Journal of Multivariate Analysis 124 (2014) 10–24

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Asymptotically efficient estimation under semi-parametric random censorship models

Gerhard Dikta

Fachhochschule Aachen, Heinrich-Mußmann Strasse 1, D-52428 Jülich, Germany

ARTICLE INFO

Article history:
Received 7 May 2013
Available online 29 October 2013

AMS 2010 subject classifications:
62n01
62n02
62E20

Keywords:
Survival analysis
Censored data
Semi-parametric random censorship model
Asymptotic efficiency

ABSTRACT

We study the estimation of some linear functionals which are based on an unknown lifetime distribution. The observations are assumed to be generated under the semi-parametric random censorship model (SRCM), that is, a random censorship model where the conditional expectation of the censoring indicator given the observation belongs to a parametric family. Under this setup a semi-parametric estimator of the survival function was introduced by the author. If the parametric model assumption is correct, it is known that the estimated functional which is based on this semi-parametric estimator is asymptotically at least as efficient as the corresponding one which rests on the nonparametric Kaplan–Meier estimator.

In this paper we show that the estimated functional which is based on this semi-parametric estimator is asymptotically efficient with respect to the class of all regular estimators under this semi-parametric model.

1. Introduction

To analyze failure-time or lifetime data, one often has to handle incomplete observations which are caused by some type of censoring. One type of censoring, which is widely accepted in practice, is described by the random censorship model (RCM). Under this model one has two independent sequences of independent and identically distributed (IID) random variables: the survival times $X_1, \ldots, X_n$ and the censoring times $Y_1, \ldots, Y_n$. These sequences define the observations: $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$, where $Z_i = \min(X_i, Y_i)$ and δ_i indicates whether the observation time Z_i is a survival time ($\delta_i = 1$) or a censoring time ($\delta_i = 0$). We assume that all these sequences are defined over some probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Denote the distribution functions (DF) of X, Y, and Z by F, G, and H, respectively, and assume that they are continuous. Nonparametric statistical inference of F based on this type of observations is usually based on the time-honored Kaplan–Meier (KM) or product limit estimator, see Kaplan and Meier [12], defined by

$$F_n^{KM}(t) = 1 - \prod_{i:\, Z_i \le t} \left(1 - \frac{\delta_i}{n - R_{i,n} + 1}\right),$$

where $R_{i,n}$ denotes the rank of Z_i among the Z-sample.

The KM-estimator can be derived as a nonparametric maximum likelihood estimator (MLE), see Johansen [11]. Wellner [24] proved that it also retains some asymptotic optimality properties one usually expects for a MLE to have in a parametric setup. Among other things, he obtained a functional Hájek–Le Cam like convolution result, see Theorem 1 in Wellner [24], to find the optimal centered Gaussian process of any regular estimator $\hat{F}_n$ of F. Since the limiting process of $\sqrt{n}(F_n^{KM}(t) - F(t))$, derived by Breslow and Crowley [4], has the same covariance structure as the optimal centered Gaussian process, the

E-mail address: dikta@fh-aachen.de.

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.10.002

KM-estimator is asymptotically optimal with respect to all regular estimators of F under SRCM. It may be noted, that the processes considered in [24,4] are restricted to the interval $[0, \tau]$ such that $\tau < \tau_H$, where $\tau_H = \inf\{t : H(t) = 1\}$.

The analysis of lifetime data is often focused on the estimation of some characteristic parameters of the underlying distribution which can be represented by an integral of a Borel-measurable function φ with respect to F, that is, by $\int_0^\infty \varphi dF$. The DF F at a fixed point t itself, the expectation, the variance, or the mean residual lifetime are some typical examples, see Section 1 in Stute and Wang [16].

Susarla and Van Ryzin [13] started the analysis of KM-integrals to estimate $\int \varphi dF$, where φ can be of unbounded variation. In particular, they studied the special case of the identity function, $\varphi(x) = x$, and obtained the asymptotic normality of the KM-integral $\int x1\,(0 \le x \le M_n)\, F_n^{KM}(dx)$, where $M_n \to \infty$ and $1\,(x \in A)$ denotes the indicator function of the set A. Among other things, Gill [10] studied the weak convergence of KM-processes on the interval $[0, \tau_H]$ and KM-integrals for nonnegative, continuous, and non-increasing φ. If φ is a monotone function, Schick, Susarla, and Koul [14] generalized Gill's weak convergence result. They provided an asymptotically linear representation of $\sqrt{n}(\int \varphi dF_n^{KM} - \int \varphi dF)$. Furthermore, they deduced from a Hájek–Le Cam like convolution result for regular estimators of $\int \varphi dF$ under RCM the asymptotic efficiency of $\int \varphi dF_n^{KM}$, see Theorem 1.1 and Theorem 1.2 of Schick et al. [14]. For arbitrary Borel-measurable φ, Stute [15] derived, under quite general assumptions, an asymptotic linear representation of $\int \varphi dF_n^{KM}$ which implies, according to the central limit theorem (CLT), the weak convergence of

$$\sqrt{n}\Big(\int \varphi(x)\, F_n^{KM}(dx) - \int_0^{\tau_H} \varphi(x)\, F(dx)\Big)$$

to a centered normal variable with variance $\sigma^2_{KM}(\varphi) \equiv \sigma^2_{KM}$.

The KM-estimator is the first choice under the RCM. Let $Z_{1:n}, \ldots, Z_{n:n}$ denote the ordered observations. F_n^{KM} attaches mass only to the uncensored observations and the amount of attached mass increases from the smallest to the largest uncensored observation. Furthermore, the increase depends on the number of censored observations between two uncensored ones, see Efron [9]. Therefore, if we have to analyze a heavily censored dataset, the KM-estimator will have only a few jumps which might lead to a rather sketchy result. In such a situation we can try to improve our data analysis by tightening the general model assumption of RCM and by using a different type of estimator which is better adapted to the restricted RCM model than the KM-estimator. For example, if it is guaranteed that F belongs to a parametric family, a maximum likelihood approach will be used to find the corresponding parametric estimator of F. Or, if a parametric model is appropriate but the data might be contaminated, the minimum Hellinger distance estimation (MHDE) of the parameter is suitable to get an estimator close to F which is a member of the parametric family. When there is no contamination, the MHDE is known to be asymptotically efficient among the class of regular estimators, see Yang [25,26].

If we have good reasons to assume that the censoring DF G is linked to the DF of the survival times F, the semi-parametric random censorship model (SRCM) is a possible restriction of RCM. Following Dikta [5,6], we assume under SRCM, additionally to RCM, that the conditional expectation of the indicator δ given the observation time $Z = z$,

$$m(z) = \mathbb{E}\big(\delta \,\big|\, Z = z\big) = \mathbb{P}\big(\delta = 1 \big| Z = z\big),$$

belongs to a parametric family

$$m(z) = m(z, \theta_0),$$

where $\theta_0 = (\theta_{0,1}, \ldots, \theta_{0,k}) \in \Theta \subset \mathbb{R}^k$. Together with SRCM, the semi-parametric (SE) estimator

$$F_n^{SE}(t) = 1 - \prod_{i:\, Z_i \le t} \Big(1 - \frac{m(Z_i, \hat{\theta}_n)}{n - R_{i,n} + 1}\Big)$$

of F was proposed in Dikta [5,6], where $\hat{\theta}_n$ is the MLE of θ_0, that is, the maximizer of the (partial) likelihood function

$$L_n(\theta) = \prod_{i=1}^{n} m(Z_i, \theta)^{\delta_i} (1 - m(Z_i, \theta))^{1-\delta_i}.$$

A discussion of possible parametric models for m can be found in these two papers. Since the parametric model of m is based on a parametric binary model, we can test the validity of this assumption by some goodness-of-fit tests. A general bootstrap based approach for such tests is given in Dikta, Kvesic, and Schmidt [8]. Note that F_n^{SE} attaches mass to every observation time Z_i and not only to the uncensored ones, like F_n^{KM} does.

The semi-parametric approach together with the corresponding estimator F_n^{SE} has been shown to be applicable and flexible enough for extensions to other statistical scenarios. For example, Subramanian [17] introduced an extension to the missing censoring-indicator model. He proved that the extended semi-parametric estimator is at least as good as an efficient nonparametric one under the missing censoring-indicator model if the parametric model assumption is correct. Sun and Zhu [20] applied this approach to truncated and censored data and Subramanian [18] for multiple imputations in the missing censoring-indicator model. Recently, Subramanian and Zhang [19] introduced a two-stage bootstrap procedure to construct confidence bands for the survival function under the semi-parametric model with and without missing censoring-indicators.

The asymptotic properties of F_n^{SE} were analyzed in Dikta [5,6]. Among other things, if the assumed model is correct, it is shown there that the asymptotic variance of $F_n^{SE}(t)$ is less than or equal to the asymptotic variance of $F_n^{KM}(t)$, where equality can only occur in exceptional cases, see Corollary 2.7 in Dikta [5]. Corresponding to the CLT obtained by Stute [15] for arbitrary Borel-measurable φ, under quite general assumptions, an asymptotic linear representation of $\int \varphi dF_n^{SE}$ was derived in Dikta, Ghorai, and Schmidt [7] which implies the weak convergence of

$$\sqrt{n}\Big(\int \varphi(x)\, F_n^{SE}(dx) - \int_0^{\tau_H} \varphi(x)\, F(dx)\Big) \tag{1}$$

to a centered normal variable with variance $\sigma_{SE}^2(\varphi) \equiv \sigma_{SE}^2$. Furthermore, if the assumed model is correct, $\sigma_{SE}^2 \leq \sigma_{KM}^2$, where equality can occur only in rare and unrealistic situations, compare Corollary 2.5 in Dikta et al. [7].

Under SRCM, if the model assumptions are correct, $\int \varphi dF_n^{SE}$ outperforms $\int \varphi dF_n^{KM}$, since it incorporates the additional parametric features. However, is $\int \varphi dF_n^{SE}$ asymptotically efficient under SRCM? The objective here is to answer this question affirmatively. We will show that the information bound based on all regular estimators of $\int \varphi(x) 1(x \leq \tau_H)\, F(dx)$ under SRCM coincides with the asymptotic variance of $\int \varphi(x)\, F_n^{SE}(dx)$.

This article is organized as follows. In Section 2 we define the semi-parametric model and outline the general approach to prove asymptotic efficiency for regular estimators under SRCM. Basic definitions and notations can be found there. In the third section, a consecutive list of lemmas verifies the technical details of the general approach under SRCM and leads to our main result in Theorem 3.9, which gives the efficient information bound for regular and asymptotically linear (RAL) estimators. A first corollary to this theorem demonstrates the asymptotic efficiency of the semi-parametric integral estimator with respect to the class of all RAL estimators. In a second corollary, by an application of the convolution theorem, this result is extended to the class of all regular estimators. All the proofs are given in the fourth section.

2. Semi-parametric setup

In this section we reflect some general ideas in proving efficiency and adapt them to our concrete SRCM setup.

Throughout this article, we denote the partial derivative $\partial m(z,\theta)/\partial\theta_r|_{\theta=\theta^*}$ by $D_r m(z,\theta^*)$ and we write for the gradient: $Dm(z,\theta^*) \equiv \mathrm{Grad}(m(z,\theta^*)) = (D_1 m(z,\theta^*), \ldots, D_k m(z,\theta^*))^\top$, where the parameter set $\Theta \subset \mathbb{R}^k$ is assumed to be an open sphere around the true parameter θ_0. Open and closed intervals are denoted by $(\cdot,\cdot)$ and $[\cdot,\cdot]$, respectively.

For an arbitrary measure λ, $L^2(\lambda)$ denotes the Hilbert-space of all square-integrable functions with respect to λ and $L_0^2(\lambda)$ the sub-space of λ-centered functions $l \in L^2(\lambda)$, that is, $\int l d\lambda = 0$.

To characterize the distribution of our observations (Z,δ) in $\mathbb{R}^+ \times \{0,1\}$ with respect to SRCM, take μ to be a σ-finite Borel-measure on $\mathbb{R}^+$ such that H has a density with respect to μ, and let ν denote the counting measure on $\{0,1\}$. Set for an arbitrary μ-density l and a $\theta \in \Theta$

$$\mathbb{R}^+ \times \{0,1\} \ni (z,j) \longrightarrow p_{\theta,l}(z,j) = m(z,\theta)^j\,(1-m(z,\theta))^{1-j}\, l(z) \in \mathbb{R}^+. \tag{2}$$

The (global) semi-parametric model can now be defined by the set of densities with respect to the product measure $\mu \otimes \nu$:

$$\mathcal{P} = \Big\{p_{\theta,l} \,:\, \theta \in \Theta,\ l \geq 0,\ \int l d\mu = 1,\ \text{with each } p_{\theta,l} \text{ satisfying (2)}\Big\}.$$

Let h and θ_0 denote the true μ-density of H and the true parameter of m, respectively. Then $p_{\theta_0,h} \equiv p$ is the true density of the distribution of (Z,δ) with respect to the product measure $\mu \otimes \nu$ on $\mathbb{R}^+ \times \{0,1\}$. We denote this density by p and we take h and θ_0 fixed throughout the rest of this article.

We now allow p to vary locally in $\mathcal{P}$ towards certain directions. Each direction defines a parametric sub-model. To be precise, take $\theta \in \mathbb{R}^k$ and a bounded and H-centered q, that is, $\int q\, dH = 0$. Then choose an $\varepsilon(q,\theta) \equiv \varepsilon > 0$ such that for all $-\varepsilon < t < \varepsilon$ and $z \in \mathbb{R}^+$:

$$1 + \varepsilon_0 > 1 + tq(z) > 1 - \varepsilon_0 > 0, \tag{3}$$

where $\varepsilon_0 > 0$ can be taken conveniently. Among other things, this implies:

$$(-\varepsilon,\varepsilon) \ni t \longrightarrow p_{\theta_0+t\theta,(1+tq)h} \equiv p_{\theta,q,t} \equiv p_t \in \mathcal{P}.$$

Let $\mathcal{P}_{\theta,q}$ denote the set of all densities in this sub-model and note that this sub-model is a one-dimensional parametric family in t with open parameter space $(-\varepsilon,\varepsilon)$ such that $p_{\theta,q,0} = p$. To simplify the notation, we will denote the densities defining the concrete sub-model $\mathcal{P}_{\theta,q}$ by p_t and we will use h_t instead of $h_{q,t} \equiv (1+tq)h$. Correspondingly, $H_t \equiv H_{q,t}$, where $H_t(z) = \int_0^z h_t(s)\,\mu(ds)$, denotes the DF with μ-density h_t. Furthermore, we will use $m_t(z)$ and $Dm_t(z)$ instead of $m(z,\theta_0+t\theta)$ and $Dm(z,\theta_0+t\theta)$, respectively.

Fix a concrete sub-model $\mathcal{P}_{\theta,q} = (p_t)_{-\varepsilon<t<\varepsilon}$ and $-\varepsilon < t_0 < \varepsilon$. Assume that the score function S_{t_0} of such a sub-model exists at $t = t_0$. Then

$$\begin{aligned} S_{t_0}(z,j) &= \frac{d}{dt}\log(p_t(z,j))\Big|_{t=t_0} \\ &= \frac{j - m_{t_0}(z)}{m_{t_0}(z)\,(1-m_{t_0}(z))}\,\theta^\top Dm_{t_0}(z) + \frac{q(z)\,h(z)}{h_{t_0}(z)}, \end{aligned} \tag{4}$$

and we get for $t_0 = 0$:

$$S_0(z,j) \equiv S_{\theta,q}(z,j) \equiv \left.\frac{d}{dt}\log(p_t(z,j))\right|_{t=0} = K(z,j)\theta^\top Dm(z,\theta_0) + q(z), \tag{5}$$

where

$$K(z,j) = \frac{j - m(z,\theta_0)}{m(z,\theta_0)\,(1 - m(z,\theta_0))}.$$

Take for each sub-model its score function at $t = 0$ to build the tangent space $\mathcal{T}$, that is, the closed linear space spanned by the set of all score functions at $t = 0$

$$\left\{S_{\theta,q} \;:\; \theta \in \mathbb{R}^k,\ \int q\,dH = 0,\ q \text{ is bounded}\right\}$$

in $L^2(p\,d\mu \otimes \nu)$. Here $p\,d\mu \otimes \nu$ denotes the probability measure on $\mathbb{R}^+ \times \{0, 1\}$ defined by the $\mu \otimes \nu$-density p.

To ensure some contiguity argument, our sub-models have to be Hellinger-differentiable at $t = 0$, where a sub-model $\mathcal{P}_{\theta,q} = (p_t)_{-\varepsilon<t<\varepsilon}$ is called Hellinger-differentiable at $t \in (-\varepsilon, \varepsilon)$ if

$$\int \left(\left(\sqrt{p_{t+s}} - \sqrt{p_t}\right)/s - S_t\sqrt{p_t}/2\right)^2 d\mu \otimes \nu \longrightarrow 0, \quad \text{as } s \to 0, \tag{6}$$

compare van der Vaart [22, p. 65, (5.38)].

Now let $\varphi \in L^1(F)$. Our goal is to estimate

$$\beta_0 = \int_0^{\tau_H} \varphi(x)\,F(dx),$$

which equals $\mathbb{E}(\varphi(X))$ if $\tau_H = \tau_F$. We can express β_0 under SRCM by

$$\beta_0 = \int \varphi(z)\gamma_0(z,p)\,p(z,1)\mu(dz), \tag{7}$$

where, for an arbitrary sub-model density $p_{\theta,q,t}$,

$$\gamma_0(z, p_{\theta,q,t}) = \exp\left(\int_0^z \frac{1 - m(u,\theta_0 + t\theta)}{1 - H_{q,t}(u)}\,h_{q,t}(u)\mu(du)\right). \tag{8}$$

Note that

$$\gamma_0(z) \equiv \gamma_0(z,p) \equiv \gamma_0(z, p_{\theta,q,0}) = 1/(1 - G(z)),$$

compare Remark 2.2 in Dikta et al. [7]. Based on (7) and (8), we define for each sub-model $\mathcal{P}_{\theta,q}$

$$\Psi_{\theta,q} :]\varepsilon, \varepsilon[\ni t \longrightarrow \Psi_{\theta,q}(t) := \int \varphi(z)\gamma_0(z,p_t)\,p_t(z,1)\mu(dz) \in \mathbb{R}. \tag{9}$$

Following van der Vaart [23, p. 363], we will show that Ψ, that is the map $(\theta, q) \longrightarrow \Psi_{\theta,q}$, is differentiable at $t = 0$ for every sub-model $\mathcal{P}_{\theta,q}$ relative to the tangent space $\mathcal{T}$ (Ψ is pathwise differentiable on $\mathcal{T}$ at $t = 0$). That is,

$$t^{-1}\left(\Psi_{\theta,q}(t) - \Psi_{\theta,q}(0)\right) \longrightarrow \Psi'(S_{\theta,q}), \quad \text{as } t \to 0,$$

where Ψ' is a continuous linear functional on the tangent space $\mathcal{T}$ which is the same for every sub-model, i.e. it does not depend on θ or on q. Note that

$$\Psi_{\theta,q}(0) = \beta_0$$

for every sub-model.

Like in Tsiatis [21], we will consider here asymptotically linear (AL) estimators $\hat\beta_n(Z_1, \delta_1, \ldots, Z_n, \delta_n)$ of β_0, that is,

$$n^{1/2}\left(\hat\beta_n - \beta_0\right) = n^{-1/2}\sum_{i=1}^n A(Z_i, \delta_i) + o_{\mathbb{P}}(1), \tag{10}$$

with influence function (IF) $A \in L_0^2(p\,d\mu \otimes \nu)$. Due to the CLT, an AL estimator converges in distribution to a centered normal distribution with asymptotic variance $\int A^2 p d\mu \otimes \nu$.

To avoid super-efficiency, we assume that a possible estimator (sequence) $\hat\beta_n$ of β_0 has to be (locally) regular at p, that is, if $(Z_{1,n}, \delta_{1,n}), \ldots, (Z_{n,n}, \delta_{n,n})$ are IID with density p, then

$$\sqrt{n}\left(\hat\beta_n\left(Z_{1,n}, \delta_{1,n}, \ldots, Z_{n,n}, \delta_{1,n}\right) - \beta_0\right)$$

converges in distribution to a random variable L, and, for every sub-model $\mathcal{P}_{\theta,q}$,

$$\sqrt{n}\Big(\hat{\beta}_n\big(Z_{1,n}, \delta_{1,n}, \ldots, Z_{n,n}, \delta_{1,n}\big) - \Psi_{\theta,q}(1/\sqrt{n})\Big),$$

where the observation are now IID with density $p_{\theta,q,1/\sqrt{n}}$, converges to the same random variable L in distribution. Here

$$\Psi_{\theta,q}(1/\sqrt{n}) = \int \varphi(z)\gamma_0(z, p_{\theta,q,1/\sqrt{n}})\, p_{\theta,q,1/\sqrt{n}}(z, 1)\mu(dz), \tag{11}$$

see van der Vaart [22, p. 365].

To give some guidelines for the technical parts in the next section here, we briefly sketch the approach to prove asymptotic efficiency for regular and asymptotically linear estimators (RAL), compare Bickel, Klaassen, Ritov, and Wellner [2], van der Vaart [23,22], Tsiatis [21]. Assume that $A \in L_0^2(p\,d\mu \otimes \nu)$ is the IF of a RAL estimator, and choose a sub-model $\mathcal{P}_{\theta,q}$. By Cauchy–Schwarz inequality we get

$$\mathrm{VAR}\big(A(Z,\delta)\big)\,\mathbb{E}\big(S_{\theta,q}^2(Z,\delta)\big) \geq \mathbb{E}^2\big(A(Z,\delta)\,S_{\theta,q}(Z,\delta)\big).$$

Based on a contiguity argument, which will be derived by the Hellinger differentiability of our sub-models, and the assumed regularity, we will conclude that

$$\mathbb{E}^2\big(A(Z,\delta)\,S_{\theta,q}(Z,\delta)\big) = \left(\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=0}\right)^2.$$

Since Ψ is assumed to be differentiable for every sub-model relative to the tangent space $\mathcal{T}$, there exists a continuous linear functional Ψ' on the tangent space such that

$$\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=0} = \Psi'(S_{\theta,q}).$$

Since the tangent space $\mathcal{T}$ is a closed linear sub-space of $L_0^2(p\,d\mu \otimes \nu)$ and therefore a Hilbert-space, Riesz representation theorem for Hilbert-spaces, see Theorem 2.E in Zeidler [27], can be applied to deduce that

$$\Psi'(S_{\theta,q}) = \mathbb{E}\big(\hat{\Psi}(Z,\delta)\,S_{\theta,q}(Z,\delta)\big),$$

where $\hat{\Psi} \in \mathcal{T}$ denotes the Riesz-density. In summary, this gives us the information bound (IB)

$$\mathrm{VAR}\big(A(Z,\delta)\big) \geq \sup_{S\in\mathcal{T}} \mathbb{E}^2\big(\hat{\Psi}(Z,\delta)S(Z,\delta)\big)/\mathbb{E}\big(S^2(Z,\delta)\big) \tag{12}$$

for RAL estimators. Furthermore, due to Riesz representation theorem we will conclude that the IB is equal to $\mathbb{E}\big(\hat{\Psi}^2(Z,\delta)\big)$.

As we will see, the semi-parametric integral estimator is asymptotically linear with IF $\hat{\Psi}$. Thus it achieves the IB given in (12) and is therefore asymptotically efficient with respect to the class of all RAL estimators. This result will be derived without an application of the convolution theorem (CT), see Theorem 25.20 in van der Vaart [22].

However, as a direct consequence of the CT, we will extend this result to obtain that the semi-parametric integral estimator is a regular asymptotically efficient estimator with respect to the class of all regular estimators under SRCM.

3. Main results

Throughout this section we assume that $\Theta \subset \mathbb{R}^k$ is an open sphere around θ_0. We will use the following regularity assumptions for our model m:

(A_1) For every $\theta \in \Theta$ and $z > 0$, $m(z,\theta)$ possesses continuous partial derivatives with respect to θ. Furthermore,

$$\sup_{\theta\in\Theta}\sup_{z>0} \|Dm(z,\theta)\| < \infty.$$

(A_2) For every $z > 0$, $m(z,\theta)$ possesses continuous partial derivatives with respect to θ, for every θ in the open sphere $\Theta \subset \mathbb{R}^k$ around θ_0, such that $Dm(z,\theta)/m(z,\theta)$ and $Dm(z,\theta)/(1-m(z,\theta))$ are well defined and continuous in θ. Furthermore, for every $1 \leq r, s \leq k$ and every $\theta \in \Theta$,

$$\left|\frac{D_r m(z,\theta)\,D_s m(z,\theta)}{m(z,\theta)(1-m(z,\theta))}\right| \leq M(z),$$

where M is H-integrable.

(A_3) The matrix $I_0 = (\sigma_{r,s})_{1\leq r,s,\leq k}$ is positive definite, where

$$\sigma_{r,s} = \mathbb{E}\left(\frac{D_r(m(Z,\theta_0))D_s(m(Z,\theta_0))}{m(Z,\theta_0)(1-m(Z,\theta_0))}\right). \tag{13}$$

Finally, some moment conditions on φ are necessary, which are listed below.

(M_1) $\int_0^{\tau_H} |\varphi(z)| \big(1 - H(z)\big)^{-(1+\kappa)} H(dz) < \infty$, for some $\kappa > 0$.

(M_2) $\int_0^{\tau_H} |\varphi(z)| \big(1 - H(z)\big)^{-(1/2+\kappa)} F(dz) < \infty$, for some $\kappa > 0$.

(M_3) $\int_0^{\tau_H} \varphi^2(z)\gamma_0(z) F(dz) < \infty$.

These conditions are similar to those discussed in Dikta et al. [7]. The moment conditions (M_1) and (M_2), however, are a bit more restrictive than the corresponding ones there. Technically, it allows us to swap integration and differentiation in some crucial parts of the proofs.

To demonstrate the applicability of the regularity conditions (A_1)–(A_3), we consider a generalization of the simple proportional hazards model (GPHM) in the following example. GPHM was introduced in Example 2.9 of Dikta [5]. Furthermore, Table 6 in Dikta et al. [8] shows its relevance in real data applications.

Example 3.1. Let $0 < a < \theta_{0,1} < b < \infty$ and $0 < c < \theta_{0,2} < d < \infty$. Set $\theta_0 = (\theta_{0,1}, \theta_{0,2})$ and define

$$m(z, \theta_0) = \theta_{0,1}/\big(\theta_{0,1} + z^{\theta_{0,2}}\big), \quad \text{for } z > 0.$$

Let Θ be some open sphere contained in $(a, b) \times (c, d)$ with $\theta_0 \in \Theta$. It is easy to check that

$$\sup_{\theta\in\Theta}\sup_{z>0} \|Dm(z, \theta)\| < \infty$$

and

$$\sup_{\theta\in\Theta}\sup_{z>0} \left| \frac{D_r m(z, \theta)\, D_s m(z, \theta)}{m(z, \theta)(1 - m(z, \theta))} \right| < \infty, \quad \text{for } 1 \le r, s, \le 2.$$

Thus (A_1) and (A_2) are fulfilled. Now set

$$\xi_r = \frac{D_r m(Z, \theta_0)}{\sqrt{m(Z, \theta_0)(1 - m(Z, \theta_0))}}, \quad \text{for } 1 \le r \le 2,$$

to get that $I_0 = \big(\mathbb{E}(\xi_r\xi_s)\big)_{1\le r,s\le 2}$. An application of Cauchy–Schwarz inequality shows that the determinant of I_0 is 0 if and only if at least one of the exceptional cases occur with probability one: (i) $\xi_1 = 0$, (ii) $\xi_2 = 0$, or (iii) $\xi_1 = c\,\xi_2$, for some constant c, see Appendix A.11.17 in Bickel and Doksum [1]. Elsewhere, the determinant of I_0 will be positive. Furthermore, by a standard algebraic argument, I_0 will then be positive definite. Since the exceptional cases can only occur if the distribution of Z is rather unrealistic, (A_3) will be fulfilled in real applications.

In the first two lemmas, we define the sub-models and analyze some continuity and differentiability aspects of Ψ with respect to these sub-models.

Lemma 3.2. *Let Θ be the open sphere around θ_0, $\theta \in \mathbb{R}^k$, and $q \in L_0(H)$ a bounded function. If H is continuous and $0 < \varepsilon_0 \le 1/2$, then an $\varepsilon(q, \theta, \varepsilon_0) \equiv \varepsilon > 0$ exists such that*

(i) *$\mathscr{P}_{\theta,q} = (p_t)_{-\varepsilon<t<\varepsilon}$ is well defined;*

(ii) *for all $z > 0$ and $-\varepsilon < t < \varepsilon$*

$$(1 - \varepsilon_0)h(z) \le h_t(z) \le (1 + \varepsilon_0)h(z).$$

(iii) *If in addition (M_1) is satisfied, ε_0 can be chosen such that the map*

$$\Psi_{\theta,q} : (\varepsilon, \varepsilon) \ni t \longrightarrow \Psi_{\theta,q}(t) := \int \varphi(z)\gamma_0(z, p_t)\, p_t(z, 1)\mu(dz) \in \mathbb{R}$$

is well defined and continuous.

Now define the subdistribution functions

$$H^1(z) = \mathbb{P}(\delta = 1, Z \le z) = \int_0^z m(t, \theta_0)\, H(dt)$$

and

$$H^0(z) = \mathbb{P}(\delta = 0, Z \le z) = \int_0^z \big(1 - m(t, \theta_0)\big)\, H(dt)$$

and note that these subdistributions have the Radon–Nikodym derivatives, respectively given by

$$H^1(dz) = m(z, \theta_0)\, H(dz) = (1 - G(z))\, F(dz),$$

and

$$H^0(dz) = (1 - m(z, \theta_0))\, H(dz).$$

Lemma 3.3. *Let Θ be the open sphere around θ_0, $\theta \in \mathbb{R}^k$, and $q \in L_0(H)$ a bounded function. If H is continuous,* (A$_1$) *and* (M$_1$) *are satisfied, then ε_0 and $\varepsilon(q, \theta, \varepsilon_0) \equiv \varepsilon$ can be chosen such that $\Psi_{\theta,q}$ is differentiable at every $-\varepsilon < t_0 < \varepsilon$ and*

$$\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=t_0} = \int \varphi(z)\frac{d}{dt}\Big(\gamma_0(z, p_t)\, m_t(z)\, h_t(z)\Big)\bigg|_{t=t_0} \mu(dz). \tag{14}$$

Now set

$$\gamma_1(u) = \frac{1}{1 - H(u)} \int_u^\infty \varphi(z)\gamma_0(z)\, H^1(dz), \tag{15}$$

and

$$\gamma_2(v) = \int_0^v \int_u^\infty \frac{\varphi(z)\gamma_0(z)}{(1 - H(u))^2}\, H^1(dz)\, H^0(du), \tag{16}$$

to obtain the following remark:

Remark 3.4. If the assumptions of Lemma 3.3 are satisfied, then

$$\begin{aligned}\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=0} &= -\int \gamma_1(u)\theta^\top Dm(u, \theta_0)\, H(du) - \int \gamma_2(v)q(v)\, H(dv) + \int \gamma_1(u)q(u)\, H^0(du)\\ &\quad + \int \varphi(z)\gamma_0(z)\theta^\top Dm(z, \theta_0)\, H(dz) + \int \varphi(z)\gamma_0(z)q(z)\, H^1(dz).\end{aligned} \tag{17}$$

From now onwards, we assume that ε and ε_0 are always chosen based on θ and q according to Lemma 3.3, if the assumptions of this lemma are satisfied. Otherwise, if the assumptions (A$_1$) and (M$_1$) are not explicitly stated, we will take ε and ε_0 according to Lemma 3.2.

Before we can verify that Ψ is differentiable relative to the tangent space $\mathcal{T}$, we have to study Hellinger-differentiability for our sub-models. This is done in the next lemma.

Lemma 3.5. *Let Θ be the open sphere around θ_0 and $\mathcal{P}_{\theta,q} = (p_t)_{-\varepsilon<t<\varepsilon}$ a sub-model. If* (A$_2$) *is satisfied, then*

(i) *the map $t \longrightarrow \sqrt{p_t(z, j)}$ is differentiable at every point $-\varepsilon < t_0 < \varepsilon$ and for every (z, j);*
(ii) *the map $t \longrightarrow I(t) := \int S_t^2 p_t\, d\mu \otimes \nu$ is well defined and continuous at every point $-\varepsilon < t_0 < \varepsilon$, where S_t is the score function at t, see (4);*
(iii) *the sub-model $\mathcal{P}_{\theta,q}$ is Hellinger-differentiable at every point $-\varepsilon < t_0 < \varepsilon$, see (6).*

To describe the consequences of this lemma, fix a sub-model $\mathcal{P}_{\theta,q} = (p_t)_{-\varepsilon<t<\varepsilon}$ and define, for $-\varepsilon < t < \varepsilon$, the n-dimensional product measure

$$P_t^{(n)} = \Big(\prod_{i=1}^n p_t\Big)\big(\mu \otimes \nu\big)^n;$$

that is, the joint distribution of the IID observations $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$, where (Z_i, δ_i) is distributed according to p_t as $\mu \otimes \nu$-density. For $t = 0$ we will use $P^{(n)}$ instead of $P_0^{(n)}$.

Remark 3.6. If the assumptions of Lemma 3.5 are satisfied, then, for every $-\varepsilon < t < \varepsilon$,

$$\int S_t p_t\, d\mu \otimes \nu = 0.$$

Furthermore, $(P^{(n)}_{1/\sqrt{n}})_{n\in\mathbb{N}}$ and $(P^{(n)})_{n\in\mathbb{N}}$ are mutually contiguous; that is, for every sequence $(A_n)_{n\in\mathbb{N}}$, where A_n is measurable with respect to the corresponding σ-Algebra of the n-dimensional product space

$$P^{(n)}_{1/\sqrt{n}}(A_n) \longrightarrow 0, \quad \text{as } n \to \infty, \text{ if and only if } P^{(n)}(A_n) \longrightarrow 0, \quad \text{as } n \to \infty.$$

Due to Lemma 3.5, we can apply Theorem 7.2 of van der Vaart [22] to conclude that $S_t \in L_0^2(p_t\, d\mu \otimes \nu)$. Furthermore, this theorem proves an asymptotic representation of the log-likelihood ratio which, together with Example 6.5 of van der Vaart [22] and Le Cam's first lemma, compare Lemma 6.4 in van der Vaart [22], yields the asserted mutual contiguity.

To show that Ψ is differentiable relative to the tangent space, we have to find a continuous linear functional Ψ' on $\mathcal{T}$ such that, for all $\theta \in \mathbb{R}^k$, and bounded q with $\int q\, dH = 0$,

$$\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=0} = \Psi'(S_{\theta,q}),$$

where Ψ' depends neither on q nor on θ. This will be done in two steps.

Lemma 3.7. *Let Θ be the open sphere around θ_0, $\theta \in \mathbb{R}^k$, and $q \in L_0(H)$ a bounded function. If H is continuous,* (A_1), (A_2), *and* (M_1)–(M_3) *are satisfied, then*

(i) $\Psi^* \in L^2(p\, d\mu \otimes \nu)$, *where*

$$\Psi^*(z,j) = \varphi(z)\gamma_0(z)1(j=1) + \gamma_1(z)1(j=0) - \gamma_2(z); \tag{18}$$

(ii) *for the derivative of $\Psi_{\theta,q}(t)$ at $t=0$ we get*

$$\frac{d}{dt}\Psi_{\theta,q}(t)\bigg|_{t=0} = \int \Psi^*(z,j)S_{\theta,q}(z,j)\, p(z,j)\, \mu(dz)\nu(dj). \tag{19}$$

The right hand side of (19) defines a continuous linear functional on the Hilbert-space $\mathcal{T}$, i.e., $\mathcal{T} \ni T \to \int \Psi^* T p\, d\mu \otimes \nu$. Riesz representation theorem, see Theorem 2.E in Zeidler [27], now guarantees the existence of a Riesz-density $\hat{\Psi} \in \mathcal{T}$ such that

$$\int \Psi^*(z,j)S_{\theta,q}(z,j)\, p(z,j)\, \mu(dz)\nu(dj) = \int \hat{\Psi}(z,j)S_{\theta,q}(z,j)\, p(z,j)\, \mu(dz)\nu(dj)$$

holds for every bounded, H-centered q and every $\theta \in \mathbb{R}^k$. Obviously, $\hat{\Psi}$ is the L^2-projection of Ψ^* onto $\mathcal{T}$.

Lemma 3.8. *Let Θ be the open sphere around θ_0. If H is continuous,* (A_1)–(A_3), *and* (M_1)–(M_3) *are satisfied, then*

(i) $\hat{\Psi} \in \mathcal{T}$, *where*

$$\hat{\Psi}(z,j) = K(z,j)\hat{\theta}^\top Dm(z,\theta_0) + \hat{q}(z), \tag{20}$$

with

$$\hat{\theta} = \int \Big(\varphi(z)\gamma_0(z) - \gamma_1(z)\Big) I_0^{-1} Dm(z,\theta_0)\, H(dz)$$

and

$$\hat{q}(z) = \varphi(z)\gamma_0(z)m(z,\theta_0) + \gamma_1(z)(1-m(z,\theta_0)) - \gamma_2(z) - \beta_0;$$

(ii) $\hat{\Psi}$ *is the L^2-projection of Ψ^* onto $\mathcal{T}$;*

(iii) *for the continuous linear functional defined through its Riesz density $\hat{\Psi}$ we get*

$$\frac{d}{dt}\Psi_{\theta,q}\bigg|_{t=0} = \int \hat{\Psi}(z,j)S_{\theta,q}(z,j)\, p(z,j)\, \mu(dz)\nu(dj), \tag{21}$$

for every $\theta \in \mathbb{R}^k$ and every bounded $q \in L_0(H)$.

Now we have all the technical details to specify the asymptotic information bound for RAL estimators under SRCM.

Theorem 3.9. *Let Θ be the open sphere around θ_0. If H is continuous,* (A_1)–(A_3), *and* (M_1)–(M_3) *are satisfied, then the asymptotic variance of any regular, asymptotically linear, estimator of $\beta_0 = \int \varphi(z)1(0 \le z < \tau_H)\, F(dz)$ under SRCM is bounded from below by the variance of $\hat{\Psi}$, that is,*

$$\int \Big(\frac{\hat{\theta}^\top Dm(z,\theta_0)}{m(z,\theta_0)} + \hat{q}\Big)^2 H^1(dz) + \int \Big(-\frac{\hat{\theta}^\top Dm(z,\theta_0)}{1-m(z,\theta_0)} + \hat{q}\Big)^2 H^0(dz), \tag{22}$$

where

$$\hat{\theta} = \int \Big(\varphi(z)\gamma_0(z) - \gamma_1(z)\Big) I_0^{-1} Dm(z,\theta_0)\, H(dz)$$

and

$$\hat{q}(z) = \varphi(z)\gamma_0(z)m(z,\theta_0) + \gamma_1(z)(1-m(z,\theta_0)) - \gamma_2(z) - \beta_0.$$

A straightforward comparison of the summands in the asymptotically linear representation of $\int \varphi dF_n^{SE}$, compare Eq. 2.3 and Theorem 2.1 in Dikta et al. [7], shows that its IF is exactly the efficient IF.

Corollary 3.10. *If the assumptions of Theorem* 2.1 *in Dikta et al.* [7] *are satisfied, then*

$$\sqrt{n}\left(\int \varphi dF_n^{SE} - \int_0^{\tau_H} \varphi dF\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\Psi}(Z_i,\delta_i) + o_{\mathbb{P}}(1). \tag{23}$$

Furthermore,

$$\sqrt{n}\left(\int \varphi dF_n^{SE} - \int_0^{\tau_H} \varphi dF\right) \longrightarrow \mathcal{N}(0, \sigma_{SE}^2), \quad \textit{in distribution},$$

where $\sigma_{SE}^2 = \mathbb{E}(\hat{\Psi}^2(Z, \delta))$ *is equal to the IB given by Eq. (22). Thus the semi-parametric integral estimator is asymptotically efficient with respect to all RAL estimators of* $\int_0^{\tau_H} \varphi dF$ *under SRCM.*

Moreover, an application of Lemma 25.23 of van der Vaart [22] shows that $\int \varphi F_n^{SE}$ is regular. Finally, the CT says that $\mathbb{E}(\hat{\Psi}^2(Z, \delta))$ is a lower bound for the variance of any regular estimator. Thus, we can extend our result to the following corollary:

Corollary 3.11. *If the assumptions of the Theorem 3.9 and of Theorem 2.1 in Dikta et al. [7] are satisfied, then the semi-parametric integral estimator is a regular estimator. Furthermore, it is asymptotically efficient with respect to the class of all regular estimators of* $\int_0^{\tau_H} \varphi dF$ *under SRCM.*

4. Proofs

Proof of Lemma 3.2. Since Θ is an open sphere around θ_0, an $\varepsilon_1(\theta) \equiv \varepsilon_1 > 0$ can be found such that $\theta_0 + t\theta \in \Theta$, for all $-\varepsilon_1 < t < \varepsilon_1$. Furthermore, $q \in L_0(H)$ is bounded. Therefore, we can find for an arbitrary $1/2 \geq \varepsilon_0 > 0$ an $\varepsilon_2(q, \varepsilon_0) \equiv \varepsilon_2 > 0$ such that $|tq(z)| < \varepsilon_0$, for all $-\varepsilon_2 < t < \varepsilon_2$ and $z > 0$. Set $\varepsilon \equiv \varepsilon(q, \theta, \varepsilon_0) = \min(\varepsilon_1, \varepsilon_2)$, and take $-\varepsilon < t < \varepsilon$. Then $p_t \geq 0$, $\int p_t d\mu \otimes \nu = 1$, and $(1 - \varepsilon_0)h(z) \leq h_t(z) \leq (1 + \varepsilon_0)h(z)$. This proves part (i) and (ii) of the lemma.
To prove (iii), check that the integrand of $\Psi_{\theta,q}(t)$

$$|\varphi(z)|\gamma_0(z, p_t)\, p_t(z, 1) \leq |\varphi(z)| \exp\Big(\int_0^z \frac{h_t(u)}{1 - H_t(u)}\, \mu(du)\Big) h_t(z).$$

Due to (ii), we can bound the right-hand side by

$$|\varphi(z)| \exp\Big(\int_0^z \frac{(1 + \varepsilon_0)h(u)}{(1 - \varepsilon_0)(1 - H(u))}\, \mu(du)\Big)(1 + \varepsilon_0)h(z),$$

which is equal to

$$|\varphi(z)|\big(1 - H(z)\big)^{-(1+\varepsilon_0)/(1-\varepsilon_0)}(1 + \varepsilon_0)h(z).$$

Take $\varepsilon_0 = \min(1/2, \kappa/(2 + \kappa))$ and the corresponding $\varepsilon(q, \theta, \varepsilon_0) \equiv \varepsilon > 0$. Check that the last term is uniformly bounded in $-\varepsilon < t < \varepsilon$, for every $z > 0$, by

$$|\varphi(z)|\big(1 - H(z)\big)^{-(1+\kappa)}(1 + \varepsilon_0)h(z).$$

Due to (M_1), this bound is μ-integrable, and the assertion (iii) follows from Lebesgue's dominated convergence theorem (DOM), which completes the proof of part (iii). □

Proof of Lemma 3.3. According to Lemma 3.2, we can start with $\hat{\varepsilon}(q, \theta, \hat{\varepsilon}_0) \equiv \hat{\varepsilon}$ and $\hat{\varepsilon}_0$ to get that $\Psi_{\theta,q}$ is well defined on $-\hat{\varepsilon} < t < \hat{\varepsilon}$ and is continuous. Note that the integrand of $\Psi_{\theta,q}$ equals $\varphi(z)\gamma_0(z, p_t)\, m_t(z)h_t(z)$.
Set $A(z, t) := \gamma_0(z, p_t)\, m_t(z)h_t(z)$, and take its derivative with respect to $t_0 \in (-\hat{\varepsilon}, \hat{\varepsilon})$ to obtain

$$\begin{aligned} \frac{d}{dt}A(z, t)\Big|_{t=t_0} &= \frac{d}{dt}\gamma_0(z, p_t)\Big|_{t=t_0} m_{t_0}(z)h_{t_0}(z) + \gamma_0(z, p_{t_0})\Big(\theta^\top D\big(m_{t_0}(z)\big)\Big)h_{t_0}(z) + \gamma_0(z, p_{t_0})m_{t_0}(z)q(z)h(z) \\ &\equiv I_1(z, t_0) + I_2(z, t_0) + I_3(z, t_0). \end{aligned}$$

An application of the chain rule yields

$$I_1(z, t_0) = \gamma_0(z, p_{t_0})\ \frac{d}{dt}\Big(\int_0^z \frac{1 - m_t(u)}{1 - H_t(u)} h_t(u)\, \mu(du)\Big)\Big|_{t=t_0} m_{t_0}(z)h_{t_0}(z).$$

Set $B(u, t) := h_t(u)(1 - m_t(u))/(1 - H_t(u))$, and note that

$$\frac{d}{dt}H_t(u)\Big|_{t=t_0} = \int_0^u q\, dH,$$

to obtain

$$\begin{aligned} \frac{d}{dt}B(u, t)\Big|_{t=t_0} &= -\frac{\theta^\top Dm_{t_0}(u)}{1 - H_{t_0}(u)}\, h_{t_0}(u) + \frac{1 - m_{t_0}(u)}{(1 - H_{t_0}(u))^2}\, h_{t_0}(u) \int_0^u q\, dH + \frac{1 - m_{t_0}(u)}{1 - H_{t_0}(u)} q(u)h(u) \\ &\equiv C_1(u, t_0) + C_2(u, t_0) + C_3(u, t_0). \end{aligned}$$

According to (A_1), we can bound $\left|\theta^\top Dm_t(u)\right|$ uniformly in $-\hat{\varepsilon} < t < \hat{\varepsilon}$ and $z > 0$ by a constant c_1. Therefore,

$$\left|C_1(u,t)\right| \le c_1 \, \frac{1+\hat{\varepsilon}_0}{1-\hat{\varepsilon}_0} \, \frac{h(u)}{1-H(u)}.$$

Since q is H-centered, we get

$$\int_0^u q(v)\, H(dv) = -\int_u^\infty q(v)\, H(dv).$$

But q is bounded by a constant c_2, which yields

$$\left|C_2(u,t)\right| + \left|C_3(u,t)\right| \le c_2 \left(\frac{1+\hat{\varepsilon}_0}{(1-\hat{\varepsilon}_0)^2} + \frac{1}{1-\hat{\varepsilon}_0}\right) \frac{h(u)}{1-H(u)}.$$

Overall, we can find a constant c_3 such that

$$\left| \frac{d}{dt} B(u,t)\Big|_{t=t_0} \right| \le c_3 \, \frac{h(u)}{1-H(u)}$$

uniformly in $-\hat{\varepsilon} < t_0 < \hat{\varepsilon}$. Since $0 < z < \tau_H$, this bound together with Lebesgue's DOM guarantees that

$$\frac{d}{dt}\Big(\int_0^z B(u,t)\, \mu(du)\Big)\Big|_{t=t_0} = \int_0^z \frac{d}{dt} B(u,t)\Big|_{t=t_0} \mu(du).$$

In total, we can bound now, uniformly over $-\hat{\varepsilon} < t_0 < \hat{\varepsilon}$,

$$I_1(z,t_0) \le \big(1-H(z)\big)^{-(1+\hat{\varepsilon}_0)/(1-\hat{\varepsilon}_0)} \, c_4 \, \log\big(1/(1-H(z))\big)\, h(z),$$

where $c_4(\theta, q) \equiv c_4$.

Similarly, we can find a bound $c_5(\theta, q) \equiv c_5$ such that, uniformly over $-\hat{\varepsilon} < t_0 < \hat{\varepsilon}$,

$$\big|\, I_2(z,t_0)\,\big| + \big|\, I_3(z,t_0)\,\big| \le \big(1-H(z)\big)^{-(1+\hat{\varepsilon}_0)/(1-\hat{\varepsilon}_0)} \, c_5 \, h(z).$$

Now choose $0 < \varepsilon_0 \le \hat{\varepsilon}_0$ such that $(1+\varepsilon_0)/(1-\varepsilon_0) = 1 + \kappa/2$ and $\varepsilon(q, \theta, \varepsilon_0) \equiv \varepsilon$. Note that there exists a constant $c_6 > 0$ such that, for all $0 < z < \tau_H$,

$$\Big|\, \log\Big(\frac{1}{1-H(z)}\Big) \Big| \le \Big(1-H(z)\Big)^{-\kappa/2} c_6.$$

Overall, this shows that uniformly in $-\varepsilon < t_0 < \varepsilon$ and $0 < z < \tau_H$

$$\left|\varphi(z)\, \frac{d}{dt} A(z,t)\Big|_{t=t_0}\right| \le \left|\varphi(z)\right| \big(1-H(z)\big)^{-(1+\kappa/2)} c_4 \Big(\big(1-H(z)\big)^{-\kappa/2} c_6 + c_5\Big)\, h(z).$$

According to (M_1), this bound is μ-integrable and Eq. (14) follows from Lebesgue's DOM. □

Proof of Remark 3.4. The proof of Lemma 3.3 shows that the derivative in the integrand of (14) is equal to

$$\gamma_0(z, p_{t_0}) m_{t_0}(z) h_{t_0}(z) \Big(-B_1(z,t_0) + B_2(z,t_0) + B_3(z,t_0)\Big) + \gamma_0(z, p_{t_0}) \big(\theta^\top Dm_{t_0}(z)\big) h_{t_0}(z) \\ + \gamma_0(z, p_{t_0}) m_{t_0}(z) q(z) h(z),$$

with

$$B_1(z,t_0) = \int_0^z \frac{\theta^\top Dm_{t_0}(u)}{1-H_{t_0}(u)} h_{t_0}(u)\, \mu(du),$$

$$B_2(z,t_0) = \int_0^z \frac{1-m_{t_0}(u)}{(1-H_{t_0}(u))^2} h_{t_0}(u) \int_0^u q(s)\, H(ds)\, \mu(du),$$

and

$$B_3(z,t_0) = \int_0^z \frac{1-m_{t_0}(u)}{1-H_{t_0}(u)} q(u) h(u)\, \mu(du).$$

Recall that for $t_0 = 0$ we have $m_0(z) = m(z, \theta_0)$, $Dm_0(z) = Dm(z, \theta_0)$, $p_0 = p$, $h_0(z) = h(z)$, and $\gamma_0(z, p_0) = \gamma_0(z)$. Furthermore, $q \in L_0(H)$ and therefore

$$\int_0^u q(s)\, H(ds) = -\int_u^\infty q(s)\, H(ds).$$

Overall, we get

$$\begin{aligned}\frac{d}{dt}\Psi_{\theta,q}(t)\Big|_{t=0} &= -\int\int_0^z \frac{\varphi(z)\gamma_0(z)\theta^\top Dm(u,\theta_0)}{1-H(u)}\,H(du)\,H^1(dz)\\ &\quad -\int\int_0^z\int_u^\infty \frac{\varphi(z)\gamma_0(z)q(v)}{(1-H(u))^2}\,H(dv)\,H^0(du)\,H^1(dz)\\ &\quad +\int\int_0^z \frac{\varphi(z)\gamma_0(z)q(u)}{1-H(u)}\,H^0(du)\,H^1(dz)\\ &\quad +\int \varphi(z)\gamma_0(z)\theta^\top Dm(z,\theta_0)\,H(dz)+\int \varphi(z)\gamma_0(z)q(z)\,H^1(dz).\end{aligned}$$

Now apply Fubini's theorem to complete the proof. $\square$

Proof of Lemma 3.5. Due to (A_2), the map $t \longrightarrow \sqrt{p_t(z,j)}$ is differentiable with respect to t at every point $-\varepsilon < t_0 < \varepsilon$ and for every (z, j).

To prove (ii), consider the function

$$\mathrm{I}: \quad (-\varepsilon,\varepsilon)\ni t \longrightarrow \mathrm{I}(t)\in\mathbb{R}^+$$

that maps t to the corresponding Fisher-information of p_t, that is,

$$\begin{aligned}\mathrm{I}(t) &= \int S_t^2 p_t\, d\mu\otimes\nu\\ &= \int S_t^2(z,1)p_t(z,1)\,\mu(dz)+\int S_t^2(z,0)p_t(z,0)\,\mu(dz)\\ &\equiv \mathrm{I}(t,1)+\mathrm{I}(t,0),\end{aligned}$$

where S_t is given under (4). For the first integral we get:

$$\begin{aligned}\mathrm{I}(t,1) &= \int\Big(\frac{\theta^\top Dm_t(z)}{m_t(z)}+\frac{q(z)h(z)}{h_t(z)}\Big)^2 m_t(z)h_t(z)\,\mu(dz)\\ &\le 2\left(\int\big(\theta^\top Dm_t(z)\big)^2/m_t(z)\,h_t(z)\,\mu(dz)+\int\big(q(z)h(z)\big)^2/h_t(z)m_t(z)\,\mu(dz)\right).\end{aligned}$$

Due to the construction of our sub-model, see Lemma 3.2(ii), we can find two constants, $c > 0$ and $C > 0$, such that $c\,h(z) \le h_t(z) \le C\,h(z)$, for every t and z. Thus, according to our assumption (A_2), the integrand of $\mathrm{I}(t, 1)$ is bounded by

$$\Big(\frac{\theta^\top Dm_t(z)}{m_t(z)}+\frac{q(z)h(z)}{h_t(z)}\Big)^2 m_t(z)h_t(z)\le 2\left(k^2\|\theta\|^2M(z)C+q^2(z)/c\right)h(z),$$

for every $z > 0$. Note that this bound does not depend on t, and that it is integrable with respect to μ.
Similarly, we can find a bound of the integrand $\mathrm{I}(t, 0)$, for every $z > 0$, which does not depend on t, and which is also integrable with respect to μ.
Overall, this shows that $\mathrm{I}(t)$ is well defined. Furthermore, according to the continuity (with respect to t) of its integrand and Lebesgue's DOM, the map $t \longrightarrow \mathrm{I}(t)$ is continuous at every point $-\varepsilon < t < \varepsilon$. This proves (ii).

Part (iii) follows from (i) and (ii) according to Lemma 7.6 of van der Vaart [22]. $\square$

Proof of Lemma 3.7. To prove part (i), we will show that each summand on the right hand side of (18) is in $L^2(p\,d\mu\otimes\nu)$. According to (M_3) and $\gamma_0(z) = 1/(1 - G(z))$, we get for the first summand

$$\begin{aligned}\int\big(\varphi(z)\gamma_0(z)1(j=1)\big)^2 p(z,j)\,\mu(du)\nu(dj) &= \int\big(\varphi(z)\gamma_0(z)\big)^2 m(z,\theta_0)\,H(dz)=\int\big(\varphi(z)\gamma_0(z)\big)^2 H^1(dz)\\ &= \int\varphi(z)^2\gamma_0(z)\,F(dz)<\infty.\end{aligned}$$

Due to (M_2), we get for the second summand

$$\begin{aligned}\int\big(\gamma_1(z)1(j=0)\big)^2 p(z,j)\,\mu(du)\nu(dj) &= \int\gamma_1^2(z)(1-m(z,\theta_0))\,H(dz)\\ &\le \int_0^{\tau_H}\Big(\frac{1}{1-H(z)}\int_z^{\tau_H}\big|\varphi(u)\big|F(du)\Big)^2 H(dz)\\ &\le \int_0^{\tau_H}\frac{1}{(1-H(z))^{1-2\kappa}}\Big(\int_0^{\tau_H}\frac{\big|\varphi(u)\big|}{(1-H(u))^{1/2+\kappa}}F(du)\Big)^2 H(dz)<\infty,\end{aligned}$$

and for the third summand

$$\begin{aligned}\int \gamma_2^2(z)p(z,j)\,\mu(du)\nu(dj) &= \int \gamma_2^2(z)\,H(dz)\\ &\le \int_0^{\tau_H}\Big(\int_0^z\int_u^\infty \frac{\varphi(s)\gamma_0(s)}{(1-H(u))^2}\,H^1(ds)\,H^0(du)\Big)^2 H(dz)\\ &\le \int_0^{\tau_H}\Big(\int_0^z (1-H(u))^{-3/2+\kappa}\,H(du)\Big)^2\\ &\quad\times\Big(\int_0^{\tau_H}\frac{|\varphi(s)|}{(1-H(s))^{1/2+\kappa}}\,F(ds)\Big)^2 H(dz) < \infty.\end{aligned}$$

This proves part (i) of the lemma.

For part (ii) of the lemma, we first observe that according to Lemma 3.5(ii) the score function $S_{\theta,q}$ is in $L^2(p\,d\mu\otimes\nu)$. Thus Hölder's inequality guarantees that the right hand side of Eq. (19) is well defined. It remains to show that the right hand side of Eq. (19) coincides with the right hand side of Eq. (17). Recall the definition of $S_{\theta,q}$ under Eq. (5) to get

$$\begin{aligned}\int \Psi^* S_{\theta,q}p\,d\mu d\nu &= \int \Psi^*(z,1)\Big(\theta^\top Dm(z,\theta_0)+m(z,\theta_0)q(z)\Big)\,H(dz)\\ &\quad+\int \Psi^*(z,0)\Big(-\theta^\top Dm(z,\theta_0)+\big(1-m(z,\theta_0)\big)q(z)\Big)\,H(dz)\\ &= \int \varphi(z)\gamma_0(z)\theta^\top Dm(z,\theta_0)\,H(dz)+\int \varphi(z)\gamma_0(z)m(z,\theta_0)q(z)\,H(dz)\\ &\quad-\int \gamma_1(z)\theta^\top Dm(z,\theta_0)\,H(dz)+\int \gamma_1(z)(1-m(z,\theta_0))q(z)\,H(dz)-\int \gamma_2(z)q(z)\,H(dz).\end{aligned}$$

Since $H^1(dz)=m(z,\theta_0)H(dz)$ and $H^0(dz)=(1-m(z,\theta_0))H(dz)$, the proof of the lemma is complete. □

Proof of Lemma 3.8. A similar argument as in the proof of Lemma 3.7 shows that $\hat\Psi\in L^2(p\,\mu\otimes\nu)$. Furthermore,

$$\int K(z,j)\hat\theta^\top Dm(z,\theta_0)p(z,j)\,\mu(dz)\nu(dj)=0$$

and, as discussed in Remark 2.2 of Dikta et al. [7], $\mathbb{E}(\hat q(Z))=0$. Thus $\hat\Psi\in L_0^2(p\,d\mu\otimes\nu)$. The specific form of $\hat\Psi$ given in Eq. (20) suggests that $\hat\Psi$ is itself a score function. Indeed, $\hat q$ is H-centered but not necessarily bounded. Nevertheless, we can easily find a sequence of H-centered, bounded functions which tends to $\hat q$. Since $\mathcal{T}$ is closed, we finally get $\hat\Psi\in\mathcal{T}$. This proves part (i) of the lemma.

If Eq. (21) holds, part (ii) of the lemma is a consequence of (i) and Eq. (19) since $\mathcal{T}$ is the closed linear space spanned by the score functions. Thus it remains to prove Eq. (21). Choose a H-centered, bounded, q and $\theta\in\mathbb{R}^k$. Then

$$\begin{aligned}\int \hat\Psi S_{\theta,q}p\,d\mu d\nu &= \int \hat\Psi(z,j)\Big(K(z,j)\theta^\top Dm(z,\theta_0)+q(z)\Big)p(z,j)\,\mu(dz)\nu(dj)\\ &= \int\Big(\frac{\hat\theta^\top Dm(z,\theta_0)}{m(z,\theta_0)}+\hat q(z)\Big)\Big(\frac{\theta^\top Dm(z,\theta_0)}{m(z,\theta_0)}+q(z)\Big)\,H^1(dz)\\ &\quad+\int\Big(-\frac{\hat\theta^\top Dm(z,\theta_0)}{1-m(z,\theta_0)}+\hat q(z)\Big)\Big(-\frac{\theta^\top Dm(z,\theta_0)}{1-m(z,\theta_0)}+q(z)\Big)\,H^0(dz).\end{aligned}$$

Decompose the last two integrals to get for their parts

$$\int \frac{\hat\theta^\top Dm(z,\theta_0)\,\theta^\top Dm(z,\theta_0)}{m(z,\theta_0)^2}\,H^1(dz)+\int\frac{\hat\theta^\top Dm(z,\theta_0)\,\theta^\top Dm(z,\theta_0)}{(1-m(z,\theta_0))^2}\,H^0(dz)=\hat\theta^\top I_0\theta,$$

$$\int \frac{\hat q(z)\theta^\top Dm(z,\theta_0)}{m(z,\theta_0)}\,H^1(dz)-\int\frac{\hat q(z)\theta^\top Dm(z,\theta_0)}{1-m(z,\theta_0)}\,H^0(dz)=0,$$

$$\int \frac{q(z)\hat\theta^\top Dm(z,\theta_0)}{m(z,\theta_0)}\,H^1(dz)-\int\frac{q(z)\hat\theta^\top Dm(z,\theta_0)}{1-m(z,\theta_0)}\,H^0(dz)=0,$$

and finally

$$\int \hat q(z)q(z)H^1(dz)+\int \hat q(z)q(z)H^0(dz)=\int \hat q(z)q(z)H(dz).$$

Overall, we obtain

$$\int \hat{\Psi} S_{\theta,q} p \, d\mu d\nu = \hat{\theta}^\top I_0 \theta + \int \hat{q} q \, dH.$$

Plugging in the definition of $\hat{\theta}$ yields

$$\hat{\theta}^\top I_0 \theta = \int \Big(\varphi(z)\gamma_0(z) - \gamma_1(z) \Big) \theta^\top Dm(z, \theta_0) \, H(dz).$$

Furthermore, from the definition of $\hat{q}$ we derive

$$\int \hat{q} q \, dH = \int \varphi(z)\gamma_0(z) q(z) \, H^1(dz) + \int \gamma_1(z) q(z) \, H^0(dz) - \int \gamma_2(z) q(z) \, H(dz) - \beta_0 \int q(z) \, H(dz).$$

Since q is H-centered, the last integral on the right hand side vanishes. Collecting all the terms derived for $\hat{\theta}^\top I_0 \theta$ and $\int \hat{q} q \, dH$, and comparing the result with Eq. (17) from Remark 3.4, finalizes the proof of this lemma. □

Proof of Theorem 3.9. Let $\hat{\beta}_n$ be a RAL estimator of β_0 with IF $A \in L_0^2(p \, d\mu \otimes \nu)$. Fix $\theta \in \mathbb{R}^k$ and a bounded, H-centered, q and the corresponding sub-model $\mathcal{P}_{\theta,q} = (p_t)_{-\varepsilon < t < \varepsilon}$.

Due to the given model assumptions and Lebesgue's DOM, we derive that the map $t \longrightarrow \int A^2 p_t \, d\mu \otimes \nu$ is continuous and $t \longrightarrow \int A p_t \, d\mu \otimes \nu$ is continuously differentiable with

$$\frac{d}{dt} \Big(\int A p_t \, d\mu \otimes \nu \Big) \Big|_{t=0} = \int A \, \frac{d}{dt} p_t \Big|_{t=0} d\mu \otimes \nu = \int A S_{\theta,q} p \, d\mu \otimes \nu. \tag{24}$$

The next step in our proof is based on a contiguity argument. For this we will switch from our IID observations to a triangular setup, that is, we will assume that

$$(Z_{1,n}, \delta_{1,n}), \ldots, (Z_{n,n}, \delta_{n,n})$$

are IID, where the DF of $(Z_{1,n}, \delta_{1,n})$ has a $\mu \otimes \nu$-density $p_{1/\sqrt{n}}$. For notational reason, we will index probabilities, expectations, etc. with the corresponding density.

Since $\hat{\beta}_n$ is an AL estimator, we have, for $n \to \infty$,

$$\sqrt{n}\big(\hat{\beta}_n - \beta_0\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A(Z_{i,n}, \delta_{i,n}) + o_p(1).$$

According to the CLT, this implies that the DF of $\sqrt{n}\big(\hat{\beta}_n - \beta_0\big)$, based on the p-density, tends to the DF of a centered normal variable with variance $\sigma^2 = \int A^2 p \, d\mu \otimes \nu$.

Due to Hellinger-differentiability (see Lemma 3.5), we can switch measures and get from Remark 3.6 that

$$\sqrt{n}\big(\hat{\beta}_n - \beta_0\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A(Z_{i,n}, \delta_{i,n}) + o_{p_{1/\sqrt{n}}}(1).$$

This results in the following asymptotic $p_{1/\sqrt{n}}$ representation

$$\begin{aligned}
\big(\hat{\beta}_n - \Psi_{\theta,q}(1/\sqrt{n})\big) &= \sqrt{n}\big(\hat{\beta}_n - \Psi_{\theta,q}(0)\big) - \sqrt{n}\big(\Psi_{\theta,q}(1/\sqrt{n}) - \Psi_{\theta,q}(0)\big) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} A(Z_{i,n}, \delta_{i,n}) - \sqrt{n}\big(\Psi_{\theta,q}(1/\sqrt{n}) - \Psi_{\theta,q}(0)\big) + o_{p_{1/\sqrt{n}}}(1) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(A(Z_{i,n}, \delta_{i,n}) - \int A p_{1/\sqrt{n}} \, d\mu \otimes \nu \Big) + \sqrt{n} \int A p_{1/\sqrt{n}} \, d\mu \otimes \nu \\
&\quad - \sqrt{n}\big(\Psi_{\theta,q}(1/\sqrt{n}) - \Psi_{\theta,q}(0)\big) + o_{p_{1/\sqrt{n}}}(1).
\end{aligned}$$

The assumed regularity of our estimator $\hat{\beta}_n$ guarantees that the DF of $\sqrt{n}\big(\hat{\beta}_n - \Psi_{\theta,q}(1/\sqrt{n})\big)$ based on the $p_{1/\sqrt{n}}$-density tends to the DF of a centered normal variable with variance σ^2. Furthermore, due to the continuity of the map $t \longrightarrow \mathrm{VAR}_{p_t}(A)$, we can apply the CLT for triangular arrays, see Theorem 27.2 of Billingsley [3], to get that the DF of

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big(A(Z_{i,n}) - \int A p_{1/\sqrt{n}} \, d\mu \otimes \nu \Big),$$

based on the $p_{1/\sqrt{n}}$-density, tends to a centered normal DF with the same variance σ^2. Consequently,

$$\sqrt{n}\int Ap_{1/\sqrt{n}}\,d\mu\otimes\nu-\sqrt{n}\big(\Psi_{\theta,q}(1/\sqrt{n})-\Psi_{\theta,q}(0)\big)\longrightarrow 0,$$

as $n\to\infty$. Note that $\int Ap\,d\mu\otimes\nu=0$ to derive that

$$\sqrt{n}\Big(\int Ap_{1/\sqrt{n}}\,d\mu\otimes\nu-\int Ap\,d\mu\otimes\nu\Big)-\sqrt{n}\big(\Psi_{\theta,q}(1/\sqrt{n})-\Psi_{\theta,q}(0)\big)\longrightarrow 0,$$

as $n\to\infty$. According to Eqs. (21) and (24) this yields

$$\int AS_{\theta,q}p\,d\mu\otimes\nu=\int\hat{\Psi}S_{\theta,q}p\,d\mu\otimes\nu. \tag{25}$$

Now apply the Cauchy–Schwarz inequality to get, from Eq. (25),

$$\begin{aligned}\int A^2p\,d\mu\otimes\nu\int S^2_{\theta,q}p\,d\mu\otimes\nu &\geq \Big(\int AS_{\theta,q}p\,d\mu\otimes\nu\Big)^2\\ &=\Big(\int\hat{\Psi}S_{\theta,q}p\,d\mu\otimes\nu\Big)^2.\end{aligned}$$

Consequently,

$$\int A^2p\,d\mu\otimes\nu\geq\sup_{S\in\mathcal{T}}\Big[\Big(\int\hat{\Psi}Sp\,d\mu\otimes\nu\Big)^2\Big/\int S^2p\,d\mu\otimes\nu\Big]. \tag{26}$$

Note that the right hand side of (26) is identical to the squared operator norm of the continuous linear functional defined through the Riesz-density $\hat{\Psi}\in\mathcal{T}$. According to Riesz representation theorem, this operator norm is identical to the L^2-norm of the Riesz-density $\hat{\Psi}$, compare Theorem 2.E of Zeidler [27]. Therefore

$$\int\hat{\Psi}^2p\,d\mu\otimes\nu=\sup_{S\in\mathcal{T}}\Big[\Big(\int\hat{\Psi}Sp\,d\mu\otimes\nu\Big)^2\Big/\int S^2p\,d\mu\otimes\nu\Big]. \tag{27}$$

To complete the proof of the theorem, notice that the information bound given in the theorem under Eq. (22) is equal to $\int\hat{\Psi}^2p\,d\mu\otimes\nu$. □

Acknowledgments

This research was supported in part by the Department of Mathematical Sciences of the University of Wisconsin, Milwaukee, while the author was visiting UW-Milwaukee in 2012. The author is also grateful to Professor Jugal Ghorai from this department for his constructive comments.

References

[1] P.J. Bickel, K.A. Doksum, Mathematical Statistics, Holden-Day Inc., San Francisco, Calif., 1976.
[2] P.J. Bickel, C.A.J. Klaassen, Y. Ritov, J.A. Wellner, Efficient and Adaptive Estimation for Semiparametric Models, Springer-Verlag, New York, 1998.
[3] P. Billingsley, Probability and Measure, third ed., in: Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc, New York, 1995.
[4] N. Breslow, J. Crowley, A large sample study of the life table and product limit estimates under random censorship, Ann. Statist. 2 (1974) 437–453.
[5] G. Dikta, On semiparametric random censorship models, J. Statist. Plann. Inference 66 (2) (1998) 253–279.
[6] G. Dikta, The strong law under semiparametric random censorship models, J. Statist. Plann. Inference 83 (1) (2000) 1–10.
[7] G. Dikta, J. Ghorai, C. Schmidt, The central limit theorem under semiparametric random censorship models, J. Statist. Plann. Inference 127 (1–2) (2005) 23–51.
[8] G. Dikta, M. Kvesic, C. Schmidt, Bootstrap approximations in model checks for binary data, J. Amer. Statist. Assoc. 101 (474) (2006) 521–530.
[9] B. Efron, The two-sample problem with censored data, in: Proceedings of the 5th Berkeley Symposium, vol. 4, 1967, pp. 831–853.
[10] R. Gill, Large sample behaviour of the product-limit estimator on the whole line, Ann. Statist. 11 (1) (1983) 49–58.
[11] S. Johansen, The product limit estimator as maximum likelihood estimator, Scandinavian J. Statist. 5 (4) (1978) 195–199.
[12] E.L. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, J. Amer. Statist. Assoc. 53 (1958) 457–481.
[13] V. Susarla, J. Van Ryzin, Large sample theory for an estimator of the mean survival time from censored samples, Ann. Statist. 8 (5) (1980) 1002–1016.
[14] A. Schick, V. Susarla, H. Koul, Efficient estimation of functionals with censored data, Statist. Decisions 6 (4) (1988) 349–360.
[15] W. Stute, The central limit theorem under random censorship, Ann. Statist. 23 (2) (1995) 422–439.
[16] W. Stute, J.-L. Wang, The strong law under random censorship, Ann. Statist. 21 (3) (1993) 1591–1607.
[17] S. Subramanian, The missing censoring-indicator model of random censorship, in: Advances in Survival Analysis, in: Handbook of Statist., vol. 23, Elsevier, Amsterdam, 2004, pp. 123–141.
[18] S. Subramanian, Multiple imputations and the missing censoring indicator model, J. Multivariate Anal. 102 (1) (2011) 105–117.
[19] S. Subramanian, P. Zhang, Model-based confidence bands for survival functions, J. Statist. Plann. Inference 143 (7) (2013) 1166–1185.
[20] L. Sun, L. Zhu, A semiparametric model for truncated and censored data, Statist. Probab. Lett. 48 (3) (2000) 217–227.
[21] A.A. Tsiatis, Semiparametric Theory and Missing Data, in: Springer Series in Statistics, Springer, New York, 2006.
[22] A.W. van der Vaart, Asymptotic statistics, in: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3, Cambridge University Press, Cambridge, 1998.

[23] A. van der Vaart, On differentiable functionals, Ann. Statist. 19 (1) (1991) 178–204.
[24] J.A. Wellner, Asymptotic optimality of the product limit estimator, Ann. Statist. 10 (2) (1982) 595–602.
[25] S. Yang, Efficient robust estimation of parameter in the random censorship model, Statist. Probab. Lett. 10 (5) (1990) 419–426.
[26] S. Yang, Minimum Hellinger distance estimation of parameter in the random censorship model, Ann. Statist. 19 (2) (1991) 579–602.
[27] E. Zeidler, Applied functional analysis, in: Applied Mathematical Sciences, vol. 108, Springer-Verlag, New York, 1995.