Theorem 2.1 is obviously connected to Theorems 2 and 3 in Shen [24]. However, in the Appendix we indicate some important gaps and inconsistencies in Shen [24] which are solved by our assumptions and argumentations. At this point, we only mention that, unlike in Shen [24], our assumptions (A1)–(A4) allow for a possibly non-continuous  $F$ , and that the joint density of  $(U, V)$  is not assumed to exist. As indicated in the Introduction, the latter is critical to include applications with interval sampling.

**Remark 2.2.** For the left-truncated setting, Woodroffe [33] investigated conditions under which the observable distributions ( $F^*$  and  $K^*$ ) determine the underlying ones ( $F$  and  $K$ ). In principle, the natural extension of his Theorem 1 to double truncation would be that, given  $F^*$ ,  $K^*$  and  $\alpha > 0$ , there exists a unique couple  $(F, K)$  with  $a_U \leq a_X \leq a_V$  and  $b_U \leq b_X \leq b_V$  giving rise to the same joint distribution of  $(X_1, U_1, V_1)$ . Discussion of (A1)–(A4) above implies a result which is ‘almost’ that one. Namely,  $(F, K)$  is unique within the class of couples satisfying  $a_U < a_X < a_V$  and  $b_U < b_X < b_V$ ,  $b_X - a_X < \infty$  and  $P(V - U \geq \gamma) = 1$  for some  $\gamma > 0$ , provided that (e.g., to simplify the statement) the densities of  $X, U$  and  $V$  exist and are continuous.

**Remark 2.3.** Theorem 2.1(b) generalizes existing results for one-sided truncation. For example, the i.i.d. representation  $n^{-1} \sum_{i=1}^n \mathcal{A}[h_{X_i, U_i, V_i}](x)$  reduces to that in Theorem 2 in Stute [27], see also Güler et al. [11], when one removes the right truncation limit  $V$ . See Remark A.2 in the Appendix for details.

**Corollary 2.1.** *Under (A1)–(A4), the process  $\sqrt{n}(F_n(x) - F(x))$ ,  $a_X \leq x \leq b_X$ , converges weakly to a zero-mean Gaussian process  $\mathcal{W}(x)$  with covariance function given by*

$$\text{Cov}(\mathcal{A}[h_{X_1, U_1, V_1}](x_1), \mathcal{A}[h_{X_1, U_1, V_1}](x_2)) = E(\mathcal{A}[h_{X_1, U_1, V_1}](x_1)\mathcal{A}[h_{X_1, U_1, V_1}](x_2)).$$

**Proof.** The result follows from Theorem 2.1 and Lemma A.4 in the Appendix. □

**Remark 2.4.** An estimator for the variance of the process  $\sqrt{n}(F_n(x) - F(x))$  can be introduced by using the i.i.d. representation in Theorem 2.1. For this, both the function  $h_{X_i, U_i, V_i}(x)$  and the operator  $\mathcal{A}$  must be estimated. The former is simply done by replacing the unknown  $a(x)$ ,  $F(x)$  and  $F^*(x)$  by their empirical counterparts  $a_n(x)$ ,  $F_n(x)$  and  $F_n^*(x)$  to give  $\hat{h}_{X_i, U_i, V_i}(x)$  say. The latter is not so immediate. Lemma A.3 in the Appendix gives  $\mathcal{A} = \sum_{r=0}^\infty A^r$  where  $A$  is a functional involving  $F$ ,  $K^*$ ,  $F^*$  and  $a$ . Introduce

$$A_n[h](x) = \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F_n(v) - F_n(u-))^2} dK_n^*(u, v) \frac{dF_n^*(t)}{a_n(t)^2}$$

and, from this,  $\mathcal{A}_n = \sum_{r=0}^{r^*} A_n^r$ , for some trimming parameter  $r^*$ . Let  $\mathcal{A}_{n,i} = \mathcal{A}_n[\hat{h}_{X_i, U_i, V_i}]$ . Then, the variance of  $\sqrt{n}(F_n(x) - F(x))$  can be approximated by  $n^{-1} \sum_{i=1}^n \mathcal{A}_{n,i}(x)^2$ . The theoretical properties and the practical performance of this estimator are interesting open questions.

Now, we provide asymptotic results for the process  $a_n(x) - a(x)$ . Write

$$a_n(x) - a(x) = a_n(x) - \bar{a}_n(x) + \bar{a}_n(x) - a(x).$$

We have

$$a_n(x) - \bar{a}_n(x) = \int_{u \leq x \leq v} \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)} + o_P(n^{-1/2})$$

uniformly in  $x$ , see the proof to Lemma A.1 in the Appendix. On the other hand, for  $\bar{a}_n(x) - a(x)$  we write

$$\begin{aligned} \bar{a}_n(x) - a(x) &= \int_{u \leq x \leq v} \frac{F(v) - F_n(v) + F_n(u-) - F(u-)}{(F(v) - F(u-))^2} dK^*(u, v) \\ &\quad + \int_{u \leq x \leq v} \frac{F(v) - F_n(v) + F_n(u-) - F(u-)}{F(v) - F(u-)} \\ &\quad \times \left[ \frac{1}{F_n(u) - F_n(v-)} - \frac{1}{F(v) - F(u-)} \right] dK^*(u, v). \end{aligned}$$

Since by Corollary 2.1  $F_n(x) - F(x) = O_P(n^{-1/2})$  uniformly in  $x \in [a_X, b_X]$ , the second term at the right-hand side is  $O_P(n^{-1})$  uniformly in  $x$  under (A2). Summarising, we have the following results.

**Theorem 2.2.** *Under the assumptions (A1)–(A4), we have*

$$a_n(x) - a(x) = \frac{1}{n} \sum_{i=1}^n \left[ \frac{I(U_i \leq x \leq V_i)}{F(V_i) - F(U_i-)} - a(x) - \tilde{\mathcal{A}}[h_{X_i, U_i, V_i}](x) \right] + o_P(n^{-1/2})$$

uniformly in  $x \in [a_X, b_X]$ , where

$$\tilde{A}[h](x) = \int_{u \leq x \leq v} \frac{\mathcal{A}[h](v) - \mathcal{A}[h](u-)}{(F(v) - F(u-))^2} dK^*(u, v).$$

**Proof.** The result is a direct consequence of the derivations above and the i.i.d. representation of  $F_n$  in Theorem 2.1. □

**Corollary 2.2.** Under (A1)–(A4), the process  $\sqrt{n}(a_n(x) - a(x))$ ,  $a_X \leq x \leq b_X$ , converges weakly to a zero-mean Gaussian process  $\mathcal{W}_0(x)$  with covariance function given by

$$\begin{aligned} & \text{Cov} \left( \left[ \frac{I(U_1 \leq x_1 \leq V_1)}{F(V_1) - F(U_1-)} - a(x_1) - \tilde{A}[h_{X_1, U_1, V_1}](x_1) \right], \right. \\ & \left. \left[ \frac{I(U_1 \leq x_2 \leq V_1)}{F(V_1) - F(U_1-)} - a(x_2) - \tilde{A}[h_{X_1, U_1, V_1}](x_2) \right] \right) \\ &= E \left( \left[ \frac{I(U_1 \leq x_1 \leq V_1)}{F(V_1) - F(U_1-)} - a(x_1) - \tilde{A}[h_{X_1, U_1, V_1}](x_1) \right] \right. \\ & \quad \left. \times \left[ \frac{I(U_1 \leq x_2 \leq V_1)}{F(V_1) - F(U_1-)} - a(x_2) - \tilde{A}[h_{X_1, U_1, V_1}](x_2) \right] \right). \end{aligned}$$

**Proof.** The result follows from Theorem 2.2 and Lemma A.5 in the Appendix. □

The i.i.d. representation for  $a_n(x) - a(x)$  in Theorem 2.2 leads now to an i.i.d. representation for (1.5). This is included in the following result.

**Theorem 2.3.** Under the assumptions (A1)–(A4), we have that  $S_n^\varphi \rightarrow S^\varphi$  in probability for any integrable function  $\varphi$ . If in addition  $\varphi$  is square-integrable, we have

$$S_n^\varphi = \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a(X_i)} - \frac{1}{n} \sum_{i=1}^n \zeta^\varphi(X_i, U_i, V_i) + o_P(n^{-1/2}),$$

where

$$\zeta^\varphi(x, u, v) = E \left( \frac{\varphi(X_1, Z_1)}{a(X_1)^2} \left[ \frac{I(u \leq X_1 \leq v)}{F(v) - F(u-)} - a(X_1) - \tilde{A}[h_{x, u, v}](X_1) \right] \right).$$

**Proof.** Write

$$\begin{aligned} S_n^\varphi &= \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a_n(X_i)} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a(X_i)} - \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{\Delta_{ni}} (a_n(X_i) - a(X_i)) \equiv \bar{S}_n^\varphi + R_{n1}^\varphi, \end{aligned}$$

where  $\Delta_{ni}$  is between  $a_n(X_i)$  and  $a(X_i)$ . Since by the Weak Law of Large Numbers,  $\bar{S}_n^\varphi \rightarrow E[\varphi(X_1, Z_1)a(X_1)^{-1}] = S^\varphi$  in probability, the weak consistency of  $S_n^\varphi$  follows provided that  $R_{n1}^\varphi =$

$o_P(1)$ , which in its turn can be ensured under (A3) if  $\|a_n - a\|_\infty = o_P(1)$ . Now, in order to prove this, write

$$a_n(x) - a(x) = \int_{u \leq x \leq v} \left[ \frac{1}{F_n(v) - F_n(u-)} - \frac{1}{F(v) - F(u-)} \right] dK_n^*(u, v) + \int_{u \leq x \leq v} \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)}.$$

The first term is  $o_P(1)$  uniformly in  $x$  under (A1)–(A4). Note that these conditions imply that  $1/\inf_{(u,v) \in \text{supp}(K^*)} (F_n(v) - F_n(u-)) = O_P(1)$  which can be used to deal with the denominator. Since the second term is  $O_P(n^{-1/2})$  uniformly in  $x$  (because the function  $(u, v) \rightarrow 1/(F(v) - F(u-))$  is bounded under (A2)), the result follows.

In order to prove the second assertion, use a second order Taylor expansion of  $v(s) = 1/s$  around  $a(X_i)$  to get

$$S_n^\varphi = \bar{S}_n^\varphi - \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a(X_i)^2} (a_n(X_i) - a(X_i)) + R_{n2}^\varphi,$$

where

$$R_{n2}^\varphi = -\frac{2}{n} \sum_{i=1}^n \varphi(X_i, Z_i) \frac{1}{\Delta_{ni}^3} (a_n(X_i) - a(X_i))^2,$$

with  $\Delta_{ni}$  between  $a_n(X_i)$  and  $a(X_i)$ . Note that the weak convergence of  $a_n(x) - a(x)$  in Corollary 2.2, together with (A3), imply that  $R_{n2}^\varphi = O_P(n^{-1}) = o_P(n^{-1/2})$ .

Now, for the term

$$W_n^\varphi \equiv \frac{1}{n} \sum_{i=1}^n \frac{\varphi(X_i, Z_i)}{a(X_i)^2} (a_n(X_i) - a(X_i))$$

use the asymptotic representation in Theorem 2.2 to get

$$W_n^\varphi = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{\varphi(X_i, Z_i)}{a(X_i)^2} \left[ \frac{I(U_j \leq X_i \leq V_j)}{F(V_j) - F(U_j-)} - a(X_i) - \tilde{A}[h_{X_j, U_j, V_j}](X_i) \right] + o_P(n^{-1/2}),$$

where we have used (A2) and (A3). The main part of  $W_n^\varphi$ ,  $W_{0,n}^\varphi$  say, is a  $U$ -statistic with Hájek projection given by

$$\hat{W}_{0,n}^\varphi = \sum_{k=1}^n E(W_{0,n}^\varphi | \mathcal{X}_k) - (n-1)E(W_{0,n}^\varphi) = \frac{1}{n(n-1)} \sum_{k=1}^n \sum_{i \neq j} E_{ijk} - (n-1)E(W_{0,n}^\varphi),$$

where

$$E_{ijk} = E \left( \frac{\varphi(X_i, Z_i)}{a(X_i)^2} \left[ \frac{I(U_j \leq X_i \leq V_j)}{F(V_j) - F(U_j-)} - a(X_i) - \tilde{A}[h_{X_j, U_j, V_j}](X_i) \right] \middle| \mathcal{X}_k \right),$$



and where  $\mathcal{X}_k = \{U_k, V_k, X_k, Z_k\}$ . Since  $E(W_{0,n}^\varphi) = 0$  and  $E_{ijk} = 0$  whenever  $j \neq k$ , we get

$$\hat{W}_{0,n}^\varphi = \frac{1}{n} \sum_{k=1}^n \zeta^\varphi(X_k, U_k, V_k).$$

Finally, since  $E(\zeta^\varphi(X_1, U_1, V_1)^2) < \infty$ , Theorem 5.3.3 in Serfling [23] implies that  $W_{0,n}^\varphi - \hat{W}_{0,n}^\varphi = o((\log n)^{(1/2+\rho)}/n)$  with probability 1 with  $\rho > 0$  and, hence, the result follows.  $\square$

**Corollary 2.3.** *Under the assumptions (A1)–(A4), for any square-integrable function  $\varphi$  we have that  $n^{1/2}(S_n^\varphi - S^\varphi) \rightarrow N(0, \sigma_\varphi)$  in law, where*

$$\sigma_\varphi^2 = \text{Var}\left(\frac{\varphi(X_1, Z_1)}{a(X_1)} - \zeta^\varphi(X_1, U_1, V_1)\right).$$

**Proof.** The expectation of the first term in the representation of Theorem 2.3 is  $S^\varphi$ , while the second term has zero mean. Then, the result follows from Theorem 2.3 and the Central Limit Theorem (CLT).  $\square$

**Remark 2.5.** Double truncation can be seen as a special case of a general biased sampling setting in which the target vector  $(X, Z)$  is observed only when  $X \in W$ ,  $W$  being a random set independent of  $(X, Z)$ . In such a general setting the function  $G$  becomes  $G(x) = P(x \in W)$ . In principle, given a random sample  $(X_i, Z_i, W_i)$ ,  $1 \leq i \leq n$ , with the distribution of  $(X, Z, W)$  conditional on  $X \in W$ , estimation of  $G$  and  $F$  by maximum likelihood is possible; just use the indicator  $J_{ji} = I(X_i \in W_j)$  in (2.1). However, the particular shape of  $W$  and its probability law relative to  $F$  are critical to ensure the existence of the NPMLE and to allow for an asymptotic analysis analogous to that accomplished in this paper. Extension of our results to more general forms of random truncation is currently under investigation.

### 3. Applications

In this section, we give some applications of our main results. These concern particular Efron–Petrosian integrals with covariates of the form as in (1.5). Specifically, applications to correlation analysis, accelerated failure time regression, Cox regression and cumulative incidence functions for competing risks will be given.

#### 3.1. Correlation analysis

The Pearson correlation coefficient depends on quantities of the type (1.4) for special  $\varphi$  functions, namely  $\varphi(x, z) = x$ ,  $\varphi(x, z) = z$ ,  $\varphi(x, z) = x^2$ ,  $\varphi(x, z) = z^2$ , and  $\varphi(x, z) = xz$ . Each of these quantities is estimated through (1.5) to get an empirical Pearson correlation coefficient, which is consistent under the conditions in Theorem 2.3. The asymptotic normality of the Pearson correlation coefficient follows then by the i.i.d. representation in Theorem 2.3, the multivariate CLT and the Delta method.

### 3.2. Accelerated failure time regression

In this section, we consider the parametric regression setting in which the goal is to estimate and to do inference for a vector of regression coefficients. Specifically, we consider the linear regression model

$$X = \beta_0^t Z + \epsilon, \tag{3.1}$$

where  $\epsilon$  is an error term such that  $E(\epsilon|Z) = 0$  and  $\beta_0$  is the vector of (true) regression coefficients. Since the response variable  $X$  is observed under random double truncation, ordinary estimation approaches (such as *e.g.*, least squares) must be properly adapted. In practice, the response variable is often a logarithm transformation of the (truncated) lifetime of ultimate interest, leading to the so-called accelerated failure time model. In such a case,  $X$  is replaced by  $\log(X)$  in (3.1).

In order to estimate  $\beta_0$ , introduce the randomly weighted least squares criterion

$$\beta \mapsto \sum_{i=1}^n (X_i - \beta^t Z_i)^2 G_n(X_i)^{-1}. \tag{3.2}$$

Denote the minimizer of (3.2) by  $\beta_n$ . It is straightforward to verify that  $\beta_n = M_{1n}^{-1} M_{2n}(X_1, \dots, X_n)$ , where  $M_{1n}$  and  $M_{2n}$  are the matrices with entries

$$M_{1n}(i, j) = \alpha_n \sum_{k=1}^n Z_k^i Z_k^j G_n(X_k)^{-1}, \quad 1 \leq i, j \leq p,$$

$$M_{2n}(i, k) = \alpha_n Z_k^i G_n(X_k)^{-1}, \quad 1 \leq i \leq p, 1 \leq k \leq n,$$

where  $Z_k^i$  denotes the  $i$ -th coordinate of  $Z_k$  and  $p$  is the dimension of  $Z_k$ . Then, the consistency of  $S_n^\varphi$  for the particular transformations  $\varphi_1^{ij}(x, z) = z^i x^j$  and  $\varphi_2^i(x, z) = z^i x$ ,  $1 \leq i, j \leq p$ , gives  $\beta_n - \beta_0 = o_p(1)$  under a first order assumption. For this, note that the limits of  $M_{1n}$  and  $M_{2n}(X_1, \dots, X_n)$  are  $E(ZZ^t)$  and  $E(ZZ^t)\beta_0$ , respectively. Besides, when the involved  $\varphi$ -functions satisfy  $E(\varphi(X, Z)^2) < \infty$ , an application of Theorem 2.3, the multivariate CLT and the Delta method gives the asymptotic normality of  $\sqrt{n}(\beta_n - \beta_0)$ .

Our main result can be applied to nonlinear regression too. Assume, rather than (3.1),

$$X = f(Z, \beta_0) + \epsilon$$

for a certain family of smooth predictors  $\{f(\cdot, \beta), \beta \in B\}$ , where again  $\beta_0$  denotes the true value of the  $q$ -dimensional vector of regression coefficients  $\beta$  ( $q \neq p$  is possible), and where  $E(\epsilon|Z) = 0$ . Furthermore, assume that

- (i)  $E(X^2) < \infty$
- (ii)  $B$  is compact
- (iii)  $\beta \rightarrow f(z, \beta)$  is continuous for each  $z$
- (iv)  $f(z, \beta)^2 \leq M(z)$  for some integrable function  $M(z)$
- (v)  $E\{f(Z, \beta) - f(Z, \beta_0)\}^2 > 0$  for  $\beta \neq \beta_0$

Then, the consistency of the Efron–Petrosian integral for the special function  $\varphi_\beta(X, Z) = \{X - f(Z, \beta)\}^2$  ensures that the minimizer  $\beta_n$  of

$$\beta \mapsto \sum_{i=1}^n (X_i - f(Z_i, \beta))^2 G_n(X_i)^{-1}.$$

satisfies  $\beta_n \rightarrow \beta_0$  in probability. Besides, the asymptotic normality of  $\beta_n$  follows from Theorem 2.3, the multivariate CLT and the Delta method, provided that the functions  $\varphi_i(X, Z) = \{X - f(Z, \beta_0)\} \frac{\partial f(Z, \beta_0)}{\partial \beta_i}$ ,  $1 \leq i \leq q$ , are square-integrable and that  $f(z, \cdot)$  is twice continuously differentiable. The straightforward details are omitted. See Stute [28] or de Uña-Álvarez [4] for similar results in the setting with right-censoring or biased sampling, respectively.

Another extension of (3.2) is obtained by considering a general function  $\rho(w)$  instead of the squared loss. That is,  $\beta_n$  becomes the minimizer of

$$\beta \mapsto \sum_{i=1}^n \rho(X_i - \beta^t Z_i) G_n(X_i)^{-1},$$

an M-estimator. As for least squares estimation, formal asymptotic theory can be established for this generalized  $\beta_n$  by using our Theorem 2.3. When the loss function  $\rho(w)$  is continuously differentiable asymptotics follow by arguments similar to those of nonlinear regression. Nevertheless, one will generally wish to include in the analysis non-differentiable loss functions; an important example is given by the least absolute deviation estimator, which is based on  $\rho(w) = |w|$ . Theory for M-estimation with possibly non-differentiable convex  $\rho(w)$  is more elaborated, and may naturally involve dealing with subgradients; see, for example, Bai et al. [3] or Du et al. [6]. With convex loss, a uniform version of our consistency result in Theorem 2.3 along the class  $\varphi_\beta(x, z) = \rho(x - \beta^t z)$ ,  $\beta \in B$ , is easily obtained (Pollard [20]), which is important for the application of standard consistency results in M-estimation (e.g., Theorem 5.7 in van der Vaart [30]). This somehow resembles the application of Theorem 2 in Jennrich [12] for continuous predictors in the nonlinear regression setting as considered by Stute [28]. Full formal derivations are out of the scope of the present paper and will be provided elsewhere.

### 3.3. Cox regression

Cox regression is based on the proportional hazards assumption

$$h(x|Z) = h_0(x)e^{\beta^t Z},$$

where  $h_0(\cdot)$  is a baseline hazard and  $\beta$  is the vector of regression coefficients attached to the covariate vector  $Z$ . Mandel et al. [14] introduced the estimating equation for  $\beta$  given by  $U_n(\beta) = 0$ , where

$$U_n(\beta) = \sum_{i=1}^n \left[ Z_i - \frac{\sum_{j=1}^n Z_j e^{\beta^t Z_j} I(X_j \geq X_i) G_n(X_j)^{-1}}{\sum_{j=1}^n e^{\beta^t Z_j} I(X_j \geq X_i) G_n(X_j)^{-1}} \right]. \tag{3.3}$$

Obviously,  $G_n(x)$  can be replaced by  $a_n(x) = \alpha_n^{-1} G_n(x)$  in (3.3). Large sample properties for the solution of  $U_n(\beta) = 0$  were derived by Mandel et al. [14] under the following couple of conjectures:

- C1.  $\max_{1 \leq i \leq n} |a_n(X_i) - a(X_i)| = o_P(1)$
- C2.  $\sqrt{n}(a_n(x) - a(x)) = n^{-1/2} \sum_{i=1}^n \zeta_{n,i}(x) + o_P(1)$  uniformly in  $x$  for zero-mean i.i.d. variables  $\zeta_{n,i}(x) = \zeta_n(X_i, U_i, V_i, x)$  with  $E((\zeta_{n,i}(x))^2) < \infty$

Our Theorem 2.2 shows that C1 and C2 hold under conditions (A1)–(A4).

### 3.4. Cumulative incidence functions

Cumulative incidence functions (CIFs) naturally appear in the competing risks setting. Let  $X$  denote the lifetime of interest, and consider the case in which  $K$  competing events are present. Define the event

indicator as a covariate  $Z$  taking values on the set  $\{1, \dots, K\}$ . The goal is the estimation of the CIFs  $F_j(x) = P(X \leq x, Z = j)$ ,  $1 \leq j \leq K$ . Note that  $F_j(x)$  is just the expectation of the transformation  $\varphi_j(X, Z) = I(X \leq x, Z = j)$ , which is obviously square-integrable. Introduce the estimators

$$F_{jn}(x) = \alpha_n \sum_{i=1}^n I(X_i \leq x, Z_i = j) G_n(X_i)^{-1}, \quad 1 \leq j \leq K. \quad (3.4)$$

Then, under conditions (A1)–(A4), we have that  $F_{jn}(x) - F_j(x) = o_P(1)$  (Theorem 2.3) and  $\sqrt{n}(F_{jn}(x) - F_j(x))$  converges in law to a Gaussian random variable (Corollary 2.3).

## 4. Simulation study

In this section, we perform a number of simulation studies related to some of the applications mentioned in Section 3. To be specific, the performance of the Pearson correlation coefficient, the weighted least squares estimator for the linear regression model, and the cumulative incidence functions for competing risks are investigated. See Mandel et al. [14] for simulations of Cox regression with a doubly truncated response, and de Uña-Álvarez [5] for more on competing risks under double truncation.

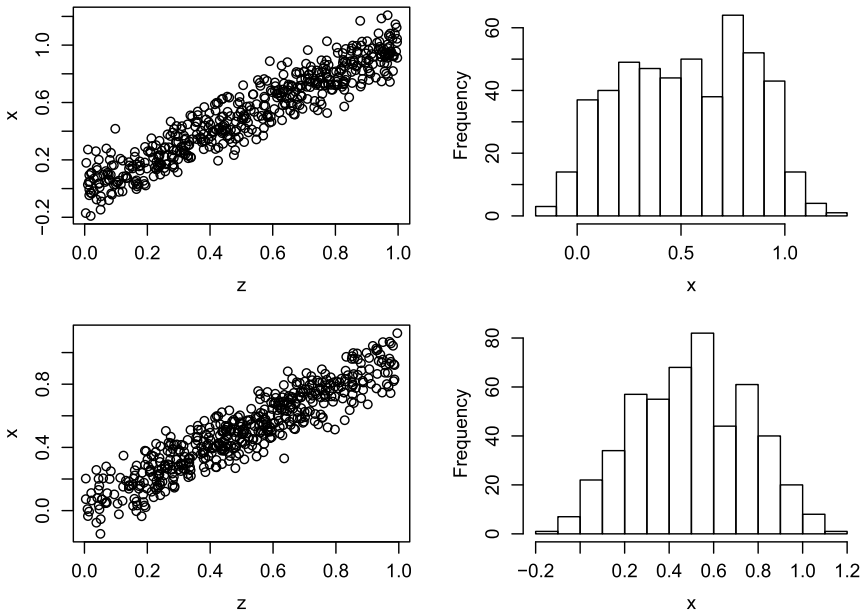
### 4.1. Pearson correlation and weighted least squares regression

We have simulated an interval sampling scenario in which the couple of interest  $(X, Z)$  is observed only when  $U \leq X \leq V$  with  $V = U + \tau$ . Specifically,  $Z \sim U(0, 1)$ ,  $X = Z + \sigma\varepsilon$ ,  $\varepsilon \sim N(0, 1)$  (independent of  $X$ ),  $\sigma = 0.1$ ,  $U \sim U(-0.25, 0.5)$  (independent of  $X$  and  $\varepsilon$ ), and  $\tau = 0.75$ . Hence, the regression function is a line with intercept  $\beta_0 = 0$  and slope  $\beta_1 = 1$ . The true correlation between  $X$  and  $Z$  is  $\rho_{XZ} = (1 + \sigma^2 / \text{Var}(Z))^{-1/2} = 0.945$ . Note that  $X$  is ‘almost’ identifiable: the probability of the event  $\{a_U < X < b_V\}$  is larger than 0.9996. The simulated scenario is illustrated in Figure 1, where a single sample of size  $n = 500$  before (top) and after (bottom) truncation is displayed by means of scatterplots and histograms. From Figure 1, it is seen that the general trend is apparently unchanged by the double truncation, and that the marginal distribution of  $X$  changes a lot; specifically, small and large values of  $X$  are observed with a relatively smaller probability.

Table 1 reports the mean and standard deviation of the ordinary Pearson correlation coefficient, denoted by  $\hat{\rho}_{XZ}$ , and of the estimator which takes the double truncation into account,  $\hat{\rho}_{XZ}$  say, along 1000 Monte Carlo trials with sample sizes  $n = 100, 250$  and  $500$ . The results of the weighted least squares estimator (3.2) are reported too. From Table 1 it is seen that the naive estimator has a moderate, albeit visible systematic bias, and that the estimators which correct for double truncation perform consistently.

### 4.2. Cumulative incidence functions

We consider a competing risks model with an initial state 0 and two absorbing states 1 and 2 corresponding to two competing events. The cause-specific hazards are  $\alpha_{01}(x) = \lambda_1$  and  $\alpha_{02}(x) = \lambda_2 x$ , which results in a hazard rate  $\alpha(x) = \lambda_1 + \lambda_2 x$  for  $X$ , with corresponding df  $F(x) = 1 - e^{-\lambda_1 x - \lambda_2 x^2/2}$ . To simulate this model, we use the inversion method for  $X$  and, given  $X = x$ , the event indicator  $I(Z = 1)$  is drawn from a Bernoulli distribution with  $p(x) \equiv P(Z = 1) = \alpha_{01}(x) / \alpha(x)$ . In this model, the cumulative incidence functions  $F_1(x)$  and  $F_2(x)$  corresponding to the competing events  $Z = 1$



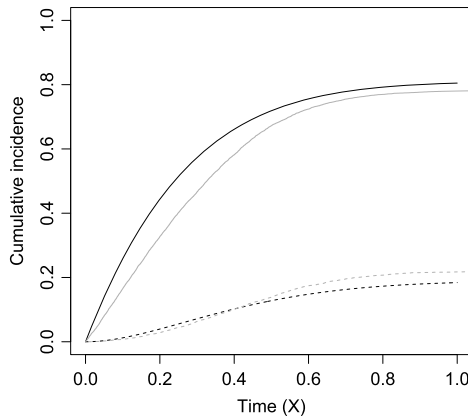
**Figure 1.** Correlation/regression setting. Scatterplots and histograms for a single sample of size  $n = 500$  from the simulated model: before truncation (top) and after truncation (bottom).

and  $Z = 2$  are  $F_1(x) = \lambda_1 e^{\lambda_1^2/(2\lambda_2)} \sqrt{2\pi/\lambda_2} (\Phi_{\mu,\sigma}(x) - \Phi_{\mu,\sigma}(0))$  and  $F_2(x) = F(x) - F_1(x)$ , where  $\Phi_{\mu,\sigma}$  denotes the df of a  $N(\mu, \sigma)$  variable,  $\mu = -\lambda_1/\lambda_2$  and  $\sigma = 1/\sqrt{\lambda_2}$ . The truncating variables  $U$  and  $V$  are generated independently of  $(X, Z)$ , with  $U \sim U(-0.25, 0.5)$  and  $V = U + 0.75$ . We take  $\lambda_1 = \lambda_2 = 3$  so the distribution of  $X$  is identifiable, except for the small proportion of values above  $b_V = 1.25$  ( $1 - F(b_V) \simeq 0.0023$ ). About 81% of the observations correspond to the event  $Z = 1$ . The simulated scenario is depicted in Figure 2, where it is seen that the naive estimator of Aalen and Johansen [1] which ignores the double truncation (grey lines) is biased.

The bias, standard deviation and mean squared error of the empirical CIFs  $F_{1n}(x)$  and  $F_{2n}(x)$  along 1000 Monte Carlo trials at the three quartiles of  $X$ , namely  $x_1 = 0.092$ ,  $x_2 = 0.211$  and  $x_3 = 0.393$ , are reported in Table 2. The sample sizes are  $n = 100$ ,  $n = 250$  and  $n = 500$ . From Table 2 it is seen that the estimators perform consistently, and that the bias is negligible compared to the standard deviation.

**Table 1.** Correlation/regression setting. Mean (standard deviation) of two estimators for the Pearson correlation coefficient and the regression parameters along 1000 Monte Carlo trials with sample size  $n$ . The naive estimators are  $\tilde{\rho}_{XZ}$ ,  $\tilde{\beta}_0$  and  $\tilde{\beta}_1$ , while the estimators adapted to double truncation are  $\hat{\rho}_{XZ}$ ,  $\hat{\beta}_0$  and  $\hat{\beta}_1$

	$n = 100$	$n = 250$	$n = 500$
$\tilde{\rho}_{XZ}$	0.925 (0.013)	0.925 (0.009)	0.925 (0.006)
$\tilde{\beta}_0$	0.038 (0.021)	0.039 (0.014)	0.038 (0.010)
$\tilde{\beta}_1$	0.923 (0.038)	0.923 (0.025)	0.923 (0.017)
$\hat{\rho}_{XZ}$	0.945 (0.010)	0.945 (0.007)	0.945 (0.005)
$\hat{\beta}_0$	0.002 (0.027)	0.001 (0.018)	0.000 (0.013)
$\hat{\beta}_1$	0.997 (0.048)	0.998 (0.033)	0.999 (0.025)



**Figure 2.** Cumulative incidences setting. True cumulative incidences in the simulated scenario (black lines) together with the naive Aalen–Johansen estimator which ignores the double truncation, computed for a single sample of size  $n = 10,000$  (grey lines).

### 5. Discussion

In this paper, Efron–Petrosian integrals for doubly truncated data with covariates have been introduced. These are empirical multivariate integrals with respect to the marginal NPMLE of the doubly truncated variable  $X$ . The main result is the consistency of the Efron–Petrosian integrals, together with an asymptotic i.i.d. representation which immediately entails convergence to a normal. Application of the main result to the asymptotic analysis of the Pearson correlation coefficient, least squares regression, Cox regression and cumulative incidence functions has been given. The finite sample performance of the Efron–Petrosian integrals has been investigated through simulations.

The asymptotic properties of the Efron–Petrosian integrals have been derived from those of its weighting process  $a_n(\cdot)$ . In its turn, the latter follow from the asymptotic properties of  $F_n$ , the marginal NPMLE of the df of  $X$ . In this paper, a formal and complete asymptotic analysis of  $F_n$  has been provided, filling some of the gaps in the related literature. In particular, primitive conditions under which

**Table 2.** Cumulative incidence setting. Bias, standard deviation (SD) and mean squared error (MSE) of  $\hat{F}_{1n}(x)$  and  $\hat{F}_{2n}(x)$  at the three quartiles  $x_1, x_2$  and  $x_3$  of  $X$  along 1000 replicates

		Bias		SD		MSE	
		$\hat{F}_{1n}$	$\hat{F}_{2n}$	$\hat{F}_{1n}$	$\hat{F}_{2n}$	$\hat{F}_{1n}$	$\hat{F}_{2n}$
$n = 100$	$x_1$	-0.0030	-0.0005	0.0604	0.0194	0.0037	0.0004
	$x_2$	-0.0033	0.0001	0.0554	0.0304	0.0031	0.0009
	$x_3$	-0.0018	0.0019	0.0459	0.0362	0.0021	0.0013
$n = 250$	$x_1$	0.0032	-0.0002	0.0396	0.0130	0.0016	0.0002
	$x_2$	0.0028	-0.0007	0.0364	0.0188	0.0013	0.0004
	$x_3$	0.0027	-0.0009	0.0295	0.0220	0.0009	0.0005
$n = 500$	$x_1$	0.0011	0.0001	0.0264	0.0093	0.0007	0.0001
	$x_2$	0.0012	0.0008	0.0246	0.0139	0.0006	0.0002
	$x_3$	0.0009	0.0008	0.0209	0.0156	0.0004	0.0002

$F_n$  follows a suitable i.i.d. asymptotic representation have been given, and the invertibility and boundedness of the involved operators have been established. The stated results are valid when the truncating couple  $(U, V)$  has a joint density and when it has not; this is interesting since the latter occurs with interval sampling, often encountered in practice.

From a practical viewpoint, a feature of much interest is the construction of confidence limits and hypothesis testing. These can be solved by using the asymptotic normality results in this paper if an estimator for the variance of the Efron–Petrosian integrals is available. The formal study of this issue is not covered in this paper; however, in principle one may apply the bootstrap (or alternative resampling procedures) to proceed. See Moreira and de Uña-Álvarez [15] for more on the bootstrap with doubly truncated data. The i.i.d. representation in Theorem 2.2 could serve to introduce an estimator for the variance too, in the spirit of Remark 2.4. This issue is currently under investigation.

Throughout the paper it has been assumed that the truncating variables are independent of the couple of interest  $(X, Z)$ ; indeed, most of the technical literature and real data applications have pivoted on such an assumption. Of course, in case of dependence between  $(U, V)$  and  $(X, Z)$  weights alternative to  $a_n(\cdot)$  must be introduced; see Moreira et al. [17] for an extended Efron–Petrosian estimator under dependent truncation. This is an interesting question which is left for future research.

### Appendix: Technical proofs

**Proof to Theorem 2.1.** Assume for a moment that Theorem 2.1(a) holds, that is  $\|F_n - F\|_\infty = o_P(1)$  (we will come back to this later). Write

$$\begin{aligned} & \mathcal{U}(F_n, F_n^*, K_n^*)(x) - \mathcal{U}(F_n, F^*, K^*)(x) \\ &= - \int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{\bar{a}_n(t)} + \int_{a_X}^x \int_{u \leq t \leq v} \frac{d(K_n^* - K^*)(u, v)}{F_n(v) - F_n(u-)} \frac{dF_n^*(t)}{a_n(t)\bar{a}_n(t)}. \end{aligned}$$

In Lemma A.1 we show that, under  $\|F_n - F\|_\infty = o_P(1)$ , the quantity above equals  $-h_n(x) + o_P(n^{-1/2})$  uniformly in  $x \in [a_X, b_X]$ , where

$$h_n(x) = \int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{a(t)} - \int_{a_X}^x \int_{u \leq t \leq v} \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)} \frac{dF^*(t)}{a(t)^2}.$$

Lemma A.2 proves the Fréchet differentiability of  $F \rightarrow \mathcal{U}(F, F^*, K^*)$ , that is, there exists a linear operator  $h \rightarrow \Gamma(F, F^*, K^*)[h]$  such that

$$\lim_{\|h\|_\infty \rightarrow 0} \frac{1}{\|h\|_\infty} \|\mathcal{U}(F + h, F^*, K^*) - \mathcal{U}(F, F^*, K^*) - \Gamma(F, F^*, K^*)[h]\|_\infty = 0.$$

Indeed, Lemma A.2 gives

$$\Gamma(F, F^*, K^*)[h](x) = h(x) - \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F(v) - F(u-))^2} dK^*(u, v) \frac{dF^*(t)}{a(t)^2}.$$

Choosing  $h = F_n - F$ , and since  $\mathcal{U}(F, F^*, K^*) \equiv 0 \equiv \mathcal{U}(F_n, F_n^*, K_n^*)$ , the results above lead to

$$\begin{aligned} \Gamma(F, F^*, K^*)[F_n - F](x) &= \mathcal{U}(F_n, F^*, K^*)(x) - \mathcal{U}(F, F^*, K^*)(x) + o_P(\|F_n - F\|_\infty) \\ &= h_n(x) + o_P(n^{-1/2}) + o_P(\|F_n - F\|_\infty) \end{aligned}$$

uniformly in  $x \in [a_X, b_X]$ .

The invertibility of the linear operator  $h \rightarrow \Gamma(F, F^*, K^*)[h]$  and the boundedness of its inverse  $\mathcal{A} = \Gamma(F, F^*, K^*)^{-1}$  are established in Lemma A.3, in which the linearity of  $\mathcal{A}$  is proved too. Now, for any linear operator  $\mathcal{B}$  it holds that  $\|\mathcal{B}[h]\|_\infty \leq \|\mathcal{B}\| \|h\|_\infty$ . Therefore, since  $\mathcal{A}$  is bounded, we obtain

$$F_n(x) - F(x) = \mathcal{A}[h_n](x) + o_P(n^{-1/2}) + o_P(\|F_n - F\|_\infty) \tag{A.1}$$

uniformly in  $x \in [a_X, b_X]$ . The weak convergence of the main term in (A.1) is established in Lemma A.4. Hence, by the continuous mapping theorem, this main term satisfies  $\|\mathcal{A}[h_n]\|_\infty = O_P(n^{-1/2})$ , from which it follows that  $\|F_n - F\|_\infty = O_P(n^{-1/2})$ . Then, the remainder in (A.1) is  $o_P(n^{-1/2})$  uniformly in  $x \in [a_X, b_X]$ . This proves Theorem 2.1(b) with

$$h_{X_i, U_i, V_i}(x) = \frac{I(X_i \leq x)}{a(X_i)} - E\left[\frac{I(X_1 \leq x)}{a(X_1)}\right] - \frac{I(U_i \leq x)}{F(V_i) - F(U_i-)} \int_{a_X \vee U_i}^{x \wedge V_i} \frac{dF^*(t)}{a(t)^2} + E\left[\frac{I(U_1 \leq x)}{F(V_1) - F(U_1-)} \int_{a_X \vee U_1}^{x \wedge V_1} \frac{dF^*(t)}{a(t)^2}\right].$$

Actually, the two expectations in this representation cancel out, and they both equal  $F(x)$ . This can be easily seen by using the relationships  $dF^*(t) = a(t) dF(t)$ ,  $dK^*(u, v) = \alpha^{-1}(F(v) - F(u-)) dK(u, v)$  and  $G(t) = \alpha a(t)$ . The proof of Theorem 2.1(b) is thus complete.

Now, in order to prove Theorem 2.1(a), note first that  $\mathcal{U}(F, F^*, K^*) \equiv 0$  is equivalent to  $H(F, F^*, K^*) \equiv 0$ , where

$$H(F, F^*, K^*)(x) = \int_{u \leq v} \frac{F(v \wedge x) - F((u-) \wedge x)}{F(v) - F(u-)} dK^*(u, v) - F^*(x).$$

By the argument in van der Laan [29], p. 122, if  $F$  has only a finite number of point masses then  $F_n$  has a subsequence which converges uniformly to a df  $F_\infty$  which has the same support as  $F$ . Let  $F_{n(\ell)}$  be such a subsequence. Lemma A.6 shows that, under (A1),  $F_\infty$  satisfies (A2) and, therefore,  $F_{n(\ell)}(v) - F_{n(\ell)}(u-)$  is bounded away from zero uniformly on the support of  $K^*$ . Conclude that  $H(F_{n(\ell)}, F_{n(\ell)}^*, K_{n(\ell)}^*)(x) - H(F_{n(\ell)}, F^*, K^*)(x)$  is bounded by  $C\|K_{n(\ell)}^* - K^*\|_\infty + \|F_{n(\ell)}^* - F^*\|_\infty$  uniformly in  $x \in [a_X, b_X]$  for a constant  $C$ . This leads to

$$\|H(F_{n(\ell)}, F^*, K^*)\|_\infty = \|H(F_{n(\ell)}, F_{n(\ell)}^*, K_{n(\ell)}^*) - H(F_{n(\ell)}, F^*, K^*)\|_\infty = O_P(n(\ell)^{-1/2}).$$

The uniform consistency of  $F_{n(\ell)}$  trivially implies  $\|H(F_{n(\ell)}, F^*, K^*) - H(F_\infty, F^*, K^*)\|_\infty = o_P(1)$  which shows that  $H(F_\infty, F^*, K^*)(x) = 0$  for  $x \in [a_X, b_X]$ . Now, since  $H(F, F^*, K^*) \equiv 0$  always holds, we get  $(I - A_\infty)[F_\infty - F] = 0$ , where

$$A_\infty[h](x) = \alpha^{-1} \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{F_\infty(v) - F_\infty(u-)} dK(u, v) \frac{dF(t)}{a_{F_\infty}(t)},$$

with

$$a_{F_\infty}(t) = \int_{u \leq t \leq v} \frac{dK^*(u, v)}{F_\infty(v) - F_\infty(u-)},$$

so  $I - A_\infty$  is analogous to the operator  $\Gamma$  above. Therefore, from Lemma A.3 we obtain the invertibility of  $I - A_\infty$ , where we need again that  $F_\infty$  satisfies (A2) (Lemma A.6). This implies the uniqueness of the solution of  $H(F, F^*, K^*) = 0$  and, consequently,  $F_\infty = F$  holds. Summarising, each subsequence of  $F_n$  has a uniformly convergent subsequence, all having the same limit  $F$ . This proves the uniform consistency of  $F_n$ .  $\square$



**Remark A.1.** Theorem 2.1(a) is related to Theorem 2 in Shen [24], p. 844, where a statement on the uniform consistency of  $F_n$  is given. Also, our Theorem 2.1(b) relates Theorem 3 in that paper. Nevertheless, the assumptions and reasonings in Shen [24] suffer from several important gaps and inconsistencies. For example, the convergence interval in Shen [24], Theorem 2, is of the form  $[a_X, \tau]$ , suggesting a potential issue at the right tail of  $F$ . This is confusing, since the denominator  $F(v) - F(u-)$  creates troubles for small  $v$ 's too. Indeed, our proof shows that there is no need to consider a consistency interval other than  $[a_X, b_X]$  as long as  $F(v) - F(u-)$  remains bounded away from zero, as stated by (A2). Shen [24] also refers to this boundedness condition in his Theorem 2, by assuming that  $F(v) - F(u-) > \delta > 0$  for  $[u, v] \in [a_X, \tau]$ ; but this is obviously impossible in his setting, where  $F$  has a density (just take  $u$  approaching  $v$  to violate Shen's condition). More importantly, in the proof to Theorem 2 in Shen [24] it is indicated that  $F_{n(\ell)}(v) - F_{n(\ell)}(u-)$  is uniformly bounded away from zero, where  $F_{n(\ell)}$  is a subsequence converging to  $F_\infty$ , with  $\text{supp}(F_\infty) = \text{supp}(F)$ . No justification is given. At this point the proof in Shen [24], Theorem 2, follows the argument in van der Laan [29], p. 122. However, the investigation of the aforementioned boundedness is far from trivial; see p. 116 in van der Laan [29] in which some artificial censoring is invoked and its possible impact on the NPMLE is discussed. We have circumvented this by proving and exploiting our Lemma A.6, which is also critical to justify that  $F_\infty$  equals  $F$ . Note that this Lemma A.6 is not for free; it requires in particular that  $v - u \geq \gamma > 0$  on the support of  $K^*$ , a condition which is missing in Shen [24].

On the other hand, our proof of Theorem 2.1(b) follows steps similar to those of Theorem 3 in Shen [24], pp. 845–847, who considered the case in which  $X$  and  $(U, V)$  have a density. However, unlike in the referred paper, we provide a formal proof of the Fréchet differentiability of  $F \rightarrow \mathcal{U}(F, F^*, K^*)$ ; see Lemma A.2. Similarly to Shen [24], it turns out that  $\mathcal{A}$  is the inverse of a certain linear operator  $\Gamma(F, F^*, K^*)$ , which is the derivative of  $\mathcal{U}$ . By deriving Volterra-type structures for  $\Gamma$ , we prove that such operator is invertible and that its inverse  $\mathcal{A}$  is bounded (Lemma A.3). We also obtain an explicit formula for  $\mathcal{A}$ , namely  $\mathcal{A} = \sum_{r=0}^\infty A^r$  where  $A[h](x) = \int_{a_X}^{b_X} L(x, w)h(w)dw$  for a certain function  $L(x, w)$ , see Lemma A.3. All these steps or results are missing in Shen [24]. Finally, compared to Theorem 3 in Shen [24], we provide primitive assumptions under which the i.i.d. representation in Theorem 2.1(b) holds true.

In the proof to his Theorem 3, Shen [24] says he can prove that the derivative of  $\mathcal{U}$  ‘has a bounded inverse in the same way as was in Section 3.4 of van der Laan [29]’ but, as mentioned, he does not provide any proof for that. In the setting with double truncation things are different from van der Laan [29]; for example, the Volterra structures appearing in the study of the invertibility of  $\Gamma(F, F^*, K^*)$  are more involved (both ‘direct’ and ‘reverse’ Volterra’s appear); we perform a thorough investigation of them in Lemma A.3. Summarising, one can say that the proof of our Theorem 2.1 fills important gaps in the proofs in Shen [24], while correcting and clarifying the needed assumptions in the statement of the results.

**Remark A.2.** As mentioned in Remark 2.3, in the left-truncated setting the i.i.d. representation  $\mathcal{A}[h_n](x)$  reduces to that in Stute [27], Theorem 2, which is  $n^{-1} \sum_{i=1}^n \tilde{h}_i(x)$  with

$$\tilde{h}_i(x) = \alpha(1 - F(x)) \left[ \frac{I(X_i \leq x)}{G(X_i)(1 - F(X_i))} - I(U_i \leq x) \int_{a_X \vee U_i}^x \frac{I(X_i \geq t) dF(t)}{G(t)(1 - F(t))^2} \right],$$

where the continuity of  $F$  is assumed for ease of presentation. Here  $G$  denotes the df of the left-truncating variable  $U$ , which equals  $P(U \leq x \leq V)$  when  $P(V \geq b_X) = 1$ . Since  $\Gamma(F, F^*, K^*) = \mathcal{A}^{-1}$ , the result follows provided that  $\Gamma(F, F^*, K^*)[\tilde{h}_i] = h_{X_i, U_i, V_i}$ ,  $1 \leq i \leq n$ . Now, in order to see this, note first that when  $P(V \geq b_X) = 1$  the operator  $\Gamma(F, F^*, K^*)$  reduces to

$$\Gamma[h](x) = h(x) + \int_{a_X}^x \int_{u \leq t} \frac{h(u) dG(u) dF(t)}{(1 - F(u))G(t)}$$

and  $G$  becomes the df of  $U$ . We have

$$\begin{aligned} & \int_{a_X}^x \int_{u \leq t} \frac{\tilde{h}_i(u) dG(u) dF(t)}{(1 - F(u))G(t)} \\ &= \int_{a_X}^x \int_{u \leq t} \frac{\alpha I(X_i \leq u) dG(u) dF(t)}{G(X_i)(1 - F(X_i))G(t)} \\ & \quad - \int_{a_X}^x \int_{u \leq t} \frac{\alpha I(U_i \leq u)}{G(t)} \int_{a_X \vee U_i}^u \frac{I(X_i \geq s) dF(s)}{G(s)(1 - F(s))^2} dG(u) dF(t) \equiv I - II. \end{aligned}$$

Straightforward calculations lead to

$$I = \frac{\alpha I(X_i \leq x)}{1 - F(X_i)} \left[ \frac{F(x) - F(X_i)}{G(X_i)} - \int_{X_i}^x \frac{dF(t)}{G(t)} \right]$$

while  $II$  equals

$$\alpha I(U_i \leq x) \left[ \int_{a_X \vee U_i}^{x \wedge X_i} \frac{(F(x) - F(s)) dF(s)}{G(s)(1 - F(s))^2} - \int_{a_X \vee U_i}^x \left( \frac{1}{1 - F(t \wedge X_i)} - \frac{1}{1 - F(U_i)} \right) \frac{dF(t)}{G(t)} \right].$$

The calculations for  $II$  involve a triple integral; the expression above is obtained by integrating first with respect to  $dG(u)$  and then with respect to  $dF(t)$ . By collecting and simplifying the several terms in  $\tilde{h}_i(x) + I - II$  one finally comes up with

$$\Gamma[\tilde{h}_i](x) = \alpha \frac{I(X_i \leq x)}{G(X_i)} - \alpha \frac{I(U_i \leq x)}{1 - F(U_i)} \int_{a_X \vee U_i}^x \frac{dF(t)}{G(t)},$$

which is just  $h_{X_i, U_i, V_i}(x)$  in the case  $P(V \geq b_X) = 1$ . This completes the proof of our assertion.

**Lemma A.1.** Assume  $\|F_n - F\|_\infty = o_P(1)$ . Under (A1)–(A4), we have

$$\int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{\bar{a}_n(t)} = \int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{a(t)} + o_P(n^{-1/2}),$$

and

$$\begin{aligned} & \int_{a_X}^x \int_{u \leq t \leq v} \frac{d(K_n^* - K^*)(u, v)}{F_n(v) - F_n(u-)} \frac{dF_n^*(t)}{a_n(t)\bar{a}_n(t)} \\ &= \int_{a_X}^x \int_{u \leq t \leq v} \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)} \frac{dF^*(t)}{a(t)^2} + o_P(n^{-1/2}) \end{aligned}$$

uniformly in  $x \in [a_X, b_X]$ .

**Proof.** For the first assertion we need to prove that

$$\int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{\bar{a}_n(t)} - \int_{a_X}^x \frac{d(F_n^* - F^*)(t)}{a(t)} = o_P(n^{-1/2}) \tag{A.2}$$

uniformly in  $x \in [a_X, b_X]$ . This follows if

$$\sup_{x \in [a_X, b_X]} |\mathbb{G}_n^*(f_{x, \eta_n} - f_{x, \eta_0})| = o_P(1), \tag{A.3}$$

where  $\mathbb{G}_n^* f = \sqrt{n}(P_n^* - P^*)f$ , and where  $P^* f = \int f dF^*$ ,  $P_n^* f = \int f dF_n^*$ ,  $f_{x,\eta}(t) = I(t \leq x)\eta(t)$ ,  $\eta_n(t) = 1/\bar{a}_n(t)$  and  $\eta_0(t) = 1/a(t)$ . Now, by Theorem 2.1 in van der Vaart and Wellner [32], (A.3) holds provided that:

- (i)  $P(\eta_n \in H_0) \rightarrow 1$  for a certain set of  $\eta$ -functions  $H_0$
- (ii)  $\{f_{x,\eta} : x \in [a_X, b_X], \eta \in H_0\}$  is Donsker
- (iii)  $\sup_{x \in [a_X, b_X]} \int (f_{x,\eta_n}(t) - f_{x,\eta_0}(t))^2 dF^*(t) = o_P(1)$

Note that condition (iii) holds under (A1)–(A4). Now, introduce the set of functions

$$H_0 = \left\{ t \mapsto \left[ \int_{u \leq t \leq v} \frac{dK^*(u, v)}{F(v) - F(u-)} \right]^{-1} : F \in \mathcal{F}^{\delta_1, \delta_2} \right\},$$

where  $\mathcal{F}^{\delta_1, \delta_2}$  is the class of df's  $F$  satisfying

$$\inf_{(u, v) \in \text{supp}(K^*)} (F(v) - F(u-)) \geq \delta_1/2, \quad \inf_{a_X \leq x \leq b_X} \int_{u \leq x \leq v} \frac{dK^*(u, v)}{F(v) - F(u-)} \geq \delta_2/(3\alpha).$$

In order to verify (ii), note that the integral over the set  $\{(u, v) : u \leq t \leq v\}$  can be written as a difference of monotone, bounded functions (integrate over  $\{(u, v) : u \leq t\}$  and then subtract the integral over  $\{(u, v) : v < t\}$ ), which are Donsker (van der Vaart and Wellner [31]). The fact that indicator functions are Donsker, and that the differences, products and inverses of (bounded away from zero) Donsker classes are Donsker (same reference) allows then to conclude.

The verification of (i) is more subtle. Take  $\epsilon = \delta_1/4$  and write

$$P(\eta_n \in H_0) = P(\eta_n \in H_0 \| \|F_n - F\|_\infty < \epsilon) P(\|F_n - F\|_\infty < \epsilon) + P(\eta_n \in H_0, \|F_n - F\|_\infty \geq \epsilon).$$

The second term is bounded from above by  $P(\|F_n - F\|_\infty \geq \epsilon)$ , which goes to zero because of Theorem 2.1(a). Similarly, the factor  $P(\|F_n - F\|_\infty < \epsilon)$  goes to one. Now, it is straightforward to see that, under (A2)–(A3), the conditional probability above equals one. Certainly,  $F_n$  is a df and, given  $\|F_n - F\|_\infty < \epsilon$ , for each  $(u, v) \in \text{supp}(K^*)$  we have under (A2)

$$F_n(v) - F_n(u-) = F_n(v) - F(v) + F(v) - F(u-) + F(u-) - F_n(u-) > \delta_1 - 2\epsilon = \delta_1/2.$$

On the other hand, given  $\|F_n - F\|_\infty < \epsilon$ ,

$$\int_{u \leq x \leq v} \frac{dK^*(u, v)}{F_n(v) - F_n(u-)} \geq \int_{u \leq x \leq v} \frac{dK^*(u, v)}{F(v) - F(u-) + 2\epsilon}.$$

Since  $\epsilon < \delta_1 \leq F(v) - F(u-)$  for each  $(u, v) \in \text{supp}(K^*)$ , the right-hand side is bounded from below by

$$\frac{1}{3} \int_{u \leq x \leq v} \frac{dK^*(u, v)}{F(v) - F(u-)} > \delta_2/(3\alpha),$$

where the last inequality follows from (A3) and  $G(x) = \alpha \int_{u \leq x \leq v} dK^*(u, v)/(F(v) - F(u-))$ . This shows that  $P(\eta_n \in H_0 \| \|F_n - F\|_\infty < \epsilon) = 1$  and thus the proof of the first assertion is finished.

For the second assertion we proceed in a similar way. Note that the aim is to prove that

$$\int_{u \leq x \wedge v} \left( \int_{a_X \vee u}^{x \wedge v} \frac{dF_n^*(t)}{a_n(t)\bar{a}_n(t)} - \int_{a_X \vee u}^{x \wedge v} \frac{dF^*(t)}{a(t)^2} \right) \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)} = o_P(n^{-1/2}).$$

This term can be written as

$$\begin{aligned} & \int_{a_X}^x \left[ \frac{1}{\bar{a}_n(t)} - \frac{1}{a_n(t)} \right] dF_n^*(t) - \int_{a_X}^x \left[ \frac{\tilde{a}_n(t)}{a(t)^2} - \frac{1}{a(t)} \right] dF^*(t) \\ &= \int_{a_X}^x \left[ \frac{1}{\bar{a}_n(t)} - \frac{1}{a(t)} \right] d(F_n^* - F^*)(t) - \int_{a_X}^x \left[ \frac{1}{a_n(t)} - \frac{1}{a(t)} \right] d(F_n^* - F^*)(t) \\ & \quad + \int_{a_X}^x \left[ \frac{1}{\bar{a}_n(t)} - \frac{1}{a_n(t)} - \frac{\tilde{a}_n(t)}{a(t)^2} + \frac{1}{a(t)} \right] dF^*(t) \\ & \equiv I + II + III, \end{aligned}$$

where, using the notation of Section 2,

$$\tilde{a}_n(t) = \int_{u \leq t \leq v} \frac{dK_n^*(u, v)}{F(v) - F(u-)}.$$

The first integral is just (A.2) and thus  $I = o_P(n^{-1/2})$  uniformly in  $x$ . For  $II$  we apply arguments similar to those above; with  $\eta_n(t) = 1/a_n(t)$  and for the same class  $H_0$  of functions, one gets (i)–(iii) above under conditions (A1)–(A4). In this case, to prove (i) one can use that

$$P(\eta_n \in H_0 \mid \|F_n - F\|_\infty < \epsilon, \|\tilde{a}_n - a\|_\infty < \epsilon) = 1 \tag{A.4}$$

for  $\epsilon < \min\{\delta_1/4, \delta_2/(2\alpha)\}$ . Certainly, note that, conditionally on  $\|F_n - F\|_\infty < \epsilon$  with  $\epsilon \leq \delta_1/4$ ,

$$\begin{aligned} \int_{u \leq x \leq v} \frac{dK_n^*(u, v)}{F_n(v) - F_n(u-)} &\geq \int_{u \leq x \leq v} \frac{dK_n^*(u, v)}{F(v) - F(u-) + 2\epsilon} \\ &\geq \frac{2}{3} \int_{u \leq x \leq v} \frac{dK_n^*(u, v)}{F(v) - F(u-)} = \frac{2}{3} \tilde{a}_n(x). \end{aligned}$$

Now, under  $\|\tilde{a}_n - a\|_\infty < \epsilon < \delta_2/(2\alpha)$ , and since  $a(x) > \delta_2/\alpha$  holds under (A3), the quantity  $2\tilde{a}_n(x)/3$  can be bounded from below by

$$\frac{2\delta_2}{3\alpha} - \frac{3\epsilon}{2} \geq \frac{2\delta_2}{3\alpha} - \frac{\delta_2}{3\alpha} = \frac{\delta_2}{3\alpha},$$

so (A.4) holds. This shows that  $II = o_P(n^{-1/2})$  uniformly in  $x$ . Finally, note that

$$III = \int_{a_X}^x \frac{a_n(t) - \bar{a}_n(t)}{a_n(t)\bar{a}_n(t)} dF^*(t) - \int_{a_X}^x \frac{\tilde{a}_n(t) - a(t)}{a(t)^2} dF^*(t).$$

Now, it holds that  $a_n(t) - \bar{a}_n(t) = \tilde{a}_n(t) - a(t) + o_P(n^{-1/2})$  uniformly in  $t$ . This can be proved by applying the aforementioned Theorem 2.1 of van der Vaart and Wellner [32] with  $f_{x,\eta}(u, v) = I(u \leq x \leq v)\eta(u, v)$ ,  $\eta_n(u, v) = (F_n(v) - F_n(u-))^{-1}$ ,  $\eta_0(u, v) = (F(v) - F(u-))^{-1}$  and the class of functions

$$H_0 = \{(u, v) \mapsto 1/(F(v) - F(u-)) : F \in \mathcal{F}^{\delta_1, 0}\},$$

where  $\mathcal{F}^{\delta_1, 0}$  is formed by the df's  $F$  such that  $\inf_{(u,v) \in \text{supp}(K^*)} (F(v) - F(u-)) \geq \delta_1/2$  (details are omitted). Thus

$$III = \int_{a_X}^x [\tilde{a}_n(t) - a(t)] \left[ \frac{1}{a_n(t)\tilde{a}_n(t)} - \frac{1}{a(t)^2} \right] dF^*(t) + o_P(n^{-1/2}),$$

uniformly in  $x$ . Under (A1)–(A4), one easily gets that  $\|1/(a_n\tilde{a}_n) - 1/a^2\|_\infty = o_P(1)$ . Furthermore, since under (A2) the class of functions  $\{(u, v) \mapsto I(u \leq t \leq v)(F(v) - F(u-))^{-1} : t \in [a_X, b_X]\}$  is Donsker, we have

$$\tilde{a}_n(t) - a(t) = \int_{u \leq t \leq v} \frac{d(K_n^* - K^*)(u, v)}{F(v) - F(u-)} = O_P(n^{-1/2})$$

uniformly in  $t$ . Hence,  $III = o_P(n^{-1/2})$  uniformly in  $x$  and the proof is complete.  $\square$

**Lemma A.2.** *Under (A2) and (A3), we have*

$$\lim_{\|h\|_\infty \rightarrow 0} \frac{1}{\|h\|_\infty} \|\mathcal{U}(F+h, F^*, K^*) - \mathcal{U}(F, F^*, K^*) - \Gamma(F, F^*, K^*)[h]\|_\infty = 0,$$

where

$$\Gamma(F, F^*, K^*)[h](x) = h(x) - \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F(v) - F(u-))^2} dK^*(u, v) \frac{dF^*(t)}{a(t)^2}.$$

**Proof.** We have

$$\begin{aligned} & (\mathcal{U}(F+h, F^*, K^*) - \mathcal{U}(F, F^*, K^*) - \Gamma(F, F^*, K^*)[h])(x) \\ &= - \int_{a_X}^x \left[ \frac{1}{a^{(h)}(t)} - \frac{1}{a(t)} \right] dF^*(t) + \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F(v) - F(u-))^2} dK^*(u, v) \frac{dF^*(t)}{a(t)^2}, \end{aligned}$$

where

$$a^{(h)}(t) = \int_{u \leq t \leq v} \frac{dK^*(u, v)}{(F+h)(v) - (F+h)(u-)}.$$

Now, since

$$\frac{1}{a^{(h)}(t)} - \frac{1}{a(t)} = \frac{1}{a^{(h)}(t)a(t)} \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{F(v) - F(u-)} \frac{dK^*(u, v)}{(F+h)(v) - (F+h)(u-)},$$

we get

$$\begin{aligned} & (\mathcal{U}(F+h, F^*, K^*) - \mathcal{U}(F, F^*, K^*) - \Gamma(F, F^*, K^*)[h])(x) \\ &= - \int_{a_X}^x \frac{1}{a(t)^2} \left[ \frac{a(t)}{a^{(h)}(t)} - 1 \right] \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{F(v) - F(u-)} \frac{dK^*(u, v)}{(F+h)(v) - (F+h)(u-)} dF^*(t) \\ & \quad - \int_{a_X}^x \frac{1}{a(t)^2} \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F(v) - F(u-))^2} \left[ \frac{F(v) - F(u-)}{(F+h)(v) - (F+h)(u-)} - 1 \right] dK^*(u, v) dF^*(t). \end{aligned}$$

Clearly, under (A2) and (A3) both terms go to zero uniformly in  $x$  as  $\|h\|_\infty \rightarrow 0$  when divided by  $\|h\|_\infty$ , which completes the proof.  $\square$

**Lemma A.3.** *Under (A2)–(A4), the linear operator  $h \rightarrow \Gamma(F, F^*, K^*)[h]$  is invertible, and its inverse  $\mathcal{A} = \Gamma(F, F^*, K^*)^{-1}$  is a linear bounded operator.*

**Proof.** Write  $\Gamma(F, F^*, K^*) = I - A$ , where

$$A[h](x) = \int_{a_X}^x \int_{u \leq t \leq v} \frac{h(v) - h(u-)}{(F(v) - F(u-))^2} dK^*(u, v) \frac{dF^*(t)}{a(t)^2}.$$

To show that  $\Gamma \equiv I - A$  is invertible we will first show that  $A[h]$  is a linear operator which involves Volterra-type structures. Note that we can restrict our attention to functions  $h$  that vanish outside the interval  $[a_X, b_X]$ , since  $h = F_n - F$  lives on that interval. For  $x, w \in [a_X, b_X]$  introduce

$$L(x, w) = E \left[ \frac{I(X_2 \leq x)I(U_1 \leq X_2 \leq V_1)}{(F(V_1) - F(U_1-))^2 a(X_2)^2} \middle| V_1 = w \right] k_{V_1}(w) \\ - E \left[ \frac{I(X_2 \leq x)I(U_1 \leq X_2 \leq V_1)}{(F(V_1) - F(U_1-))^2 a(X_2)^2} \middle| U_1 = w \right] k_{U_1}(w),$$

where  $k_{U_1}$  and  $k_{V_1}$  stand for the probability density functions of  $U_1$  and  $V_1$  respectively. Then, we have

$$A[h](x) = \int_{a_X}^{b_X} L(x, w)h(w) dw = \sum_{j=1}^2 \int_{a_X}^{b_X} L_j(x, w)h(w) dw \equiv \sum_{j=1}^2 A_j[h](x),$$

where  $L_1(x, w) \equiv L(x, w)I(x \geq w)$  and  $L_2(x, w) \equiv L(x, w)I(x < w)$ . It is clear that  $A, A_1$  and  $A_2$  are linear operators, and that  $A_1$  has a Volterra structure in the sense of Kantorovich and Akilov [13], p. 396, because  $L_1(x, w)$  vanishes for  $x < w$ . Furthermore, under (A2)–(A4) the function  $L$  is bounded. That is, there exists a constant  $M$  such that  $\sup_{a_X \leq x, w \leq b_X} |L(x, w)| \leq M$ ; this constant  $M$  serves as a bound for  $L_1$  and  $L_2$  too. Following Kantorovich and Akilov [13], p. 396, we have for  $j = 1$ ,

$$\|A_j^r[h]\|_\infty \leq \frac{M^r (b_X - a_X)^r}{(r - 1)!} \|h\|_\infty, \quad r \geq 1. \tag{A.5}$$

Interestingly, the arguments also work for the operator  $A_2$ , since it involves the function  $L_2(x, w)$  which fulfills the ‘reverse’ Volterra condition  $L_2(x, w) = 0$  for  $x \geq w$ . That is, (A.5) holds for  $j = 2$  too. Now, an application of the binomial formula gives

$$\|(A_1 + A_2)^r[h]\|_\infty \leq \sum_{k=0}^r \binom{r}{k} \|A_1^k[h]\|_\infty \|A_2^{r-k}[h]\|_\infty.$$

Use (A.5) to get  $\|(A_1 + A_2)^r[h]\|_\infty \leq C_r \|h\|_\infty^2$  where  $C_r \equiv (1 + M)^{2r}/r!$ . Recall  $A = A_1 + A_2$  so one has  $\|A^r[h]\|_\infty \leq (1 + M)^{2r}/r!$  on the class of functions  $h$  with  $\|h\|_\infty = 1$ . Since  $A^r$  is linear, we get

$$\|A^r[h]\|_\infty \leq C_r \|h\|_\infty, \quad r \geq 0.$$

Finally, since  $\sum_{r=0}^\infty C_r < \infty$ , we conclude that  $\Gamma = I - A$  is invertible and that  $\mathcal{A} = (I - A)^{-1} = \sum_{r=0}^\infty A^r$  is its inverse operator, which is linear and bounded. This completes the proof of the Lemma.  $\square$

**Lemma A.4.** *Under (A2)–(A4), the process  $\sqrt{n}A[h_n](x), a_X \leq x \leq b_X$ , converges weakly to a zero-mean Gaussian process  $\mathcal{W}(x)$  with covariance function given in Corollary 2.1.*

**Proof.** Since  $\mathcal{A}$  is linear, it is Hadamard differentiable and its derivative is  $\mathcal{A}$ . Thus, it suffices to show that the class of functions

$$\mathcal{G} = \{(t, u, v) \rightarrow h_{t,u,v}(x) : a_X \leq x \leq b_X\}$$

is Donsker (see Section 3.9 in van der Vaart and Wellner [31], on the Delta method). Here,  $h_{t,u,v}(x)$  is the function introduced in Theorem 2.1, that is,

$$h_{t,u,v}(x) = \frac{I(t \leq x)}{a(t)} - \frac{I(u \leq x)}{F(v) - F(u-)} \int_{a_X \vee u}^{x \wedge v} \frac{dF^*(s)}{a(s)^2}.$$

Now, note that  $\{t \rightarrow I(t \leq x) : a_X \leq x \leq b_X\}$  is Donsker;  $t \rightarrow 1/a(t)$  is a bounded function under (A3);  $(u, v) \rightarrow 1/(F(v) - F(u-))$  is a bounded function under (A2); the class  $\{u \rightarrow \int_{a_X}^{x \wedge u} F^*(ds)/a(s)^2 : a_X \leq x \leq b_X\}$  is Donsker because the functions are increasing and bounded. On the other hand, for  $u \leq x$ ,

$$\int_{a_X \vee u}^{x \wedge v} \frac{dF^*(s)}{a(s)^2} = \int_{a_X}^{x \wedge v} \frac{dF^*(s)}{a(s)^2} - \int_{a_X}^{a_X \vee u} \frac{dF^*(s)}{a(s)^2},$$

and the sums and products of (uniformly bounded) Donsker classes are Donsker. Therefore  $\mathcal{G}$  is Donsker. This completes the proof.  $\square$

**Lemma A.5.** Under (A2)–(A4), the leading process in Theorem 2.2 converges weakly to a zero-mean Gaussian process  $\mathcal{W}_0(x)$  with covariance function given in Corollary 2.2.

**Proof.** The first part of the leading process in Theorem 2.2 is weakly convergent because the class of functions  $(u, v) \rightarrow I(u \leq x \leq v)/(F(v) - F(u-))$ ,  $a_X \leq x \leq b_X$ , is Donsker. Consider now the term involving the linear operator  $\tilde{\mathcal{A}}$ . Since  $\mathcal{A}$  is linear, the operator  $\tilde{\mathcal{A}}$  is linear too. Then, by the Delta method, the weak convergence of the process  $\tilde{\mathcal{A}}[n^{-1/2} \sum_{i=1}^n h_{X_i, U_i, V_i}](x)$ ,  $a_X \leq x \leq b_X$ , follows from that of the process  $n^{-1/2} \sum_{i=1}^n h_{X_i, U_i, V_i}(x)$ ,  $a_X \leq x \leq b_X$ . The latter holds because the class  $\mathcal{G}$  in Lemma A.4 is Donsker.  $\square$

**Lemma A.6.** Let  $F_\infty$  be any df with the same support as  $F$ . Under (A1),  $F_\infty$  satisfies (A2).

**Proof.** Assume for a moment that  $F$  is continuous with positive density on  $(a_X, b_X)$ . Under (A1), choose  $\gamma_0 < \min\{\gamma/2, a_{V_1} - a_X, b_X - b_{U_1}\}$  and  $x_\ell = x_{\ell-1} + \gamma_0$ ,  $1 \leq \ell \leq L(\gamma_0)$ , such that  $x_0 = a_X$  and  $x_L = b_X$ . Then, given  $(u, v) \in \text{supp}(K^*)$ , we have  $v - u \geq \gamma > 2\gamma_0$  and there exists  $t \in \text{supp}(F)$  with  $u \leq t \leq v$ , so it holds  $F(v) - F(u-) \geq F(x_\ell) - F(x_{\ell-1}) \equiv J_\ell$  for some  $\ell \in \{1, \dots, L(\gamma_0)\}$ . Then (A2) holds for  $\delta_1 = \min_{1 \leq \ell \leq L(\gamma_0)} J_\ell > 0$ . Note that we have implicitly used  $\gamma_0 < a_{V_1} - a_X$  (resp.  $\gamma_0 < b_X - b_{U_1}$ ) to ensure that  $F(v)$  (resp.  $F(u-)$ ) remains away from 0 (resp. 1). If  $F$  is continuous but  $\text{supp}(F)$  is not convex, the result is proved similarly by considering separately the subintervals of  $(a_X, b_X)$  on which the density of  $F$  is positive. Finally, the possible discontinuity points of  $F$  create no difficulties. In order to see this note that, if  $(u, v) \in \text{supp}(K^*)$  is such that  $[u, v] \cap \text{supp}(F)$  contains a mass point  $y_m$  of  $F$ , then obviously  $F(v) - F(u-) \geq F\{y_m\}$ , so (A2) holds with  $\delta_1 = \min\{\min_m F\{y_m\}, \min_{1 \leq \ell \leq L(\gamma_0)} J_\ell\} > 0$ . Now, if  $F_\infty$  is any df with the same support as  $F$  argument above applies, and the proof is complete.  $\square$

## Acknowledgements

The authors thank anonymous Associate Editor and referees for comments and suggestions which have served to improve the presentation of the paper. The first author acknowledges financial support by the Grant MTM2017-89422-P (MINECO/AEI/FEDER, UE). He also acknowledges financial support from the Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019-2022) and the EU (ERDF), Ref. ED431G2019/06. The second author is financially supported by the European Research Council (2016-2021, Horizon 2020/ERC grant agreement No. 694409).

## References

- [1] Aalen, O.O. and Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5** 141–150. [MR0509450](#)
- [2] Austin, M.D., Simon, D.K. and Betensky, R.A. (2014). Computationally simple estimation and improved efficiency for special cases of double truncation. *Lifetime Data Anal.* **20** 335–354. [MR3217541](#) <https://doi.org/10.1007/s10985-013-9287-z>
- [3] Bai, Z.D., Rao, C.R. and Wu, Y. (1992).  $M$ -estimation of multivariate linear regression parameters under a convex discrepancy function. *Statist. Sinica* **2** 237–254. [MR1152307](#)
- [4] de Uña-Álvarez, J. (2003). Large sample results under biased sampling when covariables are present. *Statist. Probab. Lett.* **63** 287–293. [MR1986328](#) [https://doi.org/10.1016/S0167-7152\(03\)00093-2](https://doi.org/10.1016/S0167-7152(03)00093-2)
- [5] de Uña-Álvarez, J. (2020). Nonparametric estimation of the cumulative incidences of competing risks under double truncation. *Biom. J.* **62** 852–867. <https://doi.org/10.1002/bimj.201800323>
- [6] Du, J., Zhang, Z. and Xie, T. (2018). A weighted  $M$ -estimator for linear regression models with randomly truncated data. *Statist. Probab. Lett.* **138** 90–94. [MR3788723](#) <https://doi.org/10.1016/j.spl.2018.02.055>
- [7] Efron, B. and Petrosian, V. (1999). Nonparametric methods for doubly truncated data. *J. Amer. Statist. Assoc.* **94** 824–834. [MR1723343](#) <https://doi.org/10.2307/2669997>
- [8] Emura, T., Hu, Y.-H. and Konno, Y. (2017). Asymptotic inference for maximum likelihood estimators under the special exponential family with double-truncation. *Statist. Papers* **58** 877–909. [MR3686855](#) <https://doi.org/10.1007/s00362-015-0730-y>
- [9] Frank, G., Chae, M. and Kim, Y. (2019). Additive time-dependent hazard model with doubly truncated data. *J. Korean Statist. Soc.* **48** 179–193. [MR3958707](#) <https://doi.org/10.1016/j.jkss.2018.10.005>
- [10] Gürlér, Ü. (1996). Bivariate estimation with right-truncated data. *J. Amer. Statist. Assoc.* **91** 1152–1165. [MR1424615](#) <https://doi.org/10.2307/2291734>
- [11] Gürlér, Ü., Stute, W. and Wang, J.-L. (1993). Weak and strong quantile representations for randomly truncated data with applications. *Statist. Probab. Lett.* **17** 139–148. [MR1223536](#) [https://doi.org/10.1016/0167-7152\(93\)90008-7](https://doi.org/10.1016/0167-7152(93)90008-7)
- [12] Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.* **40** 633–643. [MR0238419](#) <https://doi.org/10.1214/aoms/1177697731>
- [13] Kantorovich, L.V. and Akilov, G.P. (1982). *Functional Analysis*, 2nd ed. Oxford-Elmsford, NY: Pergamon Press. [MR0664597](#)
- [14] Mandel, M., de Uña-Álvarez, J., Simon, D.K. and Betensky, R.A. (2018). Inverse probability weighted Cox regression for doubly truncated data. *Biometrics* **74** 481–487. [MR3825334](#) <https://doi.org/10.1111/biom.12771>
- [15] Moreira, C. and de Uña-Álvarez, J. (2010). Bootstrapping the NPMLE for doubly truncated data. *J. Nonparametr. Stat.* **22** 567–583. [MR2682208](#) <https://doi.org/10.1080/10485250903556102>
- [16] Moreira, C. and de Uña-Álvarez, J. (2010). A semiparametric estimator of survival for doubly truncated data. *Stat. Med.* **29** 3147–3159. [MR2758709](#) <https://doi.org/10.1002/sim.3938>
- [17] Moreira, C., de Uña-Álvarez, J. and Braekers, R. (2018). Nonparametric estimation of a distribution function from doubly truncated data under dependence. Available at [arXiv:1802.08579v1](https://arxiv.org/abs/1802.08579v1).
- [18] Moreira, C., de Uña-Álvarez, J. and Crujeiras, R. (2010). DTDA: An R package to analyze randomly truncated data. *J. Stat. Softw.* **37** 1–20.



- [19] Moreira, C., de Uña-Álvarez, J. and Van Keilegom, I. (2014). Goodness-of-fit tests for a semiparametric model under random double truncation. *Comput. Statist.* **29** 1365–1379. MR3266063 <https://doi.org/10.1007/s00180-014-0496-z>
- [20] Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7** 186–199. MR1128411 <https://doi.org/10.1017/S0266466600004394>
- [21] Rennert, L. and Xie, S.X. (2018). Cox regression model with doubly truncated data. *Biometrics* **74** 725–733. MR3825359 <https://doi.org/10.1111/biom.12809>
- [22] Sánchez Sellero, C., González Manteiga, W. and Van Keilegom, I. (2005). Uniform representation of product-limit integrals with applications. *Scand. J. Stat.* **32** 563–581. MR2232343 <https://doi.org/10.1111/j.1467-9469.2005.00453.x>
- [23] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. MR0595165
- [24] Shen, P. (2010). Nonparametric analysis of doubly truncated data. *Ann. Inst. Statist. Math.* **62** 835–853. MR2669740 <https://doi.org/10.1007/s10463-008-0192-2>
- [25] Shen, P. and Liu, Y. (2019). Pseudo maximum likelihood estimation for the Cox model with doubly truncated data. *Statist. Papers* **60** 1207–1224. MR4008684 <https://doi.org/10.1007/s00362-016-0870-8>
- [26] Shen, P.-S. (2010). Semiparametric analysis of doubly truncated data. *Comm. Statist. Theory Methods* **39** 3178–3190. MR2755432 <https://doi.org/10.1080/03610920903219272>
- [27] Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21** 146–156. MR1212170 <https://doi.org/10.1214/aos/1176349019>
- [28] Stute, W. (1999). Nonlinear censored regression. *Statist. Sinica* **9** 1089–1102. MR1744826
- [29] van der Laan, M.J. (1996). Nonparametric estimation of the bivariate survival function with truncated data. *J. Multivariate Anal.* **58** 107–131. MR1408131 <https://doi.org/10.1006/jmva.1996.0042>
- [30] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 <https://doi.org/10.1017/CBO9780511802256>
- [31] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Series in Statistics*. New York: Springer. MR1385671 <https://doi.org/10.1007/978-1-4757-2545-2>
- [32] van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942 <https://doi.org/10.1214/074921707000000382>
- [33] Woodroffe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13** 163–177. MR0773160 <https://doi.org/10.1214/aos/1176346584>
- [34] Xiao, J. and Hudgens, M.G. (2019). On nonparametric maximum likelihood estimation with double truncation. *Biometrika* **106** 989–996. MR4031212 <https://doi.org/10.1093/biomet/asz038>
- [35] Ying, Z., Yu, W., Zhao, Z. and Zheng, M. (2020). Regression analysis of doubly truncated data. *J. Amer. Statist. Assoc.* **115** 810–821. MR4107681 <https://doi.org/10.1080/01621459.2019.1585252>
- [36] Zhu, H. and Wang, M.-C. (2012). Analysing bivariate survival data with interval sampling and application to cancer epidemiology. *Biometrika* **99** 345–361. MR2931258 <https://doi.org/10.1093/biomet/ass009>
- [37] Zhu, H. and Wang, M.-C. (2014). Nonparametric inference on bivariate survival data with interval sampling: Association estimation and testing. *Biometrika* **101** 519–533. MR3254898 <https://doi.org/10.1093/biomet/asu005>

Received November 2019 and revised May 2020