INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2004-05

Product-Limit Estimators of the Survival Function with Twice Censored Data

V. PATILEA[1]
J.-M. ROLIN[2]

[1]CREST-ENSAI, Campus de Ker Lann, Rue Blaise Pascal, BP 37203, 35172 BRUZ Cedex, France. Email : valentin.patilea@ensai.fr
[2] Institut de Statistique, Université Catholique de Louvain, 20 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgique.
Supported by PARC Grant n° 98/03-217 of the Belgian Government.

Product-Limit Estimators of the Survival Function with Twice Censored Data

Valentin Patilea* and Jean-Marie Rolin†

*CREST - ENSAI, Campus de Ker Lann, Rue Blaise Pascal BP 37203, 35172 Bruz cedex, France. Email: valentin.patilea@ensai.fr

†Institut de Statistique, Université Catholique de Louvain, 20 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgique. Supported by PARC Grant no. 98/03-217 of the Belgian Government

Abstract

A model for competing (resp. complementary) risks survival data where the failure time can be left (resp. right) censored is proposed. Product-limit estimators for the survival functions of the individual risks are derived. We deduce the strong convergence of our estimators on the whole real half-line without any additional assumption and their asymptotic normality under conditions concerning only the observed distribution. When the observations are generated according to the double censoring model introduced by Turnbull (1974), the product-limit estimators represent upper and lower bounds for Turnbull's estimator.

Key words: twice censoring, competing and complementary risks, hazard functions, product-limit estimator, strong convergence, weak convergence, delta-method.

MSC 2000 : 62N01, 62G05, 62N02.

Résumé

Nous proposons un modèle pour les risques concurrents (resp. complémentaires) dans lequel la durée de survie peut être censurée à gauche (resp. à droite). Nous obtenons des estimateurs de type Kaplan-Meier pour les fonctions de survie. La convergence presque sure sur la droite réelle et la normalité asymptotique de nos estimateurs sont démontrées sous des conditions faibles. Nos estimateurs, de type Kaplan-Meier, peuvent être considérés comme une borne supérieure et une borne inférieure pour l'estimateur de Turnbull (1974).

Mots clefs : double censure, risques concurrents, risques complémentaires, fonctions de risque, estimateur de Kaplan-Meier, convergence presque sure, normalité asymptotique, méthode delta.

1 Introduction.

Consider the problem of nonparametric inference with competing risks survival data. The novelty we propose is that the failure time can be left-censored, for instance at the time the study starts. For simplicity, we consider two distinct competing risks of failure, the extension to more than two competing risks being straight. Let T and V_1 denote the latent independent lifetimes for each cause of failure. The failure time is $\min(T, V_1)$ and it can be censored from the left by a censoring time U_1. The observations are independent copies of a lifetime Y, a finite nonnegative random variable, and a discrete random variable A with values in $\{0, 1, 2\}$, where 2 indicates a left-censored failure time while 0 and 1 correspond to an observation equal to T and V_1, respectively. If T is the lifetime of interest, we say that Y is a twice censored observation of T. Associated with the problem of competing risks is the dual problem of complementary risks where the observed failure time is the maximum of the lifetimes for each cause of failure (e.g., Basu and Ghosh (1980)). The extension we consider here is that the failure time can be right-censored, for instance at the time the experience ends.

By the plug-in (or substitution) principle applied for the empirical distribution, the nonparametric estimation of the distribution of a latent lifetime of interest is straightforward as soon as this distribution can be expressed as an explicit function of the distribution of the observed variables. The two models we propose in this paper are shown to allow for explicit inversion formulae, that is the latent distributions of interest are explicit functionals of the distribution of the observations.

In section 2 we introduce our latent models while in section 3 we provide the inversion formulae. In section 4 we compare our model with the doubly censored data latent model proposed by Turnbull (1974). We show that the inversion formulae provide lower and upper bounds for the distribution of interest identified by Turnbull's model. Applying the inversion formulae to the empirical distribution, we deduce in section 5 the product-limit estimators. Moreover, we analyze the sample selection bias induced when discarding from the sample observations corresponding to censored failure times. In sections 6 and 7 we deduce the almost sure uniform convergence and the asymptotic normality for our functionals.

2 Latent variables models.

The random variables we consider take values in $\overline{\mathbb{R}}^+ = [0, \infty]$ endowed with $\overline{\mathcal{B}}^+$ the Borel $\sigma-$field. If X is such a variable, F_X denotes its distribution.

For the first latent model considered (call it Model 1), let T and V_1 be two lifetimes and let U_1 be a left-censoring time. Assume that T, V_1 and U_1 are independent. Suppose that Y and A are observed, where $Y = \max[\min(T, V_1), U_1]$ and

$$A = \begin{cases} 0 & \text{if} \quad U_1 < T \leq V_1; \\ 1 & \text{if} \quad U_1 < V_1 < T; \\ 2 & \text{if} \quad \min(T, V_1) \leq U_1. \end{cases}$$

Define the observed subdistributions of Y as

$$H_k(B) = P[Y \in B,\ A = k], \qquad k = 0, 1, 2,$$

where B is a Borel set in $[0, \infty]$; the distribution of Y is $H = H_0 + H_1 + H_2$. In Model 1, the subdistributions of Y can be expressed in terms of the distributions of the latent variables as follows:

$$\begin{aligned} H_0(dt) &= F_{U_1}([0, t))\ F_{V_1}([t, \infty])\ F_T(dt) \\ H_1(dt) &= F_{U_1}([0, t))\ F_T((t, \infty])\ F_{V_1}(dt) \\ H_2(dt) &= \{1 - F_T((t, \infty])\ F_{V_1}((t, \infty])\}\ F_{U_1}(dt) \end{aligned} \tag{2.1}$$

[necessarily $H_0(\{0\}) = H_1(\{0\}) = 0$]. If $S_1 = \min(T, V_1)$ and $H_{01} = H_0 + H_1$, the three equations imply

$$H_{01}(dt) = F_{U_1}([0, t))\ F_{S_1}(dt), \qquad H_2(dt) = F_{S_1}([0, t])\ F_{U_1}(dt). \tag{2.2}$$

This indicates that the problem of inverting the model, i.e., expressing the distributions of the latent variables in terms of the subdistributions of Y, can be solved in two steps. First, determine the distributions of U_1 and S_1 like in an independent left-censoring model. Next, use these distributions to recover the distribution of T.

As an application of Model 1, consider a reliability system which consists of three components U_1, T and V_1, with T and V_1 in series and U_1 in parallel with this series system (see, e.g., Meeker and Escobar (1998), chapter 15). The lifetimes of U_1, T and V_1 are independent and when the system fails we are able to determine which component failed at the same time as the system. Morales *et al.* (1991) propose the application of this model to study a certain cause death for the trees in a farm.

For our second latent model (call it Model 2), let U_2 and T be two lifetimes and let V_2 be a right-censoring time. Assume that T, U_2 and V_2 are independent. The observed variables are Y and A, where $Y = \min[\max(T, U_2),\, V_2]$ and

$$A = \begin{cases} 0 & \text{if} \quad U_2 < T \le V_2 \\ 1 & \text{if} \quad V_2 < \max(U_2, T) \\ 2 & \text{if} \quad T \le U_2 \le V_2. \end{cases}$$

In Model 2, the relationship between the subdistributions of Y and the distributions of the latent variables is described by the equations

$$\begin{aligned} H_0(dt) &= F_{U_2}([0,t))\; F_{V_2}([t,\infty])\; F_T(dt) \\ H_1(dt) &= \{1 - F_T([0,t])\; F_{U_2}([0,t])\}\; F_{V_2}(dt) \\ H_2(dt) &= F_T([0,t])\; F_{V_2}([t,\infty])\; F_{U_2}(dt) \end{aligned} \tag{2.3}$$

[necessarily $H_0(\{0\}) = 0$]. If $S_2 = \max(U_2, T)$ and $H_{02} = H_0 + H_2$, we obtain

$$H_{02}(dt) = F_{V_2}([t,\infty])\; F_{S_2}(dt), \qquad H_1(dt) = F_{S_2}((t,\infty])\; F_{V_2}(dt). \tag{2.4}$$

These relations show that Model 2 can be inverted in two steps. First, as in an independent right-censoring model, recover the distributions of V_2 and S_2 from H_{02} and H_1. Second, determine the distribution of T.

Model 2 can be interpreted as follows: consider a system consisting of three components U_2, T and V_2 with independent lifetimes. Put T and U_2 in parallel and V_2 in series with this parallel system (see also Doss *et al.* (1989), page 767). Again, assume that we are able to determine which component failed at the same time as the system.

3 Inversion formulae.

Hereafter, if μ is a nonnegative measure on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$, $\mu(t)$ is a short notation for $\mu(\{t\})$. Recall that if F is a probability distribution on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$, the associated hazard measure is $L([0,t]) = -\ln F((t,\infty])$. Two more hazard measures can be defined

$$L^-(dt) = \frac{F(dt)}{F([t,\infty])} \qquad \text{and} \qquad L^+(dt) = \frac{F(dt)}{F((t,\infty])}$$

that we call the predictable and the unpredictable hazard measure, respectively. The three hazard measures have the same continuous parts. Moreover, their point masses are in bijection: $L(t) = -\ln[1 - L^-(t)] = \ln[1 + L^+(t)]$. The probability distribution F can be expressed as

$$F((t,\infty]) = \exp\{-L([0,t])\} = \prod_{[0,t]}(1 - L^-(ds)) = [\prod_{[0,t]}(1 + L^+(ds))\,]^{-1},$$

where $\prod$ is the product-integral [e.g., Gill and Johansen (1990)]. The mass of L at infinity is irrelevant for F and $F(\infty) = \exp\{-L([0,\infty))\}$.

Similarly, by reversing time, the reverse hazard measure associated to F is $M((t,\infty]) = -\ln F([0,t])$. Moreover, the predictable and unpredictable reverse hazard measures are defined as

$$M^-(dt) = \frac{F(dt)}{F([0,t])} \qquad \text{and} \qquad M^+(dt) = \frac{F(dt)}{F([0,t))},$$

respectively. The three reverse hazard measures have the same continuous parts and their point masses satisfy $M(t) = -\ln[1 - M^-(t)] = \ln[1 + M^+(t)]$. We have

$$F([0,t]) = \exp\{-M((t,\infty])\} = \prod_{(t,\infty]}(1 - M^-(ds)) = [\prod_{(t,\infty]}(1 + M^+(ds))\,]^{-1}.$$

The mass of M at zero is irrelevant for F. Moreover, $F(0) = \exp\{-M((0,\infty])\}$.

Given a nonnegative measure on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$, we can always define a probability distribution on the same space by considering this measure as being one of L, L^- or L^+ (resp. M, M^- or M^+) and using the relations above. For instance, in the independent right-censoring model one defines $L^-(dt) = H_0(dt)/H([t,\infty])$, with H_0 the subdistribution of the uncensored data. Then, by the equations of the model, the distribution corresponding to this L^- is nothing else than the distribution of the lifetime of interest. The reverse hazard measures M, M^- and M^+ are the counterparts of L, L^- and L^+ to be used in left-censoring models.

We can invert our models using the hazard measures above. Since, apart mild conditions at the origin, the inversion formulae below apply to *any* subdistributions (H_0, H_1, H_2), we deduce them without any reference to the latent variables.

For inverting Model 1 assume $H_0(0) = H_1(0) = 0$. In view of (2.2), proceed as for inverting a left-censoring model and define the predictable reverse hazard measures

$$M_2^-(dt) = \frac{H_2(dt)}{H([0,t])}, \qquad M_{01}^-(dt) = \frac{H_{01}(dt)}{H([0,t)) + H_{01}(t)} \tag{3.1}$$

and let F_2^1 and F_{01}^1 be the corresponding distributions. By this definition we have $H([0,t]) = F_2^1([0,t])F_{01}^1([0,t])$. In the second step of the inversion, note that the first equation in (2.1) and the definition of S_1 imply $H_0(dt)/F_{U_1}([0,t))F_{S_1}([t,\infty]) = F_T(dt)/F_T([t,\infty])$. This suggests to define the predictable hazard measure

$$L_T^{1-}(dt) = \frac{H_0(dt)}{F_2^1([0,t))\ F_{01}^1([t,\infty])}. \tag{3.2}$$

Let F_T^1 be its associated distribution.

For Model 2 assume $H_0(0) = 0$. Look at the relation (2.4) and, exactly as in a right-censoring model, define the predictable hazard measures

$$L_{02}^{-}(dt) = \frac{H_{02}(dt)}{H([t,\infty])}, \qquad L_1^{-}(dt) = \frac{H_1(dt)}{H((t,\infty]) + H_1(t)}.$$

Let F_{02}^2 and F_1^2 denote the corresponding distributions. Clearly, we have $H((t,\infty]) = F_1^2((t,\infty])\ F_{02}^2((t,\infty])$. In the second step of the inversion, use the first equation in (2.3) and the definition of S_2 to deduce $H_0(dt)/\{F_{V_2}([t,\infty])F_{S_2}([0,t)) + H_0(t)\} = F_T(dt)/F_T([0,t])$. Consequently, define the predictable reverse hazard measure

$$M_T^{2-}(dt) = \frac{H_0(dt)}{F_1^2([t,\infty])\ F_{02}^2([0,t)) + H_0(t)} \tag{3.3}$$

and let F_T^2 be its associated distribution.

Now, let us consider the identification problem, that is if Model i is correct, we look for conditions ensuring that $F_T^i = F_T$ on $\overline{\mathbb{R}}^+$. Define the support of μ a nonnegative measure on $[0,\infty]$ as $supp(\mu) = \{t \,:\, \mu([0,\,t])\,\mu([t,\infty]) > 0\}$. The support of a variable X is the support of F_X. Let $B_i = \{t \,:\, F_{U_i}([0,t))\ F_{V_i}([t,\infty]) > 0\}$, $i = 1,2$. Deduce from (3.2) [resp. (3.3)] that the support of L_T^{1-} [resp. M_T^{2-}] is equal to the support of H_0. Since $supp(H_0) = B_i \cap supp(F_T)$, deduce that

$$F_T^i = F_T \quad \text{on} \quad \overline{\mathbb{R}}^+ \qquad \Leftrightarrow \qquad supp(F_T) \subset B_i, \qquad i = 1,2.$$

4 Comparisons with the doubly censored data model.

The models we propose are closely related to the model for doubly (left and right) censored observations introduced by Turnbull (1974). In Turnbull's model the lifetime T is independent of the censoring variables (L,R) and $L \leq R$. The observations are independent copies of Y and A where

$$Y = \max[\min(T, R),\ L] = \min[\max(T, L),\ R],$$

$$A = \begin{cases} 0 & \text{if} \quad L < T \le R \qquad (no\ censoring) \\ 1 & \text{if} \quad (\,L \le\,)\ R < T \qquad (right\ censoring) \\ 2 & \text{if} \quad T \le L\ (\,\le R\,) \qquad (left\ censoring)\ . \end{cases}$$

If $H_k(dt) = P(Y \in dt, A = k)$, $k = 0, 1, 2$, the equations of the model are

$$\begin{aligned} H_0(dt) &= \{F_L([0,t)) - F_R([0,t))\}\ F_T(dt) \\ H_1(dt) &= F_T((t,\infty])\ F_R(dt) \\ H_2(dt) &= F_T([0,t])\ F_L(dt). \end{aligned} \tag{4.1}$$

Note that the assumptions of the model imply

$$H([0,t]) = F_L([0,t])F_T([0,t]) + F_R([0,t])F_T((t,\infty]). \tag{4.2}$$

In Turnbull's model T is censored from the left by L and from the right by R and the observation Y is always the variable in the middle. This is different from the censoring mechanisms we consider: in Model 1 the variable $\min(T, V_1)$ is left-censored, while in Model 2 the variable $\max(U_2, T)$ is right-censored.

Turnbull (1974) proposed a nonparametric maximum likelihood estimator that can be obtained as the implicit solution of the equations (4.1). The implicit definition of Turnbull's estimator makes its asymptotic properties quite difficult [see Gu and Zhang (1993)]. Moreover, a numerical algorithm is needed for the applications.

We are interested in the relationship between our F_T^1, F_T^2 and F_T identified by Turnbull's model. In fact, for any subdistributions H_0, H_1 and H_2 with $H_0(0) = H_1(0) = 0$,

$$F_T^1([0,t]) \le F_T([0,t]) \le F_T^2([0,t]), \qquad \forall t \ge 0,$$

where F_T is the distribution of T identified by Turnbull's model. Indeed, in Model 1, use definition (3.2) and $H([0,t]) = F_2^1([0,t])F_{01}^1([0,t])$ to write

$$L_T^{1-}(dt) = \frac{H_0(dt)}{F_2^1([0,t)) - H([0,t))}.$$

In Turnbull's model [relations (4.1) and (4.2)] we have

$$L_T^{-}(dt) = \frac{H_0(dt)}{F_L([0,t)) - H([0,t))}.$$

Next, the definition of M_2^-, the last equation in (4.1) and equation (4.2) imply

$$M_2^-(dt) = \frac{F_L([0,t])\, F_T([0,t])}{F_L([0,t])\, F_T([0,t]) + F_R([0,t])\, F_T((t,\infty])}\; M_L^-(dt).$$

Deduce that the measure M_2^- is smaller than the measure M_L^-. Therefore, $F_2^1([0,t)) \geq F_L([0,t))$, $\forall t \geq 0$. Hence, the measure L_T^{1-} is smaller than the measure L_T^- which implies $F_T^1([0,t]) \leq F_T([0,t])$, $\forall t \geq 0$.

On the other hand, for Model 2, use the general relationship between M^+ and M^-, the definition (3.3) and $H((t,\infty]) = F_1^2((t,\infty])\, F_{02}^2((t,\infty])$ and write

$$M_T^{2+}(dt) = \frac{H_0(dt)}{F_1^2([t,\infty]) - H([t,\infty])}.$$

Meanwhile, in Turnbull's model [relations (4.1) and (4.2)]

$$M_T^+(dt) = \frac{H_0(dt)}{F_R([t,\infty]) - H([t,\infty])}.$$

Next, use the definition of L_1^-, the general relationship between L^+ and L^-, the second equation in (2.3) and the equality $H((t,\infty]) = F_L((t,\infty])\, F_T([0,t]) + F_R((t,\infty])\, F_T((t,\infty])$ [this is a consequence of (4.2)] to deduce

$$L_1^+(dt) = \frac{F_R((t,\infty])\, F_T((t,\infty])}{F_L((t,\infty])\, F_T([0,t]) + F_R((t,\infty])\, F_T((t,\infty])}\; L_R^+(dt).$$

Clearly, the measure L_1^+ is smaller than the measure L_R^+ and thus $F_1^2([t,\infty]) \geq F_R([t,\infty])$, $\forall t \geq 0$. Hence, the measure M_T^{2+} is smaller than the measure M_T^+ and this implies $F_T^2([0,t]) \geq F_T([0,t])$, $\forall t \geq 0$.

5 Product-limit estimators.

If we replace in the expressions of F_T^1 and F_T^2 the subdistributions H_0, H_1 and H_2 by their empirical counterparts we obtain the product-limit estimators F_{nT}^1 and F_{nT}^2, respectively. For this denote by $\{Z_j : 1 \leq j \leq M\}$ the distinct values in increasing order of Y_i in a set of independent identically distributed (iid) observations $\{(Y_i, A_i) : 1 \leq i \leq n\}$. Define

$$D_{kj} = \sum_{1\leq i\leq n} 1\!\!I_{\{Y_i = Z_j, A_i = k\}}, \qquad N_j = \sum_{1\leq i\leq n} 1\!\!I_{\{Y_i \leq Z_j\}}, \qquad \overline{N}_j = \sum_{1\leq i\leq n} 1\!\!I_{\{Y_i \geq Z_j\}},$$

$k = 0,1,2$. This means in particular, that D_{0j} (resp. D_{1j}) (resp. D_{2j}) are the number of uncensored (resp. right censored) (resp. left censored) observations at time Z_j. With these definitions, the product-limit estimator of F_T in Model 1 is

$$F_{nT}^1\left((Z_j,\infty]\right) = \prod_{1\le k\le j}\{1-\frac{D_{0k}}{U_{k-1}-N_{k-1}}\},$$

where

$$U_{j-1} = n\prod_{j\le k\le M}\{1-\frac{D_{2k}}{N_k}\}.$$

The product-limit estimator of F_T in Model 2 is given by

$$F_{nT}^2([0,Z_j]) = \prod_{j<k\le M}\{1-\frac{D_{0k}}{V_k-\overline{N}_k+D_{0k}}\},$$

where

$$V_j = n\prod_{1\le k\le j}\{1-\frac{D_{1k}}{\overline{N}_{k+1}+D_{1k}}\}.$$

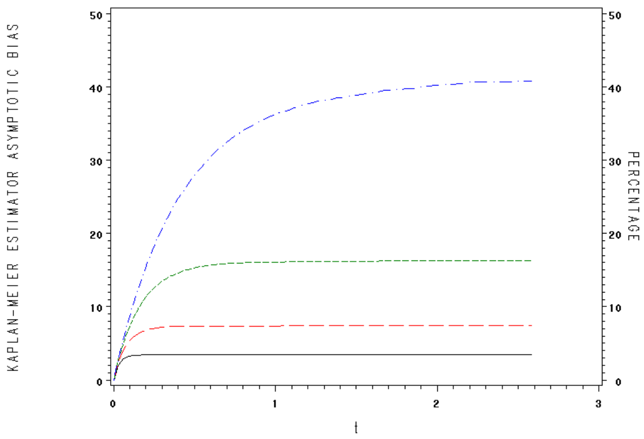
When the doubly censored data model is considered, our product-limit estimators represent lower and upper bounds for Turnbull's estimator. These bounds may serve to improve the numerical algorithms used to compute Turnbull's estimator.

Let us investigate what happens when, given Model 1 (resp. Model 2) structure data, we eliminate left (resp. right) censored failure times from the data set in order to apply the classical Kaplan-Meier estimator for right (resp. left) censored observations. For brevity, we only analyze Model 1 where we eliminate the observations with $A = 2$. The limit of the Kaplan-Meier estimator applied to the remaining observations is the probability distribution $\widetilde{F}$ with predictable hazard measure $\widetilde{L}^-(dt) = H_0(dt)/H_{01}([t,\infty])$. Since

$$F_2^1([0,t)) - H([0,t)) = H_{01}([t,\infty]) + H_2([t,\infty]) - F_2^1([t,\infty]) \le H_{01}([t,\infty]),$$

the measure $\widetilde{L}^-$ is smaller than the measure L_T^{1-}. Therefore, $F_T^1((t,\infty]) \le \widetilde{F}((t,\infty])$, $\forall t \ge 0$. We evaluated the difference between $\widetilde{F}$ and F_T^1 by simulating a large sample from Model 1 with F_{U_1}, F_T^1 and F_{V_1} exponential distributions (the parameter of F_T^1 is equal to 1). Four different values for the parameter of F_{U_1} are used and the probabilities $P(A = 2)$ obtained are 0.05, 0.1, 0.2 and 0.4. In Figure 1 we present the curves $\{\widetilde{F}((t,\infty]) - F_T^1((t,\infty])\}/F_T^1((t,\infty])$ corresponding to the four cases that are ordered from the smallest probability $P(A = 2)$ (bottom curve) to the largest one (top curve). We learn from this experiment that the sample selection bias induced when discarding the left censored failure times increases when $P(A = 2)$ increases and it may be significative when $P(A = 2)$ is not practically negligible.

Figure 1:



6 Strong convergence.

We study the uniform strong (almost sure) convergence of F_{nT}^1 and F_{nT}^2. Since, in fact, the estimators F_{nT}^1 and F_{nT}^2 are built as explicit functionals of the empirical distribution, we deduce their asymptotic behavior, in particular the strong convergence, whatever the properties of the underlying censoring mechanism are. Hereafter, we use the following rule: the subscript n indicates the empirical version of the quantities we consider. Moreover, if μ is a nonnegative measure on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$ and f is a measurable function, $\mu(f) = \int f(t)\,\mu(dt)$.

For the strong convergence we recall a result of Rolin (2001), an extension of the strong law under right-censorship proved by Stute and Wang (1993). Let $H = \sum_{1\le r\le g} H_r$ be a probability distribution decomposed into g subdistributions. If $I \subset K = \{r : 1 \le r \le g\}$, let $H_I = \sum_{r\in I} H_r$. For $J_k \subset I_k \subset K$, $k = 1, 2$, define

$$L_k^-(dt) = \frac{H_{J_k}(dt)}{H((t,\infty]) + H_{I_k}(t)}$$

and consider the measure $G(dt) = \exp\{-L_2([0,t))\}L_1^-(dt)$.

Theorem 6.1 *If $G(f) < \infty$, then $G_n(f) \to G(f)$ a.s. and in the mean.*

The same result holds if we define the predictable reverse hazard measures

$$M_k^-(dt) = \frac{H_{J_k}(dt)}{H([0,t)) + H_{I_k}(t)}, \qquad k = 1, 2,$$

and consider the measure $G(dt) = \exp\{-M_2((t,\infty])\}M_1^-(dt)$.

Let us extend the number of hazard measures associated to Model 1 by defining

$$M_0^-(dt) = \frac{H_0(dt)}{H([0,t)) + H_{01}(t)}, \qquad M_1^-(dt) = \frac{H_1(dt)}{H([0,t)) + H_1(t)}.$$

Consider F_0^1, F_1^1 the corresponding distributions. Deduce that $M_{01} = M_0 + M_1$, where M_0, M_1 and M_{01} are the reverse hazard measures associated to M_0^-, M_1^- and M_{01}^- [see (3.1)], respectively. In view of equations (2.2) deduce $H([0,t)) = F_2^1([0,t))F_{01}^1([0,t))$ and $H_{01}(t) = F_2^1([0,t))F_{01}^1(t)$. Therefore, we can write

$$H_0(dt) = F_2^1([0,t))F_{01}^1([0,t])M_0^-(dt).$$

Consequently, in the expression of the predictable hazard measure defining F_T^1 [see (3.2)], we get rid of F_2^1 and obtain

$$L_T^{1-}(dt) = \frac{F_{01}^1([0,t])}{F_{01}^1([t,\infty])} M_0^-(dt).$$

Theorem 6.2 *If f is a nonnegative Borel measurable function on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$ such that $L_T^{1-}(f) < \infty$, then, almost surely as $n \to \infty$, $L_{nT}^{1-}(f) \to L_T^{1-}(f)$.*

Theorem 6.2 is a direct consequence of the following lemma.

Lemma 6.1 *i) If $L_T^{1-}(f\, 1\!\!1_{[0,t]}) < \infty$ and $F_{01}^1([t,\infty]) > 0$, then, almost surely*

$$L_{nT}^{1-}(f\, 1\!\!1_{[0,t]}) \to L_T^{1-}(f\, 1\!\!1_{[0,t]}), \qquad n \to \infty.$$

ii) If $L_T^{1-}(f\, 1\!\!1_{[t,\infty]}) < \infty$ and $F_2^1([0,t)) > 0$, then, almost surely as $n \to \infty$,

$$L_{nT}^{1-}(f\, 1\!\!1_{[t,\infty]}) \to L_T^{1-}(f\, 1\!\!1_{[t,\infty]}).$$

Proof. Proof *i)* First, Theorem 6.1 implies that any empirical distribution function defined by the empirical reverse hazard measures of Model 1 converges uniformly on $[0,\infty]$. Now,

$$\begin{aligned} &| L_{nT}^{1-}(f\, 1\!\!1_{[0,t]}) - \int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n01}^1([0,s])M_{n0}^-(ds) | \\ &\leq \frac{|| F_{n01}^1 - F_{01}^1 ||}{F_{n01}^1([t,\infty])} \int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n01}^1([0,s])M_{n0}^-(ds). \end{aligned}$$

The second member of the inequality tends to zero almost surely because $F_{n01}^1([t,\infty]) \to F_{01}^1([t,\infty]) > 0$ a.s. and, by Theorem 6.1 applied for $G(ds) = \exp\{-M_{01}((s,\infty])\}M_0^-(ds)$,

$$\int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n01}^1([0,s])M_{n0}^-(ds) \to L_T^{1-}(f\, 1\!\!1_{[0,t]}), \quad a.s.$$

ii) First, looking at the definition of M_{01}^{-}, a simple computation shows that

$$H_{01}([s,\infty]) \leq F_{01}^{1}([s,\infty]) \leq \frac{H_{01}([s,\infty])}{H_{01}([0,\infty])}.$$

Using definition (3.2) for the predictable hazard measure defining F_T^1, we have

$$\begin{aligned} \mid L_{nT}^{1-}(f\, 1\!\!I_{[t,\infty]}) \quad &- \quad \int_{[t,\infty]} \frac{f(s)}{F_2^1([0,s))} \frac{H_{n0}(ds)}{F_{n01}^1([s,\infty])} \mid \\ &\leq \quad \frac{\|F_{n2}^1 - F_2^1\|}{F_{n2}^1([0,t))} \int_{[t,\infty]} \frac{f(s)}{F_2^1([0,s))} \frac{H_{n0}(ds)}{F_{n01}^1([s,\infty])} \\ &\leq \quad \frac{\|F_{n2}^1 - F_2^1\|}{F_{n2}^1([0,t))} \int_{[t,\infty]} \frac{f(s)}{F_2^1([0,s))} \frac{H_{n0}(ds)}{H_{n01}([s,\infty])}. \end{aligned}$$

Now, almost surely $F_{n2}^1([0,t)) \to F_2^1([0,t))$ which is strictly positive. Since

$$\int_{[t,\infty]} \frac{f(s)}{F_2^1([0,s))} \frac{H_0(ds)}{H_{01}([s,\infty])} \leq H_{01}([0,\infty])^{-1} L_T^{1-}(f\, 1\!\!I_{[t,\infty]}) < \infty,$$

a new application of Theorem 6.1 provides the result. $\Box$

Denote by t_{0k} the left endpoint and by t_{1k} the right endpoint of the support of H_k, $k = 0, 1, 2$. We have the following corollary of Theorem 6.2. Note that the strong uniform convergence of F_{nT}^1 is obtained without any additional assumption, apart that of iid observations and condition $H_0(0) = H_1(0) = 0$.

Corollary 6.3 *a) If $L_T^{1-}([0,t_{10})) < \infty$, then, almost surely,*

$$\sup_{0 \leq t < t_{10}} \mid L_{nT}^1([0,t]) - L_T^1([0,t]) \mid \to 0$$

and $L_{nT}^1(t_{10}) \to L_T^1(t_{10})$. If $L_T^{1-}([0,t_{10})) = \infty$, then, almost surely,

$$\sup_{0 \leq s \leq t} \mid L_{nT}^1([0,s]) - L_T^1([0,s]) \mid \to 0$$

for all $t < t_{10}$ and $L_{nT}^1([0,t_{10})) \to \infty$.

b) Almost surely, $\|F_{nT}^1 - F_T^1\| = \sup_{0 \leq t \leq \infty} \mid F_{nT}^1([0,t]) - F_T^1([0,t]) \mid \to 0$.

Proof. Proof The Glivenko-Cantelli theorem provides the result in *a)* with L_T^1 and L_{nT}^1 replaced by L_T^{1-} and L_{nT}^{1-}, respectively. The similar result for the hazard measure L_{nT}^1 is obtained by taking care of the fact that $L_T^1(t_{10}) = \infty$ if $L_T^{1-}(t_{10}) = 1$. This happens if $t_{10} \geq t_{11}$, $H_0(t_{10}) > 0$ and $H_1(t_{10}) = 0$. The convergence of F_{nT}^1 is implied by the convergence of the associated hazard measure L_{nT}^1. $\Box$

The strong uniform convergence of F^2_{nT} can be obtained in a similar way. Define

$$L_0^-(dt) = \frac{H_0\,(dt)}{H([t,\infty])}, \qquad L_2^-(dt) = \frac{H_2(dt)}{H((t,\infty]) + H_1(t) + H_2(t)}$$

and consider F_0^2, F_2^2 the corresponding distribution. After some manipulations, we can get rid of F_1^2 in the definition (3.3):

$$M_T^{2-}(dt) = \frac{F_2^2([t,\infty]F_0^2([t,\infty])}{1 - F_2^2([t,\infty])F_0^2((t,\infty])} L_0^-(dt).$$

Next, apply Theorem 6.1 [see Patilea and Rolin (2001) for the details].

7 Asymptotic normality.

Let $(D[a,b], \|\cdot\|)$ be the space of càdlàg functions defined on $[a,b] \subset [0,\infty]$, endowed with the supremum norm. $BV_C[a,b] \subset D[a,b]$ is the set of càdlàg functions with total variation bounded by C. The integrals with respect to functions which are not of bounded variation have to be understood via partial integration. Finally, weak convergence is denoted by $\rightsquigarrow$ and is in the sense considered by Pollard (1984), that is $D[a,b]$ is endowed with the ball $\sigma-$field.

Given the explicit form of F^i_{nT}, $i = 1, 2$, a convenient approach for proving weak convergence is the delta method [e.g., Gill (1989), van der Vaart and Wellner (1996), section 3.9]. For proving Hadamard differentiability, the denominators appearing in the maps used to define F^1_T and F^2_T should stay away from zero. Therefore, we have to complete the delta method with a tool for treating the endpoints of the intervals on which weak convergence is finally proved.

Lemma 7.2 *(Pollard (1984), page 70) Let $X, X_1, X_2...$ be random elements of $(D[a,b], \|\cdot\|)$ with the distribution of X concentrated on a separable set. Suppose, for each $\varepsilon > 0$ and $\delta > 0$, there exists approximating random elements $AX, AX_1, AX_2...$ such that $AX_n \rightsquigarrow AX$, $P(\|X - AX\| > \delta) < \varepsilon$ and*

$$\limsup_{n\to\infty} P\left(\|X_n - AX_n\| > \delta\right) < \varepsilon. \tag{7.1}$$

Then $X_n \rightsquigarrow X$.

For brevity we only consider the asymptotic normality of F^1_{nT}; similar arguments apply for F^2_{nT} [see Patilea and Rolin (2001) for the details].

Note that the empirical central limit theorem yields

$$\sqrt{n}\,(H_n - H,\, H_{0n} - H_0,\, H_{2n} - H_2) \rightsquigarrow (G,\, G_0,\, G_2) \qquad in \quad D^3([0,\infty]).$$

Now, we prove that $\sqrt{n}(M^-_{n2} - M^-_2)$ and $\sqrt{n}(F^1_{n2} - F^1_2)$ converge weakly to Gaussian limits. The computation of the covariance structures for the limit processes in this section is elementary, albeit tedious [see Patilea and Rolin (2001) for some formulae].

Lemma 7.3 *Let $M^-_{2t} = M^-_2((t,\infty])$ and M^-_{n2t} be the corresponding estimator. Assume that*

$$\int_{(t_{00},\infty]} \frac{M^-_2(du)}{H([0,u])} = \int_{(t_{00},\infty]} \frac{H_2(du)}{H([0,u])^2} < \infty, \tag{7.2}$$

where $t_{00} = \inf\{t : H_0([0,t]) > 0\}$. Then,

$$\sqrt{n}\left(H_n - H,\, H_{n0} - H_0,\, M^-_{n2} - M^-_2\right) \rightsquigarrow (G,\, G_0,\, G_M) \;\; in \; D^3[t_{00},\infty], \tag{7.3}$$

where $(G,\, G_0,\, G_M)$ is a zero-mean Gaussian process with

$$G_{Mt} = \int_{(t,\infty]} \frac{dG_{2u}}{H\left([0,u]\right)} - \int_{(t,\infty]} \frac{G_u}{H\left([0,u]\right)^2} H_2(du). \tag{7.4}$$

Moreover, if $F^1_{2t} = F^1_2([0,t])$, then

$$\sqrt{n}\left(H_n - H,\, H_{n0} - H_0,\, F^1_{n2} - F^1_2\right) \rightsquigarrow (G,\, G_0,\, G_3) \qquad in \quad D^3[t_{00},\infty],$$

where $(G,\, G_0,\, G_3)$ is a zero-mean Gaussian process with

$$G_{3t} = -F^1_2\left([0,t]\right) \int_{(t,\infty]} \frac{dG_{Mu}}{1 - M^-_2(u)}\,. \tag{7.5}$$

Proof. Proof The map $(A,B) \to \int_{(\cdot,\infty]} (1/A) dB$ is Hadamard-differentiable on a domain of the type $\{(A,B) : A \in D[a,b],\, B \in BV_C[a,b],\, A \geq \epsilon\}$, $C,\epsilon > 0$, at every point such that $1/A$ is of bounded variation. The derivative map is given by $(\alpha,\beta) \to \int_{(\cdot,\infty]} (1/A) d\beta - \int_{(\cdot,\infty]} (\alpha/A^2) dB$. Therefore, the delta-method for the map $(H, H_0, H_2) \to (H, H_0, M^-_2)$ yields the weak convergence of $\sqrt{n}(H_n - H,\, H_{n0} - H_0,\, M^-_{n2} - M^-_2)$ in $D^3[\sigma,\infty]$, provided that $H([0,\sigma]) > 0$.

For the weak convergence in $D^3[t_{00},\infty]$, consider the pathwise limit of $G_{M\sigma}$ as $\sigma \downarrow t_{00}$, which exists in view of (7.2). It remains to verify (7.1) when $H\left([0,t_{00}]\right) = 0$. It suffices to prove: *a)* for any $\varepsilon, \delta > 0$, there exists $\sigma = \sigma(\varepsilon,\delta) > t_{00}$ such that

$$\overline{\lim}_{n\to\infty} P\,(\sup_{U \leq t \leq \sigma} \sqrt{n}\left|M^-_{n2}([t,\sigma)) - M^-_2\left([t,\sigma)\right)\right| > \delta) < \varepsilon; \tag{7.6}$$

and *b)* $\sqrt{n}M_2^-((t_{00},U)) \to 0$, in probability, where $U = \min_i Y_i$. To ensure *a)*, reverse the time and apply the arguments usually used to check the "tightness at $\tau_H = \sup\{t : H([0,t]) < 1\}$" when proving weak convergence for Nelson-Aalen and Kaplan-Meier estimators [see Fleming and Harrington (1991), Theorem 6.2.1, Gill (1983)]. For *b)*, first note that (7.2) ensures $M_2^-((t_{00},\infty])$ finite. This implies $F_2^1([0,t_{00}]) > 0$ [use, for instance, arguments as in Lemma 6 of Gill and Johansen (1990)]. Since in general M^- is smaller than M, deduce

$$M_2^-((t_{00},U)) \le M_2((t_{00},U)) = \ln\frac{F_2^1([0,U))}{F_2^1([0,t_{00}])} \le \frac{F_2^1((t_{00},U))}{F_2^1([0,t_{00}])}.$$

Let $u_n^\lambda = \sup\{s : \sqrt{n}\, F_2^1((t_{00},s)) \le \lambda\}$ [see also Ying (1989)]. We have

$$\begin{aligned} P(\sqrt{n} \quad F_2^1 \quad ((t_{00},U)) > \lambda) &\le P(U > u_n^\lambda) = H((u_n^\lambda,\infty])^n \\ &\le \quad \left\{1 - F_2^1((t_{00},u_n^\lambda])F_{01}^1([0,u_n^\lambda])\right\}^n \le \left(1 - \frac{\lambda^2}{n}\frac{F_{01}^1([0,u_n^\lambda])}{F_2^1((t_{00},u_n^\lambda])}\right)^n \to 0. \end{aligned}$$

The convergence to zero is true because, in view of (7.2),

$$\frac{F_2^1((t_{00},u_n^\lambda])}{F_{01}^1([0,u_n^\lambda])} \le \int_{(t_{00},u_n^\lambda]} \frac{F_2^1(ds)}{F_{01}^1([0,s])} = \int_{(t_{00},u_n^\lambda]} \frac{M_2^-(ds)}{H([0,s])} \to 0, \qquad n \to \infty.$$

Now, *b)* is clear. For the last part of the lemma, apply the delta-method for the map $A \to \prod_{(\cdot,\infty]}(1 - A(ds))$ defined on $BV_C[t_{00},\infty]$, for some $C > 0$. $\square$

Remark. In view of the variance of the process G_3, it seems possible to relax condition (7.2) when $F_2^1([0,t_{00}]) = 0$ [see also Gill (1983)]. However, in the following, due to the lack of an obvious martingale structure for $L_{nT}^{1-} - L_T^{1-}$ it is convenient to keep the denominator appearing in the definition of L_T^{1-} away from zero when $t \downarrow t_{00}$. For this we have to impose $F_2^1([0,t_{00}]) > 0$ and in this case (7.2) is needed to bound the variance of G_{3t} when t decreases to t_{00}.

Now, we state the asymptotic normality for L_{nT}^{1-} and F_{nT}^1. The notation A_- means that we consider the left-limits of the process A.

Theorem 7.1 *Suppose condition (7.2). Let $t_{00} < \tau$ such that $H_{01}([0,\tau)) < 1$. If $L_{Tt}^{1-} = L_T^{1-}([0,t])$, then $\sqrt{n}(L_{nT}^{1-} - L_T^{1-}) \rightsquigarrow V$ in $D[0,\tau]$, where*

$$V_t = \int_{(0,t]} \frac{dG_{0u}}{(F_2^1 - H)([0,u))} - \int_{(0,t]} \frac{G_{3u-} - G_{u-}}{(F_2^1 - H)^2([0,u))} H_0(du), \qquad t \in [0,\tau],$$

is a zero-mean Gaussian process. Moreover, if $F^1_{Tt} = F^1_T([0,t])$, then we have $\sqrt{n}(F^1_{nT} - F^1_T) \rightsquigarrow W$ in $D[0,\tau]$, with W a zero-mean Gaussian process given by

$$W_t = F^1_T\left((t,\infty]\right) \int_{(0,t]} \frac{dV_u}{1 - L^{1-}_T(u)}.$$

Proof. Proof Since $F^1_{01}([\tau,\infty]) > 0$ and, by (7.2), $F^1_2([0,t_{00}]) > 0$, we have $\inf_{(t_{00},\tau]}(F^1_2 - H)([0,s)) > \epsilon$, for some $\epsilon > 0$. Thus, if $H_0(t_{00}) = 0$, the weak convergence of $\sqrt{n}(L^{1-}_{nT} - L^{1-}_T)$ is obtained by the delta-method for the map $(A,B) \to \int_{(t_{00},\cdot]}(1/A_-)dB$ [see van der Vaart and Wellner (1996), pages 382-4].

When $H_0(t_{00}) > 0$ (hence, necessarily $t_{00} > 0$), in the definition of L^{1-}_T we also have to take into account $F^1_2([0,t_{00}))$. For this extend the weak convergence in (7.3) on $D^3[0,\infty]$ by considering a modified predictable reverse hazard function

$$M^-_{2t} = M^-_2\left((t,\infty]\right) = \int_{(t,\infty]} \frac{H_2(du)}{H\left([0, u \vee t_{00}]\right)}, \qquad t \in [0,\infty].$$

Let M^-_{n2} be the empirical counterpart. Since the denominator in the last display stays away from zero, the weak convergence of $\sqrt{n}(H_n - H,\, H_{n0} - H_0,\, F^1_{n2} - F^1_2)$ in $D^3[0,\infty]$ is easily obtained by the delta-method, where now F^1_2, F^1_{n2} correspond to the modified M^-_2, M^-_{n2}, respectively. Note that now $F^1_2([0,t_{00})) > 0$. The processes G_M and G_3 are still defined according to (7.4) and (7.5), respectively. Since the modification of M^-_{2t} and M^-_{n2} do not change the definitions of L^{1-}_T and L^{1-}_{nT}, the delta-method yields the weak convergence of $\sqrt{n}\left(L^{1-}_{nT} - L^{1-}_T\right)$. The last part of the theorem is obtained by the delta-method for the product-integration map. $\square$

References

[1] Basu, A.P. and Ghosh, J.K. (1980). Identifiability of distributions under competing risks and complementary risks model. *Commun. Statist.-Theory Meth.* **A9(14)**, 1515-1525.

[2] Doss, H., Freitag, S., and Prochan, F. (1989). Estimating jointly system and component reliabilities using mutual censorship approach. *Ann. Statist.* **17**, 764-782.

[3] Fleming, T.R. and Harrington, P. (1991). *Counting Processes and Survival Analysis.* New York: John Wiley & Sons.

[4] Gill, R. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* **11**, 49-58.

[5] Gill, R. (1989). Non- and semi-parametric maximum likelihood and the von-Mises method (part I). *Scand. J. Statist.* **16**, 97-128.

[6] Gill, R. and Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. *Ann. Statist.* **18**, 1501-1555.

[7] Gu, M.G. and Zhang, C.-H. (1993). Asymptotic properties of self-consistent estimators based on doubly censored data. *Ann. Statist.* **21**, 611-624.

[8] Meeker, W.Q., and Escobar, L.A. (1998). *Statistical Methods for Reliability Data.* New York: John Wiley & Sons.

[9] Morales, D., Pardo, L. and Quesada, V. (1991). Bayesian survival estimation for incomplete data when the life distribution is proportionally related to the censoring time distrbution. *Commun. in Statist. - Theory Meth.* **20(3)**, 831-850.

[10] Patilea, V. and Rolin, J.-M. (2001). Product-limit estimators for doubly censored data. DP 0131, Institut de Statistique, Louvain-la-Neuve, http://www.stat.ucl.ac.be.

[11] Pollard, D. (1984). *Convergence of Stochastic Processes.* New-York: Springer Verlag.

[12] Rolin, J.-M. (2001). Nonparametric Competing Risks Models: Identification and Strong Consistency. DP 0115, Institut de Statistique, Louvain-la-Neuve, http://www.stat.ucl.ac.be.

[13] Stute, W. and Wang, J. (1993). The Strong Law Under Random Censorship. *Ann. Statist.* **21**, 1591-1607.

[14] Turnbull, B.W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. *J. Amer. Statist. Assoc.* **69**, 169-173.

[15] Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak convergence and empirical processes.* New-York: Springer Verlag.

[16] Ying, Z. (1989). A Note on the Asymptotic properties of the product-limit estimator on the whole line. *Statist. Prob. Letters* **7**, 311-314.