INSTITUT DE STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION PAPER

0131

PRODUCT LIMIT ESTIMATORS OF THE SURVIVAL FUNCTION FOR DOUBLY CENSORED DATA

VALENTIN PATILEA and JEAN-MARIE ROLIN

http://www.stat.ucl.ac.be

Product Limit Estimators of the Survival Function for Doubly Censored Data

Valentin Patilea * and Jean-Marie Rolin **

July 2001

Abstract

Starting from left and right censored data, two product-limit estimators of the distribution function of a lifetime variable are introduced. They correspond to two latent models with independent variables L (left censoring), T (lifetime of interest) and R (right censoring). Basically, each of our estimators is obtained by combining two Kaplan-Meier type estimators. When the observations are generated according to the classical double censoring model introduced by Turnbull (1974), the product-limit estimators represent close upper and lower bounds for Turnbull's estimator. We deduce the strong convergence of our estimators on the whole real half-line without any additional assumption. Moreover, their asymptotic normality is obtained by the delta-method under conditions concerning only the observed distribution.

Keywords: left and right censoring, hazard functions, product-limit estimator, strong convergence, delta-method

MSC 2000 Primary: 62N01 Secondary: 62G05, 62N02

---

* LEO, Université d'Orléans, Orléans, France. E-mail: Valentin.Patilea@univ-orleans.fr
** Institut de statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgium. E-mail: rolin@stat.ucl.ac.be
Financial support from the contract "Projet d'Actions de Recherche Concertées" no. 98/03-217 of the Belgian Government is gratefully acknowledged.

1 Introduction

A common problem in survival analysis is the determination of the underlying lifetime distribution in the presence of a censoring mechanism. Herein we are concerned with the nonparametric estimation of the distribution F_T of a real lifetime T, a positive (possibly infinite) random variable, when the observation of this lifetime may be doubly censored, that is either from the left or from the right. The observation consists therefore of an observed lifetime Y, a finite positive random variable, and of a discrete random variable A with values in $K = \{0, 1, 2\}$. Typically, by definition, when $A = 0$, Y is equal to T, when $A = 1$, T is greater than Y and, when $A = 2$, T is smaller than Y. The joint distribution of the random vector (Y, A), denoted by $\tilde{H}$, is then a probability measure on the measurable space $(S, \mathcal{S}) = (I\!\!R^+ \times K, \mathcal{B}^+ \bigotimes 2^K)$ where $\mathcal{B}^+$ denotes the Borel subsets of $I\!\!R^+ = [0, \infty)$ and 2^K is the collection of all subsets of K.

If we observe an i.i.d. sample of size n of the random vector (Y, A), that is $(Y_1^n, A_1^n) = \{(Y_i, A_i) : 1 \leq i \leq n\}$, the nonparametric estimation of $\tilde{H}$ is straightforward. Indeed, the nonparametric maximum likelihood estimator (NPMLE) of $\tilde{H}$, in the sense of Kiefer and Wolfowitz (1956), is given by $\tilde{H}_n$, the empirical distribution of the observed sample. Using the plug-in (or substitution) principle, the nonparametric estimation of F_T will also be straightforward as soon as we can express it, explicitly or not, as a function of $\tilde{H}$. The existence of such a function requires a precise description of the censoring mechanism that is generally achieved by introducing 'latent' variables and by making some reasonable and mild assumptions on their distributions. The most popular model (hereafter called the classical one), introduced by Turnbull (1974), provides the identification of F_T by $\tilde{H}$, but does not provide an explicit function expressing this relation. In this paper, two slightly different latent models with more stringent assumptions are shown to provide explicit inversion formulae, that is explicit functions relating F_T to $\tilde{H}$. This entails generalizations of the well-known Kaplan-Meier estimator and facilitates the study of the distributional and asymptotic properties of these estimators. Typically, the two underlying models we introduce consider that the three latent variables, that is the lifetime of interest and the left and right censoring times, are independent.

More generally, without any model assumption, given a distribution for the observed variables (Y, A), we can always apply our two inversion formulae. In this way we build two pseudo-true distribution functions of the lifetime of interest which are functionals of the observed distribution $\tilde{H}$. If the experiment under observation is compatible with the hypothesis of one of our latent models, the true F_T can be exactly recovered from the observed distribution. Otherwise, we can only approximate the true lifetime distribution. If the true model is the classical one, it is shown that the true F_T lies between our two pseudo-true distributions. In particular, Turnbull estimator of the survival function obtained by self-consistency procedures deduced from the classical model lies between the two obtained estimators. Empirical studies show that these estimates are very close to Turnbull estimate. The same appears to be true for their limits.

The paper is organized as follows. In section 2 the equations characterizing the classical model are recalled and two independent models are introduced. We obtain the inversion formulae by using suitably defined hazard measures. Next, we show that the independence assumption can be relaxed and thus the inversion formulae can still recover the true lifetime distribution under slightly weaker hypotheses. The identification problem inside our latent models is also analyzed. At the end of section 2 some comparisons with Turnbull's model are presented. Applying the inversion formulae deduced previously to the empirical distribution, we provide in section 3 the expressions of two product-limit estimators of F_T. In section 4 some empirical studies using real and simulated data are presented. Finally, in section 5 the asymptotic properties our the two product-limit estimators are studied. We deduce the strong convergence of our product-limit estimators on the whole real half-line from a generalization of the arguments of Stute and Wang (1993) due to Rolin (2001). The asymptotic normality is obtained through the delta-method [see Gill (1989)] completed with martingale arguments as those used by Gill (1983) and Ying (1989). The conditions we use for the weak convergence of the product-limit estimators concern only the observed distribution $\tilde{H}$, without any reference to a latent model.

2 Latent Variables Models

Consider $(\Omega, \mathcal{A}, P)$ an arbitrary probability space on which the random variables are defined.

2.1 The Classical Model

Let L and R be positive random variables where L represents the left censoring time and R the right censoring time and therefore L is supposed to not exceed R. Now, the observed lifetime Y may be expressed as

$$Y = \max(\min(T, R),\ L) \tag{2.1}$$

or equivalently

$$Y = \min(\max(T, L),\ R). \tag{2.2}$$

This implies that $A = 0$ when T lies between L and R, that $A = 1$ when T is greater than R and that $A = 2$ when T is less than L. If the three random variables T, L and R are supposed to be continuous, the model is then perfectly defined. However, when these variables are not continuous or when applying the model to real data, the inequalities (strict or not) must be more precisely defined leading to the notions of what has been called in the literature exclusive or inclusive censoring [see, e.g., Ferguson and Phadia (1979)]. Here, the following model is considered [see Turnbull (1974) and Gu and Zhang (1993); for a slightly different version of this model see Chang and Yang (1987)]:

$$A = \begin{cases} 0 & \text{if} \quad L < T \leq R \qquad (\textit{no censoring}) \\ 1 & \text{if} \quad (L \leq)\, R < T \qquad (\textit{right censoring}) \\ 2 & \text{if} \quad T \leq L\, (\leq R) \qquad (\textit{left censoring})\ . \end{cases} \tag{2.3}$$

Concerning the joint distribution of (T, L, R), it is assumed that T is independent of the vector (L, R), i.e.,

$$T \perp\!\!\!\perp (L, R),$$

and that $L \leq R$. If we consider the subdistributions of Y defined as

$$H_k(B) \doteq \tilde{H}(B \times \{k\}) = P\,[Y \in B,\ A = k],$$

for $k \in K$ and $B \in \mathcal{B}^+$, and the distribution of Y denoted as

$$H(B) \doteq \tilde{H}(B \times K) = P\,[Y \in B],$$

for $B \in \mathcal{B}^+$, we clearly have $H = H_0 + H_1 + H_2$.

In this classical model, the subdistributions of Y may be expressed in terms of the (marginal) distributions of the latent variables as follows:

$$\begin{aligned} H_0(dt) &= \{F_L([0,t)) - F_R([0,t))\}\ F_T(dt) \\ H_1(dt) &= F_T((t,\infty])\ F_R(dt) \\ H_2(dt) &= F_T([0,t])\ F_L(dt). \end{aligned} \tag{2.4}$$

Moreover,

$$H([0,t]) = F_L([0,t])F_T([0,t]) + F_R([0,t])F_T((t,\infty]). \tag{2.5}$$

This model has been analyzed, amongst others, by Turnbull (1974), Chang and Yang (1987), Chang (1990) and Gu and Zhang (1993). In his pioneering paper, Turnbull (1974) proposed a nonparametric maximum likelihood estimator defined as the implicit solution of the equations (2.4) and provided a self-consistent algorithm for computing it. Chang and Yang (1987) proved the identifiability of the model and the convergence of Turnbull's estimator. Chang (1990) obtained its asymptotic normality under some additional assumptions. Gu and Zhang (1993) substantially weakened conditions of continuity of the latent random variables and proved identifiability, convergence, asymptotic normality and asymptotic efficiency. A closely related model for doubly censored data was introduced by Huang (1999).

The main difficulty in obtaining all these asymptotic results relies on the fact that Turnbull's estimator was defined as an implicit solution of an estimating equation. Samuelsen (1989) overcomes this problem by assuming that the left censoring times are always observed and derived asymptotic results under this additional condition. Van der Laan and Gill (1999) introduce a sieve-NPMLE. This sieve-NPMLE still requires an iterative computational procedure, but allows easier derivations of the asymptotic results. In particular, Van der Laan and Gill prove the asymptotic efficiency of their estimator.

It is important to realize that the obtained identification results only concern the marginal distributions of the latent random variables. The joint distribution of L and R cannot be identified since these two random variables are never simultaneously observed. Therefore, untestable supplementary assumptions can be made on the joint distribution of L and R. It will be proved in the sequel that the assumption of independence between these two random variables entails some interesting results. Clearly, the assumption that $L \leq R$ can no longer hold and the model has to describe the value of Y when $R < L$. Moreover in such a case the first definition of Y (formula (2.1)) is no longer equivalent to the second definition (formula (2.2)). Indeed, when $R < L$, Y is equal to L for the first definition and equal to R for the second definition. Therefore two models stipulating the independence of L and R can be considered. It will be shown that these two models are invertible in the sense that the distribution of T is an explicit function of the distribution of the observation. This will provide Kaplan-Meier type estimators close to the Turnbull estimator. Statistically, these models *cannot* be rejected while they can be rejected by the description of the experiment. However, certain experiments easily support such assumptions [see the end of the following subsection for some examples].

2.2 Independent Models

In this section we consider two independent models, i.e., models in which two latent independent random variables independent of T are introduced to describe the observation. Many other versions of the independent models below can be built by changing the inequalities in the definition of the function A (strict with not strict and viceversa). Nevertheless, the versions we consider are the only ones for which the definition of $A = 0$ is the same as in Turnbull (1974) and for which we can derive inversion formulae.

2.2.1 Model 1

Let L_1 and R_1 be two positive random variables such that

$$T \perp\!\!\!\perp L_1 \perp\!\!\!\perp R_1,$$

$$Y \;=\; \max(\min(T\,,\,R_1\,)\,,\,L_1\,)\,,$$

and A is defined as follows :

$$A = \begin{cases} 0 & \text{if} \quad L_1 < T \leq R_1 \\ 1 & \text{if} \quad L_1 < R_1 < T \\ 2 & \text{if} \quad T \leq L_1 < R_1 \quad \text{or} \quad R_1 \leq L_1. \end{cases}$$

Note that, under the supplementary condition $L_1 < R_1$, Model 1 is equivalent to the classical model.

In this model, the subdistributions of Y may be expressed in terms of the distributions of the latent variables as follows :

$$\begin{aligned} H_0(dt) &= F_{L_1}([0,t))\; F_{R_1}([t,\infty])\; F_T(dt) \\ H_1(dt) &= F_{L_1}([0,t))\; F_T((t,\infty])\; F_{R_1}(dt) \\ H_2(dt) &= \{1 - F_T((t,\infty])\; F_{R_1}((t,\infty])\}\; F_{L_1}(dt). \end{aligned} \tag{2.6}$$

Moreover,

$$H([0,t]) = F_{L_1}([0,t])\;\{1 - F_T((t,\infty])\; F_{R_1}((t,\infty])\}. \tag{2.7}$$

These relations may be simplified by defining the random variable

$$S_1 \doteq \min(T, R_1)$$

and considering

$$H_{01} \doteq H_0 + H_1.$$

Indeed,

$$\begin{aligned} H_{01}(dt) &= F_{L_1}([0,t))\; F_{S_1}(dt) \\ H_2(dt) &= F_{S_1}([0,t])\; F_{L_1}(dt) \\ H([0,t]) &= F_{L_1}([0,t])\; F_{S_1}([0,t]). \end{aligned} \tag{2.8}$$

These relations show 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. Firstly, the determination of the distributions of L_1 and S_1 and secondly, the distributions of R_1 and T. Before solving this problem, let us consider the second model.

2.2.2 Model 2

Let L_2 and R_2 be two positive random variables such that

$$T \perp\!\!\!\perp L_2 \perp\!\!\!\perp R_2 \,,$$

$$Y = \min(\max(T,\, L_2),\, R_2),$$

and A is defined as follows :

$$A = \begin{cases} 0 & \text{if} \quad L_2 < T \leq R_2 \\ 1 & \text{if} \quad L_2 \leq R_2 < T \quad \text{or} \quad R_2 < L_2 \\ 2 & \text{if} \quad T \leq L_2 \leq R_2 . \end{cases}$$

Note that, under the supplementary condition $L_2 \leq R_2$, Model 2 is equivalent to the classical model.

Now, in this second independent model, the subdistributions of Y may be expressed in terms of the distributions of the latent variables as follows :

$$\begin{aligned} H_0(dt) &= F_{L_2}([0,t))\, F_{R_2}([t,\infty])\, F_T(dt) \\ H_1(dt) &= \{1 - F_T([0,t])\, F_{L_2}([0,t])\}\, F_{R_2}(dt) \\ H_2(dt) &= F_T([0,t])\, F_{R_2}([t,\infty])\, F_{L_2}(dt). \end{aligned} \tag{2.9}$$

Moreover,

$$H(\,(t,\infty]\,) \;=\; F_{R_2}(\,(t,\infty]\,)\,\{1 \;-\; F_T(\,[0,t]\,)\,F_{L_2}(\,[0,t]\,)\}\,. \tag{2.10}$$

These relations may be simplified by defining the random variable

$$S_2 \doteq \max(T, L_2)$$

and considering

$$H_{02} \doteq H_0 + H_2.$$

Indeed,

$$\begin{aligned} H_{02}(dt) &= F_{R_2}([t,\infty])\, F_{S_2}(dt) \\ H_1(dt) &= F_{S_2}((t,\infty])\, F_{R_2}(dt) \\ H((t,\infty]) &= F_{R_2}((t,\infty])\, F_{S_2}((t,\infty]). \end{aligned} \tag{2.11}$$

As in Model 1, these relations show that the problem of inverting the model can be solved in two steps. Firstly, the determination of the distributions of R_2 and S_2 and secondly, the distributions of L_2 and T.

These two models may receive some interpretation. The first one is a natural extension of the usual right-censoring model in the sense that the failure may occur at an unknown time before the experiment starts. Moreover, the failure cause is unknown for the experimenter [see Fernandez, *et al.* (1997) for an example of application when studying a certain cause death for the trees in a farm; see also Gu and Zhang (1993), page 615, for a closely related example]. We argue that in many applications the individuals for which the failure time is previous to the start of the experiment and the failure cause is unknown are excluded from the sample and thus an asymptotic bias is induced. Our first model may be a suitable alternative in such cases. The second model we propose is an extension of a left-censoring model where only the maximum of two independent lifetimes is observed. Herein, an additional independent right-censoring may appear, that is the maximum is known to be greater than the right-censoring time, but it is not known which of the two lifetimes will achieve this maximum.

2.3 Associated Hazard Measures

Relations (2.6), (2.8), (2.9) and (2.11) suggest how to define hazard measures in order to recover the distributions of the latent variables [see Appendix A.1 for a review of the definitions of hazard measures].

Let us define the support of the (sub)distributions of Y. In general, if μ is a positive finite measure on $(\overline{\mathbb{R}}^+, \overline{\mathcal{B}}^+)$, where $\overline{\mathbb{R}}^+ = [0, \infty]$ and $\overline{\mathcal{B}}^+$ denotes the corresponding Borel subsets, its support is defined as

$$supp(\mu) \doteq \{t \,:\, \mu([0,\, t])\, \mu([t, \infty]) > 0\}.$$

In particular, for the subdistributions of Y let

$$\begin{aligned} t_{0k} &\doteq \inf\{t \,:\, H_k([0,t]) > 0\} \\ t_{1k} &\doteq \inf\{t \,:\, H_k((t,\infty]) = 0\} \end{aligned}$$

and for the distribution of Y set

$$\begin{aligned} t_0 &\doteq \inf\{t \,:\, H([0,t]) > 0\} = \min(t_{00},\, t_{01},\, t_{02}) \\ t_1 &\doteq \inf\{t \,:\, H((t,\infty]) = 0\} = \max(t_{10},\, t_{11},\, t_{12}). \end{aligned}$$

The support of H_k, is then equal to the interval with endpoints t_{0k} and t_{1k} open (resp. closed) at an endpoint if H_k has no mass (resp. a positive mass) at this endpoint. The same rule applies for the support of H.

In this paper, for any measure μ, $\mu(t)$ is a short notation for $\mu(\{t\})$.

2.3.1 Hazard Measures associated to Model 1

Since the first two equations in (2.6) can be satisfied only if $H_0(0) = H_1(0) = 0$, this condition will be implicit in the sequel when using equations (2.6).

In view of relation (2.8), let us define the following predictable reverse hazard measures:

$$M^-(dt) \doteq \frac{H(dt)}{H([0,t])},$$

$$M_2^-(dt) \doteq \frac{H_2(dt)}{H([0,t])} \tag{2.12}$$

$$M_{01}^-(dt) \doteq \frac{H_{01}(dt)}{H([0,t)) + H_{01}(t)} \tag{2.13}$$

Now, considering the usual reverse hazard measures, we obtain

$$H([0,t]) = \exp\{-M((t,\infty])\}$$

and

$$M \;=\; M_2 \;+\; M_{01}\,.$$

If we define

$$\begin{aligned} F_2^1([0,t]) &\doteq \exp\{-M_2((t,\infty])\} \\ F_{01}^1([0,t]) &\doteq \exp\{-M_{01}((t,\infty])\}, \end{aligned} \tag{2.14}$$

then

$$H([0,t]) = F_2^1([0,t])\; F_{01}^1([0,t]). \tag{2.15}$$

If relations (2.8) hold, it follows that

$$M_2^-(dt) = M_{L_1}^-(dt) \tag{2.16}$$

for all t in the support of H_2 and that for t in the support of H_{01},

$$M_{01}^-(dt) = M_{S_1}^-(dt).$$

In the second step of the inversion, in view of relation (2.6), we define the following predictable hazard measure

$$L_T^{1-}(dt) \doteq \frac{H_0(dt)}{F_2^1([0,t))\, F_{01}^1([t,\infty])} \tag{2.17}$$

and its associated distribution

$$F_T^1((t,\infty]) \doteq \exp\{-L_T^1([0,t])\}. \tag{2.18}$$

Now, if relations (2.6) hold, in section 2.5, it will be shown in a more general setting that, for all t in the support of H_0,

$$L_T^{1-}(dt) = L_T^-(dt).$$

2.3.2 Hazard Measures associated to Model 2

In view of the first equation in (2.9), hereafter, when using this set of equations, we implicitly assume $H_0(0) = 0$.

Looking at the relation (2.11), we define the following predictable hazard measures:

$$L^-(dt) \doteq \frac{H(dt)}{H([t,\infty])}$$

$$L_{02}^-(dt) \doteq \frac{H_{02}(dt)}{H([t,\infty])} \tag{2.19}$$

$$L_1^-(dt) \doteq \frac{H_1(dt)}{H((t,\infty]) + H_1(t)}. \tag{2.20}$$

Now, considering the usual hazard measures, we obtain

$$H((t,\infty]) = \exp\{-L([0,t])\}$$

and

$$L = L_1 + L_{02}.$$

Like in the first independent model, if we define

$$\begin{aligned} F_1^2((t,\infty]) &\doteq \exp\{-L_1([0,t])\} \\ F_{02}^2((t,\infty]) &\doteq \exp\{-L_{02}([0,t])\}, \end{aligned} \tag{2.21}$$

then

$$H((t,\infty]) = F_1^2((t,\infty])\, F_{02}^2((t,\infty]). \tag{2.22}$$

If relations (2.11) hold, it follows that

$$L_1^-(dt) = L_{R_2}^-(dt) \tag{2.23}$$

for all t in the support of H_1 and that for t in the support of H_{02},

$$L_{02}^-(dt) = L_{S_2}^-(dt).$$

In the second step of the inversion, in view of relations (2.9), we define the following predictable reverse hazard measure

$$M_T^{2-}(\,dt\,) \doteq \frac{H_0\,(\,dt\,)}{F_1^2\,(\,[t,\infty]\,)\,F_{02}^2\,(\,[0,t)\,)\,+\,H_0\,(t)}\,. \tag{2.24}$$

and its associated distribution

$$F_T^2\,(\,[0,t]\,) \doteq \exp\{-M_T^2\,(\,(t,\infty]\,)\} \tag{2.25}$$

Now, if relations (2.9) hold, in section 2.5, it will be shown in a more general setting that, for all t in the support of H_0,

$$M_T^{2-}\,(\,dt\,) \;=\; M_T^-\,(\,dt\,)\,.$$

2.4 Weakening of the Independence Assumptions

As pointed out before, the independence between L_i and R_i in Model i ($i = 1, 2$) is an untestable assumption. Without this assumption, it may happen that the relations determining F_T are still valid.

In Model 1, if we only assume that, as in the classical model,

$$T \perp\!\!\!\perp (L_1, R_1),$$

the subdistributions of Y may be expressed in terms of the joint distribution of the latent variables as follows:

$$\begin{aligned} H_0(dt) &= P\,[L_1 < t \le R_1]\,F_T(dt) \\ H_1(dt) &= P\,[L_1 < t \mid R_1 = t]\,F_T((t,\infty])\,F_{R_1}(dt) \\ H_2(dt) &= \{1 - F_T((t,\infty])\,P\,[R_1 > t \mid L_1 = t]\}\,F_{L_1}(dt). \end{aligned}$$

Moreover,

$$H([0,t]) = F_{L_1}([0,t]) - F_T((t,\infty])\,P\,[L_1 \le t < R_1].$$

This entails the following proposition whose proof is provided in Appendix A.2.

Proposition 2.1 *Let* $I_1 = \{t : P[L_1 < t \leq R_1] > 0\}$*. If*

$$P[R_1 \leq t \mid L_1 = t] = P[R_1 \leq t \mid L_1 \leq t] \qquad a.s.\ F_{L_1},$$

then

$$L_T^{1-}(dt) = \begin{cases} L_T^-(dt) & \forall t \in I_1 \\ 0 & \forall t \in I_1^c. \end{cases}$$

A simple example of a model fulfilling these assumptions is the following: let X_1 and X_2 be two independent random variables and define $L_1 = X_1$ and $R_1 = \max(X_1, X_2)$. Then the model (T, L_1, R_1) is observationally equivalent to the model (T, X_1, X_2). Indeed, $\{L_1 < R_1\} = \{X_1 < X_2\}$. Note that, since $P[L_1 = R_1] > 0$, this simple model does not coincide with Turnbull's model.

In Model 2, if we only assume that

$$T \perp\!\!\!\perp (L_2, R_2),$$

the subdistributions of Y may be expressed in terms of the joint distribution of the latent variables as follows:

$$\begin{aligned} H_0(dt) &= P[L_2 < t \leq R_2]\, F_T(dt) \\ H_1(dt) &= \{1 - F_T([0,t])\, P[L_2 \leq t \mid R_2 = t]\}\, F_{R_2}(dt) \\ H_2(dt) &= P[R_2 \geq t \mid L_2 = t]\, F_T([0,t])\, F_{L_2}(dt). \end{aligned}$$

Moreover,

$$H((t,\infty]) = F_{R_2}((t,\infty]) - F_T([0,t])\, P[L_2 \leq t < R_2].$$

This entails the following proposition whose proof is given in Appendix A.2.

Proposition 2.2 *Let* $I_2 = \{t : P[L_2 < t \leq R_2] > 0\}$*. If*

$$P[L_2 \leq t \mid R_2 = t] = P[L_2 \leq t \mid R_2 > t] \qquad a.s.\ F_{R_2},$$

then

$$M_T^{2-}(dt) = \begin{cases} M_T^-(dt) & \forall t \in I_2 \\ 0 & \forall t \in I_2^c. \end{cases}$$

Recall that Model 2 under the supplementary condition that $L_2 \leq R_2$, is equivalent to the classical model. Therefore, as a corollary of the above proposition, it is seen that the classical model is equivalent to an independent model if

$$P[L_2 > t \mid R_2 > t] = 0 \qquad a.s.\ F_{R_2}.$$

This requires that the (essentially) maximal value of L_2 is smaller or equal than the (essentially) minimal value of R_2. This is quite natural because $L_2 \leq R_2$ and $L_2 \perp\!\!\!\perp R_2$ require such a condition.

Finally, note that, as in the classical model [see (2.4)], in each Model i,

$$P[L_i < t \leq R_i] = P[A = 0 \mid T = t], \qquad a.s.\ F_T, \qquad i = 1, 2.$$

2.5 Identification in the Independent Models

In view of the conclusions of Propositions 2.1 and 2.2, some parts of the distribution of interest F_T are identified in each of the two models. In order to obtain the full identification, assumptions have to be made on the support of the latent variables. In fact, full identification will be obtained if and only if the support of F_T is included in the support of H_0. Indeed, it may be seen, in definitions (2.17) and (2.24), that the supports of the corresponding hazard and reverse hazard measures are the same as the supports of the numerators. In particular, the support of L_T^{1-} and M_T^{2-} is equal to the support of H_0.

To make these statements more precise, we define for any random variable U

$$\begin{aligned} s_{0U} &\doteq \inf\{\, t : F_U(\,[0,t]\,) > 0 \,\} \\ s_{1U} &\doteq \inf\{\, t : F_U(\,[0,t]\,) = 1 \,\}. \end{aligned}$$

The support of F_U is then the interval with endpoints s_{0U} and s_{1U}, open (resp. closed) at an endpoint if F_U has no mass (resp. a positive mass) at this endpoint. Clearly, the sets I_i, defined in Propositions 2.1 and 2.2 become

$$I_i = \{t : \ F_{L_i}([0,t))\ F_{R_i}([t,\infty]) > 0\}, \qquad i = 1, 2.$$

Using the above definitions, we obtain

$$I_i = \begin{cases} (s_{0L_i}, s_{1R_i}) & \text{if} \quad F_{R_i}(s_{1R_i}) = 0 \\ \\ (s_{0L_i}, s_{1R_i}] & \text{if} \quad F_{R_i}(s_{1R_i}) > 0. \end{cases} \tag{2.26}$$

Therefore, in Model i, F_T will be identified, that is $F_T^i = F_T$ on $[0,\infty]$, if and only if

$$supp(F_T) \subset I_i.$$

Indeed, since $supp(H_0) = I_i \cap supp(F_T)$, this implies that L_T^{1-} (resp. M_T^{2-}) is equal to L_T^- (resp. M_T^-).

In conclusion, the identification of F_T in any of the two models will require the following two conditions.

Left endpoint condition: either $s_{0L_i} < s_{0T}$ or $s_{0L_i} = s_{0T}$ and $F_T(s_{0T}) = 0$.

Right endpoint condition: either $s_{1R_i} > s_{1T}$ or $s_{1R_i} = s_{1T}$ and, either $F_{R_i}(s_{1R_i}) > 0$ or $F_{R_i}(s_{1R_i}) = 0$ and $F_T(s_{1T}) = 0$.

Consequently, if the distribution of T satisfies $F_T(s_{0T}) = 0$ and $F_T(s_{1T}) = 0$, F_T will be identified as soon as $s_{0L_i} \leq s_{0T}$ and $s_{1R_i} \geq s_{1T}$. Moreover, it may be seen from Appendix A.1 that, in general, the mass of L^- (resp. M^-) at infinity (resp. at zero) does not play any role in the definition of the probability distribution. Therefore, in Model 1, F_T will be identified if

$$L_T^{1-} = L_T^- \qquad \text{on} \qquad [0, \infty).$$

This implies that if $s_{1R_1} = \infty$ only the condition on the left endpoint has to hold. Similarly, in Model 2, F_T will be identified if

$$M_T^{2-} = M_T^- \qquad \text{on} \qquad (0, \infty].$$

This implies that if $s_{0L_2} = 0$ only the condition on the right endpoint has to be satisfied. At last, note that a mass of F_T at zero cannot be identified in Model 1 and a mass of F_T at infinity will be identified in Model 2 only if $s_{1R_2} = \infty$ and if $F_{R_2}(\infty) > 0$.

The validity of all these identification conditions can also be easily deduced from the expressions of F_T provided in Appendix A.3.

2.6 Comparison with the Classical Model

In section 2.3, F_T^1 and F_T^2 have been expressed in terms of the subdistributions of Y. Therefore, they are defined in *any* doubly censored model. It is then interesting to look at possible interpretations of these two distributions in the classical model.

In Model 1, using definition (2.17) and relation (2.15), the predictable hazard measure of F_T^1 may be rewritten as

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

while in the classical model (relations (2.4) and (2.5)), it may be seen that, on the support of H_0,

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

Now, the predictable reverse hazard measure of F_2^1 (formula (2.13)) may be written in the classical model as

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

It is then clear that the measure M_2^- is smaller than the measure M_L^-. The same is true for the usual reverse hazard measures and therefore

$$F_2^1([0,t)) \geq F_L([0,t)), \qquad \forall\, t \geq 0.$$

Hence, the measure $L_T^{1\,-}$ is smaller than the measure L_T^- and the measure L_T^1 is then smaller than the measure L_T. Therefore,

$$F_T^1([0,t]) \leq F_T([0,t]), \qquad \forall t \geq 0.$$

In Model 2, using definition (2.24) and relation (2.22), the unpredictable reverse hazard measure of F_T^2 may be rewritten as

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

while in the classical model (relations (2.4) and (2.5)), it may be seen that, on the support of H_0,

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

Now the unpredictable hazard measure of F_1^2 (formula (2.20)) may be written in the classical model as

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

It is then clear that the measure L_1^+ is smaller than the measure L_R^+. The same is true for the usual hazard measures and therefore

$$F_1^2([t,\infty]) \geq F_R([t,\infty]) \qquad \forall\, t \geq 0.$$

Hence, the measure $M_T^{2\,+}$ is smaller than the measure M_T^+ and the measure M_T^2 is then smaller than the measure M_T. Therefore,

$$F_T^2([0,t]) \geq F_T([0,t]), \qquad \forall\, t \geq 0.$$

3 The Product-Limit Estimators

To summarize the preceding sections, we have two distributions that are functionals of the distribution of the observations $\tilde{H}$ and that are closely related (equal under the identification conditions) to the distribution of the lifetime F_T in some independent (or equivalent to some independent) latent variables models.

Now, by the plug-in principle, if we replace in the expressions of F_T^1 and F_T^2 , $\tilde{H}$ by the empirical distribution of the observations $\tilde{H}_n$, we obtain generalizations of the Kaplan-Meier estimator. That is, F_{nT}^1 is the NPMLE of F_T in Model 1 and F_{nT}^2 is the NPMLE of F_T in Model 2. Moreover, in the classical model, F_{nT}^1 is an estimator of F_T whose distribution function is smaller than the Turnbull NPMLE and F_{nT}^2 is an estimator of F_T whose distribution function is greater than the Turnbull NPMLE.

To obtain product-limit forms of these estimators, we define the following statistics : $\{ Z_j : 1 \leq j \leq M \}$ are the distinct values in increasing order of the sample $\{ Y_i : 1 \leq i \leq n \}$. Now,

$$D_{kj} = \sum_{1 \leq i \leq n} \mathbb{1}_{\{ Y_i = Z_j \, , \, A_i = k \}}$$

$$N_j = \sum_{1 \leq i \leq n} \mathbb{1}_{\{ Y_i \leq Z_j \}}$$

$$\overline{N}_j = \sum_{1 \leq i \leq n} \mathbb{1}_{\{ Y_i \geq Z_j \}} \, .$$

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. Thanks to these definitions, the NPMLE of F_T in Model 1 is given by

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

where

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

The NPMLE of F_T in Model 2 is given by

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

where

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

4 Examples and Simulations

In this section we carry out several estimations using the product-limit estimators introduced above for both real and simulated data. We show that when applied to real doubly-censored data, our estimators are quite close to Turnbull NPMLE. We also investigate the asymptotic bias of our estimators when the data are drawn in Turnbull's model, that is we compute the pseudo-true distribution functions corresponding to the independent models when the true model is the classical one. Again, it appears that our independent models represent close approximations for Turnbull model.

First, we apply our product-limit estimators for the data on African infant precocity used by Turnbull (1974). These data are obtained from a study aiming to establish norms for infant development for a community in Kenya [see Leiderman *et al.* (1973)]. The sample consisted of children born between July 1st and December 31, 1969. Starting in January 1970, each child was tested monthly to see if he had learned to accomplish certain tasks. The variable T represents the time from birth to first learn to perform a particular task. Left-censoring occurred when it was found that, at the very first test, some children could already perform the task; whereas right-censoring occurred when some infants were still unsuccessful by the end of the study. The data is provided in Table 1.

Table 1. Data for Infant Development

Age (year)	Z_1	Z_2	Z_3	Z_4
D_{0j}	**12**	**6**	**2**	**3**
D_{1j}	**3**	**2**	**0**	**3**
D_{2j}	**2**	**4**	**2**	**5**

The nonparametric maximum likelihood estimation for Turnbull model and for our two independent models yields the results given in Table 2.

Table 2. NPMLE Estimations for Infant Development Data

Age (year)	Z_1	Z_2	Z_3	Z_4
$F_{nT}^1([0, Z_j])$	**0.431**	**0.665**	**0.753**	**0.876**
$F_{nT}([0, Z_j])$	**0.462**	**0.705**	**0.790**	**0.905**
$F_{nT}^2([0, Z_j])$	**0.547**	**0.808**	**0.889**	**1.000**

Next, we consider a set of data on drug use by California high school students [see Hamburg *et al.* (1975); see also Wellner and Zhang (1995) where this data set is reproduced]. Each of the 191 adolescents was asked for an answer to the question "When did you first use marijuana?". Any direct answer as "12, 15 or 17" yields an exact observation of the lifetime variable T. Meanwhile, a student answering "I never used it" would correspond to an observation censored on the right at the student's current age. Similarly, if a student answered "I have used it but I cannot remember when was the first time", this gives rise to a left-censored observation by student's age at the time of the study. The data is given in Table 3.

Table 3. Data for California Students

Z_j	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
D_{0j}	**4**	**12**	**19**	**24**	**20**	**13**	**3**	**1**	**0**	**4**
D_{1j}	**0**	**0**	**2**	**15**	**24**	**18**	**14**	**6**	**0**	**0**
D_{2j}	**0**	**0**	**0**	**1**	**2**	**3**	**2**	**3**	**1**	**0**

Wellner and Zhang (1995) used Turnbull model in order to estimate the distribution function of lifetime corresponding to the first use of the drug. Our product-limit estimations and Wellner and Zhang results are presented in Table 4.

Table 4. NPMLE Estimations for California Students Data

Z_j	$F^1_{nT}([0, Z_j])$	$F_{nT}([0, Z_j])$	$F^2_{nT}([0, Z_j])$
10	**.023**	**.023**	**.028**
11	**.091**	**.094**	**.113**
12	**.199**	**.206**	**.247**
13	**.337**	**.349**	**.419**
14	**.468**	**.484**	**.578**
15	**.587**	**.608**	**.655**
16	**.631**	**.655**	**.760**
17	**.665**	**.692**	**.789**
18	**.665**	**.692**	**.789**
19	**1.00**	**1.00**	**1.00**

Finally, we generate large samples ($n = 250.000$) from a lifetime distribution F_T following a double censoring mechanism as in Turnbull model. Afterwards, we estimate the distribution function F_T using our two product estimators in order to obtain very close approximations of the pseudo-true distribution functions correponding to the independent models. Several type of distributions are considered for this simulation study: each of L, T and R can be continuous, discontinuous or purely discontinuous. We define the right-censoring variable as $R = L + D$, where $D \geq 0$ and L and D are independent. For each of the cases we retained for presentation the proportions of left-censored, uncensored and right-censored observation were, approximately, 0.3, 0.5 and 0.2, respectively.

In Figure 1 we report the results for three cases we considered; the other cases we studied produced similar patterns. The first picture corresponds to the case of F_T exponential distribution function of parameter one; L has an exponential distribution of parameter two, while D is an exponential of parameter 0.4. For the second picture we considered F_T a mixture between an exponential distribution (with parameter one) and point masses at 0.5, 1 and 1.5. The weight of the continuous part is 0.6. The distributions of censoring variables are also obtained as mixtures between some exponential distributions and some point masses distributions. The probability of the event $\{L = R\}$ is 0.06. The last picture corresponds to the case of purely discontinuous distributions for L, T and R. In this last case, the probability of the event $\{L = R\}$ is 0.08.

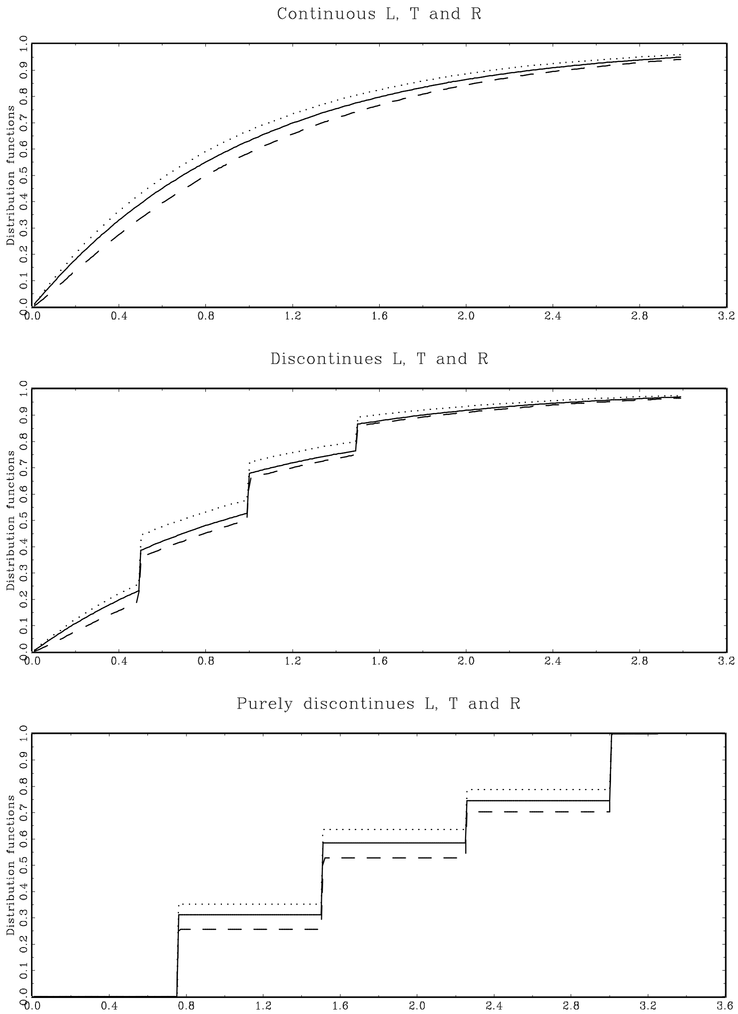

Figure 1:

Continuous line - true distribution function.
Dashed line - pseudo-true distribution function for the independent model 1.
Dotted line - pseudo-true distribution function for the independent model 2.

5 Asymptotic Theory

In this section we obtain the uniform strong (almost sure) convergence and the asymptotic normality for the two estimators introduced in section 3. As we proved in section 2.3, to any distribution $\tilde{H}$ we can associate the distributions F_T^1 and F_T^2. We shall therefore simply look at F_T^1 and F_T^2 as functionals of the observed distribution $\tilde{H}$ and proceed to a model-free study of the corresponding plug-in estimators. In this way we can deduce the asymptotic behavior of the estimators whatever the properties of the underlying censoring mechanism are.

For the strong convergence we use a general result of Rolin (2001) obtained from reverse supermartingales arguments. The asymptotic normality is obtained through the delta method [see Gill (1989); see also van der Vaart and Wellner (1996), section 3.9]. The space $(D([a,b]), \|\cdot\|)$ is the set of all càdlàg functions defined on $[a,b] \subset [0,\infty]$, endowed with the supremum norm. Weak convergence is denoted by $\rightsquigarrow$ and is to be understood in the sense considered by Pollard (1984), that is the space of càdlàg functions is endowed with the $\sigma-$field generated by the open balls.

Hereafter, we use the following rule: the subscript n indicates the empirical version of the quantities we consider.

5.1 Strong Convergence

The main tool to obtain strong convergence results in censored data is an extension of the result of Stute and Wang (1993) which can be stated as follows and whose proof may be found in Rolin (2001). Let H be a probability distribution decomposed into g subdistributions as follows:

$$H = \sum_{1 \leq r \leq g} H_r.$$

For $I \subset K = \{r : 1 \leq r \leq g\}$, let

$$H_I = \sum_{r \in I} H_r.$$

If we define for $J_k \subset I_k \subset K$ the predictable hazard measure

$$L_k^-(dt) = \frac{H_{J_k}(dt)}{H((t,\infty]) + H_{I_k}(t)}, \qquad k = 1, 2,$$

and consider the following measure

$$G(dt) = e^{-L_2([0,t))} L_1^-(dt),$$

then we have the following result :

Theorem 5.1 *If $G(f) < \infty$, then, almost surely and in the mean,*

$$G_n(f) \to G(f).$$

The same result holds, if we define the two 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

$$G(dt) = e^{-M_2((t,\infty])} M_1^-(dt).$$

5.1.1 Strong Convergence for F_{nT}^1

Let us extend the number of hazard measures associated to Model 1 by defining the following predictable reverse hazard measures :

$$\begin{aligned} M_0^-(dt) &\doteq \frac{H_0(dt)}{H([0,t)) + H_0(t)} \\ M_1^-(dt) &\doteq \frac{H_1(dt)}{H([0,t)) + H_{01}(t)} \end{aligned}$$

Then, just as before we obtain $M_{01} = M_0 + M_1$. Hence, if

$$\begin{aligned} F_0^1([0,t]) &\doteq \exp\{-M_0((t,\infty])\} \\ F_1^1([0,t]) &\doteq \exp\{-M_1((t,\infty])\} \end{aligned}$$

we can express the subdistribution H_0 as follows:

$$H_0(dt) = F_2^1([0,t))F_1^1([0,t))F_0^1([0,t])M_0^-(dt).$$

Consequently, in the expression of the predictable hazard measure defining F_T^1 [see (2.17)], we get rid of F_2^1 and we obtain

$$L_T^{1-}(dt) = \frac{F_1^1([0,t))F_0^1([0,t])}{F_{01}^1([t,\infty])} M_0^-(dt).$$

We can now state the following result:

Theorem 5.2 *If f is a positive Borel measurable function on $(\mathbb{R}^+, \mathcal{B}^+)$ such that $L_T^{1-}(f) < \infty$, then, almost surely as $n \to \infty$,*

$$L_{nT}^{1-}(f) \to L_T^{1-}(f).$$

The Glivenko-Cantelli theorem provides the following corollary :

Corollary 5.3 *If $L_T^{1-}([0,t_{10})) < \infty$, then, almost surely,*

$$|| L_{nT}^{-} - L_T^{-} || = \sup_{0\le t\le t_{10}} | L_{nT}^{1-}([0,t]) - L_T^{1-}([0,t]) | \to 0, \qquad n\to\infty.$$

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

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

for all $t < t_{10}$ and

$$L_{nT}^{1-}([0,t_{10})) \to \infty, \qquad n\to\infty.$$

The proof of Theorem 5.2 is a straightforward consequence of the following two lemmas.

Lemma 5.4 *If $L_T^{1-}(f\, \mathbb{1}_{[0,t]}) < \infty$ and $F_{01}^1([t,\infty]) > 0$, then, almost surely*

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

Proof: Note first that Theorem 5.1 implies that any empirical distribution function defined by the empirical reverse hazard measures of Model 1 converges uniformly on $[0,\infty]$. Now,

$$\begin{aligned} &| L_{nT}^{1-}(f\, \mathbb{1}_{[0,t]}) - \int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n1}^1([0,s))F_{n0}^1(ds) | \\ &\le \frac{|| F_{n01}^1 - F_{01}^1 ||}{F_{n01}^1([t,\infty])} \int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n1}^1([0,s))F_{n0}^1(ds). \end{aligned}$$

The second member of the inequality tends to zero because

$$F_{n01}^1([t,\infty]) \to F_{01}^1([t,\infty])$$

which is strictly positive and

$$\int_{(0,t]} \frac{f(s)}{F_{01}^1([s,\infty])} F_{n1}^1([0,s))F_{n0}^1(ds) \to L_T^{1-}(f\, \mathbb{1}_{[0,t]})$$

by Theorem 5.1. Indeed,

$$F_1^1([0,s))F_0^1(ds) = e^{-M_{01}((s,\infty])}\, [1 - M_1^-(s)\,]\, M_0^-(ds)$$

has the form of the measure G of Theorem 5.1. ■

Note that this Lemma suffices to prove Theorem 5.2 and Corollary 5.3 in the case $F^1_{01}([t_{10},\infty]) > 0$ or, equivalently, if $M_{01}([t_{10},\infty]) > 0$. This will be the case if either $t_{10} < t_{11}$ or $t_{10} \geq t_{11}$ and $H_{01}(t_{10}) > 0$.

Lemma 5.5 *If $L^{1-}_T(f\,1\!\!1_{[t,\infty]}) < \infty$ and $F^1_2([0,t)) > 0$, then, almost surely as $n \to \infty$,*

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

Proof: First of all, looking at the definition of M^-_{01} (formula (2.13)) and using the trivial fact that $0 \leq H_2\,([0,s)) \leq H_2([0,\infty])$, a simple computation shows that

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

Now, using the first expression of the predictable hazard measure defining F^1_T (formula (2.17)),

$$\begin{aligned}
&|\ L^{1-}_{nT}\,(\,f\,1\!\!1_{[t,\infty]}\,) - \int_{[t,\infty]} \frac{f\,(s)}{F^1_2\,(\,[0,s)\,)}\,\frac{H_{n0}\,(ds)}{F^1_{n01}\,(\,[s,\infty]\,)}\ |\\
&\leq \frac{||\ F^1_{n2} - F^1_2\ ||}{F^1_{n2}\,(\,[0,t)\,)} \int_{[t,\infty]} \frac{f\,(s)}{F^1_2\,(\,[0,s)\,)}\,\frac{H_{n0}\,(ds)}{F^1_{n01}\,(\,[s,\infty]\,)}\\
&\leq \frac{||\ F^1_{n2} - F^1_2\ ||}{F^1_{n2}\,(\,[0,t)\,)} \int_{[t,\infty]} \frac{f\,(s)}{F^1_2\,(\,[0,s)\,)}\,\frac{H_{n0}\,(ds)}{H_{n01}\,(\,[s,\infty]\,)}\,.
\end{aligned}$$

Now,

$$F^1_{n2}\,(\,[0,t)\,) \ \to\ F^1_2\,(\,[0,t)\,)$$

which is strictly positive and since

$$\int_{[t,\infty]} \frac{f\,(s)}{F^1_2\,(\,[0,s)\,)}\,\frac{H_0\,(ds)}{H_{01}\,(\,[s,\infty]\,)} \ \leq\ H_{01}\,(\,[0,\infty]\,)^{-1}\,L^{1-}_T\,(\,f\,1\!\!1_{[t,\infty]}\,)\ <\ \infty$$

a new application of Theorem 5.1 provides the result. ■

Note that this Lemma suffices to prove Theorem 5.2 and Corollary 5.3 in the case $F^1_2([0,t_{00})) > 0$ or, equivalently, if $M_2([t_{00},\infty]) < \infty$. This will be the case if either $t_{00} > t_0$ or $t_{00} = t_0$ and $H_{01}(t_0) > 0$.

From the above results one easily obtain a Glivenko-Cantelli theorem for the corresponding usual hazard measure in 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 $H_1(t_{10}) = 0$.

Corollary 5.6 *If $L_T^{1-}([0,t_{10})) < \infty$, then, almost surely,*

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

and

$$L_{nT}(t_{10}) \to L_T(t_{10}).$$

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

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

for all $t < t_{10}$ and

$$L_{nT}([0,t_{10})) \to \infty.$$

A trivial consequence is the almost sure uniform convergence on $\mathbb{R}^+$ of the NPMLE of F_T in Model 1.

Corollary 5.7 *Almost surely,*

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

5.1.2 Strong Convergence for F_{nT}^2

If we define the following predictable hazard measures associated to Model 2 :

$$\begin{aligned} L_0^-(dt) &\doteq \frac{H_0(dt)}{H([t,\infty])} \\ L_2^-(dt) &\doteq \frac{H_2(dt)}{H((t,\infty]) + H_1(t) + H_2(t)} \end{aligned}$$

just as before we obtain $L_{02} = L_0 + L_2$. Therefore, defining

$$\begin{aligned} F_0^2((t,\infty]) &\doteq \exp\{-L_0([0,t])\} \\ F_2^2((t,\infty]) &\doteq \exp\{-L_2([0,t])\}, \end{aligned}$$

we can express the subdistribution H_0 as follows:

$$H_0(dt) = F_0^2([t,\infty])F_2^2([t,\infty])F_1^2([t,\infty])L_0^-(dt).$$

In the expression of the predictable reverse hazard measure defining F_T^2 (formula (2.24)) we get rid of F_1^2 and obtain after some manipulations,

$$M_T^{2-}(dt) = \frac{F_2^2([t,\infty]\, F_0^2([t,\infty])}{1 - F_2^2([t,\infty])\, F_0^2((t,\infty])} L_0^-(dt). \tag{5.1}$$

We can now state the following result :

Theorem 5.8 *If f is a positive Borel measurable function on $(\mathbb{R}^+, \mathcal{B}^+)$ such that $M_T^{2-}(f) < \infty$, then, almost surely as $n \to \infty$,*

$$M_{nT}^{2-}(f) \;\to\; M_T^{2-}(f).$$

As in the preceding subsection, this entails Glivenko-Cantelli type corollaries for the predictable and the usual reverse hazard measures and the almost sure uniform convergence on $\mathbb{R}^+$ of the NPMLE of F_T in Model 2.

Corollary 5.9 *Almost surely as $n \to \infty$,*

$$\| F_{nT}^2 - F_T^2 \| = \sup_{0 \le t \le \infty} |\, F_{nT}^2([0,t]) - F_T^2([0,t])\, | \to 0\,.$$

Just as before, the proof of Theorem 5.8 is a straightforward consequence of the following two lemmas.

Lemma 5.10 *If $M_T^{2-}(f\, \mathbb{1}_{[t,\infty]}) < \infty$ and $F_2^2([t,\infty])F_0^2((t,\infty]) < 1$, then, almost surely as $n \to \infty$,*

$$M_{nT}^{2-}(f\, \mathbb{1}_{[t,\infty]}) \;\to\; M_T^{2-}(f\, \mathbb{1}_{[t,\infty]}).$$

Proof: Once again, Theorem 5.1 implies that any empirical distribution function defined by the empirical hazard measures of Model 2 converges uniformly on $[0,\infty]$. Now,

$$|\, M_{nT}^{2-}(f\, \mathbb{1}_{[t,\infty]}) - \int_{[t,\infty)} \frac{f(s)}{1 - F_2^2([s,\infty])\, F_0^2((s,\infty])} F_{n2}^2([s,\infty])\, F_{n0}^2(ds)\, |$$

$$\le \frac{\| F_{n2}^2 - F_2^2 \| + \| F_{n0}^2 - F_0^2 \|}{1 - F_{n2}^2([t,\infty])\, F_{n0}^2((t,\infty])}$$

$$\times \int_{[t,\infty)} \frac{f(s)}{1 - F_2^2([s,\infty])\, F_0^2((s,\infty])}\, F_{n2}^2([s,\infty])\, F_{n0}^2(ds).$$

The second member of the inequality tends to zero because

$$F_2^2([s,\infty])\, F_0^2(ds) = e^{-L_{02}([0,s))}\, L_0^-(ds)$$

has the form of the measure G of Theorem 5.1. ∎

This Lemma suffices to prove Theorem 5.8 if $F_2^2([t_{00},\infty])\, F_0^2((t_{00},\infty]) < 1$ or equivalently if $L_2([0,t_{00})) + L_0(t_{00}) > 0$. This will be the case if either $t_{00} > t_{02}$ or $t_{00} \le t_{02}$ and $H_0(t_{00}) > 0$.

Lemma 5.11 *If $M_T^{2-}(f\, 1\!\!I_{[0,t]}) < \infty$ and $F_1^2([t,\infty]) > 0$, then, almost surely as $n \to \infty$,*

$$M_{nT}^{2-}(f\, 1\!\!I_{[0,t]}) \to M_T^{2-}(f\, 1\!\!I_{[0,t]}).$$

Proof: First of all, looking at the definition of L_{02}^- (formula (2.19)) and using the trivial fact that $0 \le H_1([s,\infty]) \le H_1([0,\infty])$, a simple computation shows that

$$H_{02}([0,s)) + H_0(s) \le 1 - F_2^2([s,\infty])\, F_0^2((s,\infty]) \le \frac{H_{02}([0,s)) + H_0(s)}{H_{02}([0,\infty])}.$$

Now, combining the two expressions of the predictable reverse hazard measure defining F_T^2 (formulae (2.24) and (5.1)),

$$\begin{aligned} &\left|\, M_{nT}^{2-}(f\, 1\!\!I_{[0,t]}) - \int_{[0,t]} \frac{f(s)}{F_1^2([s,\infty])} \frac{H_{n0}(ds)}{1 - F_{n2}^2([s,\infty])\, F_{n0}^2((s,\infty])} \,\right| \\ &\le \frac{\|\, F_{n1}^2 - F_1^2 \,\|}{F_{n1}^2([t,\infty])} \int_{[0,t]} \frac{f(s)}{F_1^2([s,\infty])} \frac{H_{n0}(ds)}{1 - F_{n2}^2([s,\infty])\, F_{n0}^2((s,\infty])} \\ &\le \frac{\|\, F_{n1}^2 - F_1^2 \,\|}{F_{n1}^2([t,\infty])} \int_{[0,t]} \frac{f(s)}{F_1^2([s,\infty])} \frac{H_{n0}(ds)}{H_{n02}([0,s)) + H_{n0}(s)} \end{aligned}$$

Now, since

$$\int_{[0,t]} \frac{f(s)}{F_1^2([s,\infty])} \frac{H_0(ds)}{H_{02}([0,s)) + H_0(s)} \le H_{02}([0,\infty])^{-1}\, M_T^{2-}(f\, 1\!\!I_{[0,t]})$$

which is finite, a new application of Theorem 5.1 provides the result. ∎

Note that this Lemma suffices to prove Theorem 5.8 if $F_1^2([t_{10},\infty]) > 0$ or equivalently if $L_1([0,t_{10})) < \infty$. This will be the case if either $t_{10} < t_1$ or $t_{10} = t_1$ and $H(t_1) > 0$.

5.2 Asymptotic Normality

Given the explicit form of our estimators, a natural idea is to use the delta method combined with the empirical central limit theorem for $\tilde{H}$. For the Hadamard differentiability, the denominators appearing in the transformations used to express F_T^1 and F_T^2 as functions $\tilde{H}$ should stay away from zero. This makes that the delta method has to be combined with a special treatment at the endpoints of the interval on which the weak convergence is expected to hold. For the extension at these endpoints we make use of the following lemma [see Pollard (1984), page 70].

Lemma 5.12 *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*
i) $P(\|X - AX\| > \delta) < \varepsilon$;
ii) $\limsup_{n\to\infty} P(\|X_n - AX_n\| > \delta) < \varepsilon$;
iii) $AX_n \rightsquigarrow AX$.
Then $X_n \rightsquigarrow X$.

Note that verifying the condition *ii)* of the previous lemma (a kind of tightness condition) will be a rather delicate matter in our doubly censoring framework. This because, contrarily to the case of the classical random right-censorship model, there is no obvious (continuous time) martingale structure for the estimators we propose.

Let us introduce some more notation: $BV_C([a,b])$ is the set of all càdlàg functions $t \to A_t$ defined on an interval $[a,b] \subset [0,\infty]$ with total variation $\int |dA|$ bounded by C (equipped with the uniform norm). If E is any of the spaces $D([a,b])$ or $BV_C([a,b])$ and $k \geq 1$, then E^k denotes the product space (k times) equipped with the product norm. Let $\prod_{(t,u]}(1 + A(ds))$ denote the product-integral of the bounded variation càdlàg function A over $(t,u]$ [see Gill and Johansen (1990); see also van der Vaart and Wellner (1996, section 3.9.4.5)]. Finally, in this section the integrals with respect to functions which are not of bounded variation have to be understood via partial integration.

5.2.1 Asymptotic Normality for F_{nT}^1

Assume that $H_0(0) = H_1(0) = 0$ and recall that $t_{00} = \inf\{t : H_0([0,t]) > 0\}$. First, we prove that $\sqrt{n}(M_{n2}^- - M_2^-)$ converges weakly to a Gaussian limit and from this we deduce the same result for $\sqrt{n}(F_{n2}^1 - F_2^1)$.

Lemma 5.13 *Let*

$$M_{2t}^- = M_2^-((t,\infty]) = \int_{(t,\infty]} \frac{H_2(du)}{H\left([0,u]\right)}, \qquad t \geq 0,$$

and M_{n2t}^- the corresponding plug-in estimator. Assume that

$$\int_{(t_{00},\infty]} \frac{M_2^-(du)}{H([0,u])} = \int_{(t_{00},\infty]} \frac{H_2(du)}{H([0,u])^2} < \infty. \tag{5.2}$$

Then,

$$\sqrt{n}\left(H_n - H,\, H_{n0} - H_0,\, M_{n2}^- - M_2^-\right) \rightsquigarrow (G,\, G_0,\, G_M) \;\; in \;\; D^3([t_{00},\infty]), \tag{5.3}$$

where $(G,\, G_0,\, G_M)$ is a zero-mean Gaussian process with $G_{Mt} = A_t - B_t$ and

$$A_t = \int_{(t,\infty]} \frac{dG_{2u}}{H\left([0,u]\right)}, \qquad B_t = \int_{(t,\infty]} \frac{G_u}{H\left([0,u]\right)^2} H_2(du). \tag{5.4}$$

The covariance structure of $(G,\, G_0,\, G_M)$ is

$$E\left(G_s G_t\right) = H\left([0, s\wedge t]\right) - H\left([0,s]\right) H\left([0,t]\right),$$

$$E\left(G_{0s} G_{0t}\right) = H_0\left([0, s\wedge t]\right) - H_0\left([0,s]\right) H_0\left([0,t]\right),$$

$$E\left(G_{Ms} G_{Mt}\right) = \int_{(s\vee t,\infty]} \frac{1 - M_2^-(u)}{H([0,u])} M_2^-(du),$$

$$E\left(G_s G_{0t}\right) = H_0\left([0, s\wedge t]\right) - H\left([0,s]\right) H_0\left([0,t]\right),$$

$$E\left(G_s G_{Mt}\right) = -H([0,s]) \int_{(s\vee t,\infty]} \frac{M_2^-(du)}{H([0,u])},$$

$$E\left(G_{0s} G_{Mt}\right) = -\int_{(t,\infty]} \frac{H_0\left([0, s\wedge u]\right)}{H([0,u])} M_2^-(du),$$

where $s\wedge t = \min(s,t)$, $s \vee t = \max(s,t)$.

Proof. Let $\epsilon > 0$ and $t_{00} \leq \sigma$ such that $H\left([0,\sigma]\right) > \epsilon$. The map

$$(H, H_0, H_2) \;\rightarrow\; \left(H, H_0, M_2^-\right)$$

is Hadamard differentiable on the domain

$$\widetilde{D} = \left\{(A_1, A_2, A_3) \in BV_1^3([\sigma,\infty]) \,:\, A_3 \geq 0,\, A_1 \geq \epsilon\right\},$$

with values in $BV_C^3([\sigma,\infty])$ [cf Lemma A.1 ii)]. From a classical empirical central limit theorem we deduce that

$$\sqrt{n}\left(H_n - H,\, H_{n0} - H_0,\, H_{n2} - H_2\right) \rightsquigarrow (G,\, G_0,\, G_2) \qquad in \quad D^3([0,\infty]).$$

Moreover, almost surely, the empirical (sub)distributions (H_n, H_{n0}, H_{n2}) restricted on $[\sigma, \infty]$ belong to $\tilde{D}$, for sufficiently large n. Applying the delta method [see Theorem 3 of Gill (1989)] we deduce the weak convergence of $\sqrt{n}(H_n - H,\, H_{n0} - H_0,\, M_{n2}^- - M_2^-)$ in $D^3([\sigma, \infty])$. The computation of the covariance structure of the Gaussian limit process is given in the appendix.

For the weak convergence in $D^3([t_{00}, \infty])$ we apply Lemma 5.12. In order to extend G_M to a Gaussian process on $[t_{00}, \infty]$, consider the pathwise limit of $G_{M\sigma}$ as $\sigma \downarrow t_{00}$, which exists in view of the condition (5.2). Moreover, the extended process $(G_t,\, G_{0t},\, G_{Mt})$, $t \in [t_{00}, \infty]$ is a Gaussian process and the condition *i)* of Lemma 5.12 is fulfilled. It remains to verify the condition *ii)* of this lemma when $H([0, t_{00}]) = 0$. It suffices to prove that, for any $\varepsilon > 0$ and $\delta > 0$, there exists $\sigma = \sigma(\varepsilon, \delta) > t_{00}$ such that $H_2(\sigma) = 0$ and

$$\overline{\lim}_{n\to\infty} P\,(\sup_{t_{00} < t \leq \sigma} \sqrt{n} \left| M_{n2}^-([t, \sigma)) - M_2^-\,([t, \sigma)) \right| > \delta) < \varepsilon. \tag{5.5}$$

The idea is to reverse the time and to interpret $M_{n2}^-([t, \infty])$ as a Nelson-Aalen estimator. Use (H_2, H_{01}) and define a classical right-censorship model. Denote by L_1 and S_1 the corresponding independent lifetime and right-censoring variables, respectively. Take h a strictly monotone transformation reversing the time, for instance, $h(t) = 1/(1+t)$. Now, consider that we observe an i.i.d. sample $(\widetilde{T}_i, \Delta_i)$, $i = 1, ..., n$ where $\widetilde{T}_i = \min[h(L_{1i}), h(S_{1i})]$ and $\Delta_i = 1\!\!I\{h(L_{1i}) \leq h(S_{1i})\} = 1\!\!I\{S_{1i} \leq L_{1i}\}$. Moreover, the variables $h(L_1)$ and $h(S_1)$ are independent. Let π denote the distribution function of $\widetilde{T}_i$ and let π_n be its empirical counterpart. Let also Λ^- be the predictable hazard function corresponding to the variable $h(L_1)$. It is easy to see that $M_{n2}^-([t, \infty])$ is nothing else than $\Lambda_n^-([0, h(t)])$, that is the Nelson-Aalen estimator of $\Lambda^-([0, h(t)])$ which in turn coincides with $M_2^-([t, \infty])$. Thus, the inequality (5.5) is equivalent with the following one

$$\overline{\lim}_{n\to\infty} P\,(\sup_{h(\sigma) \leq u < h(t_{00})} \sqrt{n} \left| \int_{(h(\sigma), u]} (\Lambda_n^- - \Lambda^-)(ds) \right| > \delta) < \varepsilon. \tag{5.6}$$

Now, for proving this inequality we can use the 'classical' techniques [see Gill (1983) and Ying (1989); see also Fleming and Harrington (1991), Theorems 6.2.1 and 6.3.2].

Define the random time $U = \max_i \widetilde{T}_i$ and consider the stopped process $(\Lambda_n^- - \Lambda^-)^U([0, t]) = (\Lambda_n^- - \Lambda^-)([0, t \wedge U])$ which is a squared integrable (continuous time) martingale. Set $t' < h(t_{00})$. Applying Lenglart's inequality [see, e.g., Corollary 3.4.1 of Fleming and Harrington (1991)], for any $\delta, \eta > 0$, and f a left-continuous function with right-hand limits,

$$P\Big(\sup_{h(\sigma)\le u\le t'\wedge U}\sqrt{n}\left|\int_{(h(\sigma),u]} f(s)(\Lambda_n^- - \Lambda^-)(ds)\right| > \delta\Big)$$
$$\le \frac{\eta}{\delta^2} + P\Big(\int_{(h(\sigma),t'\wedge U]} f^2(s)\frac{1-\Lambda^-(s)}{\pi_n([s,\infty])}\Lambda^-(ds) \ge \eta\Big)$$
$$\le \frac{\eta}{\delta^2} + P\Big(\int_{(h(\sigma),t'\wedge U]} f^2(s)\frac{\Lambda^-(ds)}{\pi_n([s,\infty])} \ge \eta\Big).$$

Let $\gamma \in (0,1)$ and set

$$\eta = \int_{[h(\sigma),\, h(t_{00}))} \frac{f^2(s)}{\gamma\,\pi([s,\infty])}\Lambda^-(ds).$$

From Lemma 2.7 of Gill (1983) we obtain

$$P\Big(\sup_{h(\sigma)\le u\le t'\wedge U}\sqrt{n}\left|\int_{(h(\sigma),u]} f(s)(\Lambda_n^- - \Lambda^-)(ds)\right| > \delta\Big) \tag{5.7}$$
$$\le \frac{\eta}{\delta^2} + (e/\gamma)e^{-1/\gamma} + P\Big(\int_{(h(\sigma),\,t']} \frac{f^2(s)}{\gamma\,\pi([s,\infty])}\Lambda^-(ds) \ge \eta\Big)$$
$$= \frac{\eta}{\delta^2} + (e/\gamma)e^{-1/\gamma}.$$

Take $f \equiv 1$. In view of condition (5.2), for any $\gamma \in (0,1)$, we deduce $\eta \to 0$ as $\sigma \downarrow t_{00}$. Thus the last quantity in the display can be made arbitrarily small by a suitable choice of γ and σ.

On the other hand, condition 5.2 implies that $M_2^-((t_{00},\infty])$ is finite. From this deduce that $F_2^1([0,t_{00}]) > 0$ [use, for instance, the same type of arguments as in Lemma 6 in Gill and Johansen (1990)]. Thus, if $\widetilde{F}$ denotes the distribution of the variable $h(L_1)$, we have $\widetilde{F}([h(t_{00}),\infty]) > 0$. Now, we can write

$$\sqrt{n}\,\Lambda^-((U,h(t_{00}))) \le \sqrt{n}\,\ln\frac{\tilde{F}((U,\infty])}{\tilde{F}([h(t_{00}),\infty])} \le \sqrt{n}\,\frac{\tilde{F}((U,h(t_{00})))}{\tilde{F}([h(t_{00}),\infty])}.$$

Let G be the distribution function of the variable $h(S_1)$. Following the lines of the proof of Lemma 3 of Ying (1989), for any $\lambda > 0$, we can define

$$u_n^\lambda = \inf\{s:\ \sqrt{n}\widetilde{F}((s,h(t_{00}))) \le \lambda\}$$

and write

$$P\left(\sqrt{n}\,\widetilde{F}((U,h(t_{00}))) > \lambda\right)$$
$$\le P\left(U < u_n^\lambda\right)$$
$$\le \left[1-\widetilde{F}([u_n^\lambda,h(t_{00})))G([u_n^\lambda,\infty])\right]^n$$
$$\le \left[1-\frac{\lambda^2}{n}\frac{G([u_n^\lambda,\infty])}{\widetilde{F}([u_n^\lambda,h(t_{00})))}\right]^n \to 0,$$

since

$$\frac{\widetilde{F}([u_n^\lambda, h(t_{00})))}{G([u_n^\lambda, \infty])} \leq \int_{[u_n^\lambda, h(t_{00}))} \frac{\widetilde{F}(ds)}{G([s,\infty])} \leq \int_{[u_n^\lambda, h(t_{00}))} \frac{\Lambda^-(ds)}{\pi([s,\infty])}$$

and, in view of condition (5.2), the last integral converges to zero as $n \to \infty$. Thus, $\sqrt{n}\Lambda^-((U, h(t_{00}))) \to 0$, in probability. Now the proof is complete. ∎

Lemma 5.14 *Assume the condition (5.2). If $F_{2t}^1 = F_2^1([0,t])$, then*

$$\sqrt{n}\left(H_n - H,\, H_{n0} - H_0,\, F_{n2}^1 - F_2^1\right) \rightsquigarrow (G,\, G_0,\, G_3) \qquad in \quad D^3([t_{00}, \infty]),$$

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

$$G_{3t} = -F_2^1([0,t]) \int_{(t,\infty]} \frac{dG_{Mu}}{1 - M_2^-(u)} \,. \tag{5.8}$$

We have the following covariances

$$E(G_s G_{3t}) = F_2^1([0,t]) \int_{(s \vee t, \infty]} \frac{1}{\left(1 - M_2^-(u)\right)} \frac{H([0,s])}{H([0,u])} M_2^-(du)$$

$$E(G_{0s} G_{3t}) = F_2^1([0,t]) \int_{(t,\infty]} \frac{1}{\left(1 - M_2^-(u)\right)} \frac{H_0([0, s \wedge u])}{H([0,u])} M_2^-(du)$$

$$E(G_{3s} G_{3t}) = F_2^1([0,s])\, F_2^1([0,t]) \int_{(s \vee t, \infty]} \frac{M_2^-(du)}{\left(1 - M_2^-(u)\right) H([0,u])}.$$

Proof. Since

$$F_2^1([0,t]) = \prod_{(t,\infty]} (1 - M_2^-(ds)), \qquad t \in [0,\infty],$$

we deduce this result as a consequence of Lemma 5.13 and the Hadamard differentiability of the product-integration [see Lemma A.1 i)]. The computation of the covariance structure of the Gaussian limit process is given in the appendix. ∎

In view of the variance of the process G_3 and of the Theorem 1.2 in Gill (1983), one may ask if the condition (5.2) could be relaxed when proving the last lemma above. This is perhaps possible if $F_2^1([0,t_{00}]) = 0$. Nevertheless, in the following theorem, due to the lack of a useful (local) martingale structure for $L_{nT}^{1-} - L_T^{1-}$ it is quite convenient to ensure that the denominator appearing in the definition of L_T^{1-} stays 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 (5.2) is needed to bound the variance of G_{3t} when t decreases to t_{00}.

Now, we are able to state the asymptotic normality for L_{nT}^{1-} and F_{nT}^1. Using the same type of arguments as in the previous lemmas, the computation of the covariance structures of their Gaussian limits is straightforward, albeit tedious. Therefore, we shall omit the explicit form of the covariances structures of the limits of our estimators. In the following, the notation G_{3u-} and G_{u-} means that we consider the left-limits of G_{3u} and G_u, respectively.

Theorem 5.15 *Suppose condition (5.2). Let* $t_{00} < \tau$ *such that* $H_{01}([0,\tau)) < 1$. *If* $L_{Tt}^{1-} = L_T^{1-}([0,t])$, *then*
i)

$$\sqrt{n}\left(L_{nT}^{1-} - L_T^{1-}\right) \rightsquigarrow V \qquad in \quad D([0,\tau]),$$

where

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

is a zero-mean Gaussian process.
ii) If $F_{Tt}^1 = F_T^1([0,t])$, *then*

$$\sqrt{n}(F_{nT}^1 - F_T^1) \rightsquigarrow W \qquad in \quad D([0,\tau]),$$

with

$$W_t = F_T^1((t,\infty]) \int\limits_{(0,t]} \frac{dV_u}{1 - L_T^{1-}(u)}$$

a zero-mean Gaussian process.

Proof. *i)* Since condition (5.2) implies $F_2^1([0,t_{00}]) > 0$ and the definition of τ ensures $F_{01}^1([\tau,\infty]) > 0$, there exists $\epsilon > 0$ such that $\inf_{(t_{00},\tau]}(F_2^1 - H)([0,s)) > \epsilon$. Moreover, from the strong uniform convergence of $F_{n2}^1([0,s))$ [see section 5.1] we deduce that, almost surely, $\inf_{(t_{00},\tau]}(F_{n2}^1 - H_n)([0,s)) \geq \epsilon$ for sufficiently large n. Thus, in the case where $H_0(t_{00}) = 0$, the weak convergence of $\sqrt{n}\left(L_{nT}^{1-} - L_T^{1-}\right)$ is easily obtained by the delta-method [see Lemma A.1 ii) in the appendix and Theorem 3 of Gill (1989); see also van der Vaart and Wellner (1996), example 3.9.19, page 383].

When $H_0(t_{00}) > 0$ (hence, necessarily $t_{00} > 0$), in the definition of L_T^{1-} and L_{nT}^{1-} we also have to take into account $F_2^1([0,t_{00}))$ and $F_{n2}^1([0,t_{00}))$, respectively. In this case we can extend the weak convergence result in (5.3) on $D^3([0,\infty])$ by modifying in a suitable way the predictable reverse hazard function on $[0,t_{00})$. For instance, redefine M_2^- by

$$M_{2t}^- = M_2^-((t,\infty]) = \int_{(t,\infty]} \frac{H_2(du)}{H([0,u \vee t_{00}])}, \qquad t \in [0,\infty],$$

and let M_{n2}^{-} be its empirical counterpart. Since the denominator in the last display stays away from zero, the weak convergence of $\sqrt{n}(H_n - H,\, H_{n0} - H_0,\, F_{n2}^1 - F_2^1)$ on $D^3([0,\infty])$ can be easily obtained by the delta-method, where now F_2^1 and F_{n2}^1 correspond to the modified M_2^- and M_{n2}^-, respectively. The processes G_M and G_3 are still defined according to (5.4) and (5.8), respectively. Note that the definitions of $M_2^-(t_{00})$ (which is now necessarily smaller than one) and $M_{n2}^-(t_{00})$ are unchanged and therefore $F_2^1([0,t_{00}))$ (which is now necessarily positive) and $F_{n2}^1([0,t_{00}))$ remain the same. As a consequence, L_T^{1-} and L_{nT}^{1-} do not change. Finally, we may invoke one more time the delta-method and deduce the weak convergence of $\sqrt{n}\left(L_{nT}^{1-} - L_T^{1-}\right)$.

ii) This is easily obtained from *i)* and the Hadamard differentiability of the product-integration using the delta-method. ∎

5.2.2 Asymptotic Normality for F_{nT}^2

Assume that $H_0(0) = 0$ and recall that $t_{10} = \inf\{t : H_0((t,\infty]) = 0\}$. In order to work with transformations for which the Hadamard differentiability property is already available, we shall slightly change the inversion formula for F_T^2. Let M_T^{2+} be the unpredictable reverse hazard function corresponding to M_T^{2-} [see (2.24)]. Then we can write

$$\left(\overline{H}, H_0, H_1\right) \to \left(\overline{H}, H_0, L_1^+\right) \to \left(\overline{H}, H_0, \overline{F}_1^2\right) \to M_T^{2+} \to F_T^2,$$

where $\overline{H}_t = H((t,\infty])$,

$$L_{1t}^+ = L_1^+([0,t]) = \int_{[0,t]} \frac{H_1(ds)}{H((s,\infty])},$$

$$\overline{F}_{1t}^2 = F_1^2((t,\infty]) = \frac{1}{\prod_{[0,t]}(1 + L_1^+(ds))},$$

$$M_{Tt}^{2+} = M_T^{2+}((t,\infty]) = \int_{(t,\infty]} \frac{H_0(ds)}{F_1^2([s,\infty]) - H([s,\infty])}$$

$$F_{Tt}^2 = F_T^2([0,t]) = \frac{1}{\prod_{(t,\infty]}(1 + M_T^{2+}(ds))}.$$

For the asymptotic normality of F_{nT}^2 we proceed as follows. The central limit theorem yields

$$\sqrt{n}\left(\overline{H}_n - \overline{H},\, H_{n0} - H_0,\, H_{n1} - H_1\right) \rightsquigarrow \left(\overline{U},\, U_0,\, U_1\right) \qquad in \quad D^3([0,\infty]),$$

with $(\overline{U}, U_0, U_1)$ a zero mean Gaussian process. Let $\sigma' > 0$ such that $H((\sigma', \infty]) > 0$. Use the delta-method and deduce the weak convergence of $\sqrt{n}(L_{n1}^+ - L_1^+)$ in $D([0,\sigma'])$. Next, deduce the same result for $\sqrt{n}(\overline{F}_{n1}^2 - \overline{F}_1^2)$. Take $0 < \tau' < \sigma'$ such that $H_{02}([0,\tau')) > 0$, which means that $F_{02}^2([0,\tau')) > 0$. Since $(F_1^2 - H)([\cdot,\infty])$ stays away from zero on $[\tau',\sigma']$, the delta-method easily provides the weak convergence of $\sqrt{n}(M_{nT}^{2+}((\cdot,\sigma']) - M_T^{2+}((\cdot,\sigma']))$ in $D([\tau',\sigma'])$. Apply Lemma 5.12 to obtain the asymptotic normality of $M_{nT}^{2+}((\cdot,\infty])$ in $D([\tau',\infty])$. For this step we use the following assumption:

$$\int_{[0,t_{10})} \frac{H_1(ds)}{H((s,\infty])^2} < \infty. \tag{5.9}$$

In particular, this condition implies that $L_1^+([0,t_{10})) < \infty$ and therefore, $F_1^2([t_{10},\infty]) > 0$. Consequently, the denominator appearing in the definition of $M_T^{2+}((t,\infty])$ stays away from zero as $t \uparrow t_{10}$. Finally, deduce the asymptotic normality for F_{nT}^2. For the sake of a shorter presentation, we shall omit the covariances functions of the limit processes appearing in this subsection.

Lemma 5.16 *Let $\sigma' > 0$ such that $H((\sigma',\infty]) > 0$.*

i)

$$\sqrt{n}\left(\overline{H}_n - \overline{H},\, H_{n0} - H_0,\, L_{n1}^+ - L_1^+\right) \rightsquigarrow \left(\overline{U},\, U_0,\, V_1\right) \qquad \text{in} \quad D^3([0,\sigma']),$$

where

$$V_{1t} = \int_{[0,t]} \frac{dU_{1s}}{H((s,\infty])} - \int_{[0,t]} \frac{\overline{U}_s}{H((s,\infty])^2} H_1(ds);$$

is a zero-mean Gaussian process.

ii)

$$\sqrt{n}\left(\overline{H}_n - \overline{H},\, H_{n0} - H_0,\, \overline{F}_{n1}^2 - \overline{F}_1^2\right) \rightsquigarrow \left(\overline{U},\, U_0,\, V_2\right) \qquad \text{in} \quad D^3([0,\sigma']),$$

with

$$V_{2t} = -F_1^2((t,\infty]) \int_{[0,t]} \frac{dV_{1s}}{1 + L_1^+(s)}$$

a zero-mean Gaussian process.

Proof. *i)* It suffices to use the Hadamard differentiability of the mapping $(\overline{H}, H_0, H_1) \to (\overline{H}, H_0, L_1^+)$. See, e.g., Lemmas 3.9.17 and 3.9.25 in van der Vaart and Wellner (1996).

ii) Use the Hadamard differentiability of the map $A \to 1/A$ composed with the product-integration mapping [see, e.g., Lemmas 3.9.25 and 3.9.30 in van der Vaart and Wellner (1996)]. ∎

Now, we can prove the asymptotic normality for M_{nT}^{2+} and F_{nT}^{2} .

Theorem 5.17 *Suppose (5.9) holds. Let $\tau' > 0$ such that $H_{02}([0,\tau')) > 0$. Then,*

i)

$$\sqrt{n}\left(M_{nT}^{2+} - M_T^{2+}\right) \rightsquigarrow V_M \qquad in \quad D([\tau',\infty]),$$

where

$$V_{Mt} = \int_{(t,\infty]} \frac{dU_{0s}}{(F_1^2 - H)([s,\infty])} - \int_{(t,\infty]} \frac{V_{2s-} - \overline{U}_{s-}}{(F_1^2 - H)([s,\infty])^2} H_0(ds);$$

is a zero-mean Gaussian process.

ii)

$$\sqrt{n}(F_{nT}^2 - F_T^2) \rightsquigarrow V_F \qquad in \quad D([\tau',\infty]),$$

with

$$V_{Ft} = -F_T^2\,([0,t]) \int_{(t,\infty]} \frac{dV_{Mu}}{1 + M_T^{2+}(u)}$$

a zero-mean Gaussian process.

Proof. Note that if $t_{10} < t_1 = \inf\{t : H((t,\infty]) = 0\}$, then we can take $t_{10} \le \sigma'$ in Lemma 5.16 and the result is an immediate consequence of the delta-method since the map $(H, H_0, \overline{F}_1^2) \to M_T^{2+}$ is Hadamard differentiable.

If $t_{10} = t_1$, let $\sigma' > 0$ such that $H((\sigma',\infty]) > 0$. After applying Lemma 5.16 ii) and the delta-method on $D([\tau',\sigma'])$, in view of Lemma 5.12, it remains to prove: *a)* for any $t \ge \tau'$, the variance of V_{Mt} remains bounded as $\sigma' \uparrow t_1$; and *b)* for any $\varepsilon > 0$ and $\delta > 0$, there exists $\sigma' = \sigma'(\varepsilon,\delta) < t_1$ such that

$$\overline{\lim}_{n\to\infty} P\,(\sup_{u\in[\sigma',t_1)} \sqrt{n}\left|M_{nT}^{2+}((u,t_1]) - M_T^{2+}\,((u,t_1])\right| > \delta) < \varepsilon. \qquad (5.10)$$

Recall that $(F_1^2 - H)([s,\infty]) = F_1^2([s,\infty])F_{02}^2([0,s))$ and that condition (5.9) implies $F_1^2([t_1,\infty]) > 0$. In view of the definition of U_0, $\overline{U}_{s-}$ and V_{2s-} it is not difficult to see that condition (5.9) ensures *a)* [see also the covariance calculations appearing in Model 1]. To verify *b)*, note that the probability of the event in (5.10) is bounded by

$$P\Big(\sup_{u \in [\sigma', t_{1n})} \sqrt{n} \left| M^{2+}_{nT}((u, t_{1n}]) - M^{2+}_{T}((u, t_{1n}]) \right| \; > \; \delta/2)$$

$$+ P\, (\sqrt{n} M^{2+}_{T}((t_{1n}, t_1]) \; > \; \delta/2) = I + II,$$

where t_{n1} is the largest observation, i.e., $t_{n1} = \inf\{t : H_n((t, \infty]) = 0\}$. We have

$$II \leq P\, \left(\sqrt{n} H\, ((t_{1n}, t_1]) > \delta/2 F^2_1([t_1, \infty]) \right)$$

and the last quantity approaches zero as $n \to \infty$ [see, e.g., Lemma 2 in Ying (1989)]. On the other hand, note that

$$M^{2+}_T(dt) = \frac{H_0(dt)}{(F^2_1 - H)([t, \infty])} = f(t)\, L^-_0(dt),$$

with $f(t) = F^2_{02}([t, \infty])/F^2_{02}([0, t))$ and $L^-_0(dt) = H_0(dt)/H([t, \infty])$. Thus,

$$\begin{aligned}
\sqrt{n} | M^{2+}_{nT}((u, t_{1n}]) \; &- \; M^{2+}_{T}((u, t_{1n}]) \,| \\
&= \; \sqrt{n} \left| \int_{(u, t_{1n}]} f_n(t) L^-_{n0}(dt) - \int_{(u, t_{1n}]} f(t) L^-_0(dt) \right| \\
&\leq \; \sqrt{n} \left| \int_{(u, t_{1n}]} (f_n(t) - f(t)) L^-_{n0}(dt) \right| \\
&+ \; \sqrt{n} \left| \int_{(u, t_{1n}]} f(t) (L^-_{n0} - L^-_0)(dt) \right| \\
&= \; A_{1u} + A_{2u},
\end{aligned}$$

where $f_n(t) = F^2_{n02}([t, \infty])/F^2_{n02}([0, t))$. Using the same kind of arguments as in (5.7) deduce that $P(\sup_{u \in [\sigma', t_{1n})} A_{2u} > \delta/4)$ can be made arbitrarily small by a suitable choice of σ'. For this note also that

$$\int_{(\sigma', t_1]} \frac{f^2(s)}{H([s, \infty])} L^-_0(ds) = \int_{(\sigma', t_1]} \frac{H_0(ds)}{F^2_1([s, \infty])^2 F^2_{02}([0, s))^2}$$

decreases to zero as $\sigma' \uparrow t_1$, provided that $F^2_1([t_1, \infty]) > 0$. Finally, write

$$A_{1u} \leq \{ A_{2u} + \int_{(\sigma', t_{1n}]} f(t) L^-_0(dt) \} \sup_{(\sigma', t_1]} |f_n/f - 1|$$

and remark that F^2_{n02} is nothing else than a Kaplan-Meier estimator. Use Lemma 2.6 in Gill (1983) to bound for $\sup_{(\sigma', t_{1n}]} |f_n/f - 1|$ [the assumption of continuous distributions is not needed; see also Shorack and Wellner (1986), p 317]. Deduce that I can be made arbitrarily small by a good choice of σ'. Now, the proof of *i)* is complete.

ii) The same arguments as in Lemma 5.16 ii) apply. ■

Appendix

A.1 Hazard and Reverse Hazard Measures

In general, if F is a probability distribution on $(\overline{I\!R}^+, \overline{\mathcal{B}}^+)$ where $\overline{I\!R}^+ = [0,\infty]$ and $\overline{\mathcal{B}}^+$ denotes its associated Borel $\sigma-$field, three hazard measures may be defined: the "usual" hazard measure

$$L([0,t]) \doteq -\ln F((t,\infty]),$$

the "predictable" hazard measure

$$L^-(dt) \doteq \frac{F(dt)}{F([t,\infty])}$$

and the "unpredictable" hazard measure

$$L^+(dt) \doteq \frac{F(dt)}{F((t,\infty])}.$$

It may be seen that these three hazard measures have the same support and the same set of discontinuity points as F. Moreover their continuous parts are the same and their point masses are in bijection. Indeed,

$$L(t) = -\ln[1 - L^-(t)] = \ln[1 + L^+(t)]$$

(recall that for any measure μ, $\mu(t)$ is a short notation for $\mu(\{t\})$). This leads to the following representations of the probability distribution F:

$$\begin{aligned} F((t,\infty]) &= \exp\{-L([0,t])\} \\ &= \exp\{-\tilde{L}([0,t])\} \times \prod_{0\le s\le t} [1 - L^-(s)] \\ &= \prod_{[0,t]}(1 - L^-(ds)) \end{aligned}$$

where $\tilde{L}$ denotes the continuous part of L, i.e.,

$$\tilde{L}(B) = L(B) - \sum_{t\in B} L(t)$$

for any $B \in \overline{\mathcal{B}}^+$, and $\prod$ is the product-integral of the measure L^- [see, e.g., Gill and Johansen (1990)]. Note that in these representations, the mass of L at infinity is irrelevant and that F may have a supplementary mass at infinity. Indeed,

$$F(\infty) = \exp\{-L([0,\infty))\}.$$

Note also that, given an arbitrary measure on $(\overline{I\!R}^+, \overline{\mathcal{B}}^+)$, we can always define a probability distribution on the same space by considering this arbitrary measure as being the usual, the predictable or the unpredictable hazard measure and using the relations above.

Similarly, by reversing time, three reverse hazard measures may be associated to F: the "usual" reverse hazard measure defined as

$$M((t,\infty]) \doteq -\ln F([0,t])$$

the "predictable" reverse hazard measure defined as

$$M^-(dt) \doteq \frac{F(dt)}{F([0,t])}$$

and the "unpredictable" reverse hazard measure defined as

$$M^+(dt) \doteq \frac{F(dt)}{F([0,t))} \tag{A.1}$$

Once again, these three reverse hazard measures have the same support and the same set of discontinuity points as F. Moreover their continuous parts, denoted as $\tilde{M}$, are the same and their point masses are in bijection. Indeed,

$$M(t) = -\ln[1 - M^-(t)] = \ln[1 + M^+(t)]$$

This leads to the following representations of the probability distribution F:

$$\begin{aligned}
F([0,t]) &= \exp\{-M((t,\infty])\} \\
&= \exp\{-\tilde{M}((t,\infty])\} \times \prod_{t<s\leq\infty} [1 - M^-(s)] \\
&= \prod_{(t,\infty]} (1 - M^-(ds)).
\end{aligned}$$

As before, note that in these representations, the mass of M at zero is irrelevant and that F may have a supplementary mass at zero. Indeed,

$$F(0) = \exp\{-M((0,\infty])\}.$$

Moreover, note that given an arbitrary measure on $(\overline{I\!R}^+, \overline{\mathcal{B}}^+)$, we can always define a probability measure by regarding this arbitrary measure as a usual, predictable or unpredictable reverse hazard measure and using the suitable relationships above.

Clearly, the hazard measure and the reverse hazard measure are also in bijection, since

$$1 - e^{-L([0,t])} = e^{-M((t,\infty])}$$

and, in particular,

$$M^{-}(dt) = \frac{e^{-L([0,t))}}{1 - e^{-L([0,t])}} L^{-}(dt)$$

and

$$L^{-}(dt) = \frac{e^{-M((t,\infty])}}{1 - e^{-M([t,\infty])}} M^{-}(dt).$$

A.2 Proofs of section 2

A.2.1 Proof of Proposition 2.1

From the assumption, we obtain

$$H_2(dt) = \{1 - F_T((t,\infty])\, P[R_1 > t \mid L_1 \leq t]\}\, F_{L_1}(dt).$$

Moreover

$$H([0,t]) = F_{L_1}([0,t])\, \{1 - F_T((t,\infty])\, P[R_1 > t \mid L_1 \leq t]\}.$$

Therefore, $M_2^{-}(dt) = M_{L_1}^{-}(dt)$ for all t such that

$$1 - F_T((t,\infty])\, P[R_1 > t \mid L_1 \leq t] > 0.$$

In particular, this will be true for all t such that $F_T([0,t]) > 0$, i.e., for $t > s_{0T}$ and for $t = s_{0T}$ if $F_T(s_{0T}) > 0$ [see section 2.5 for the definition of s_{0T}]. Hence, for all such t,

$$F_2^1([0,t)) = F_{L_1}([0,t)).$$

Now, for $t \in supp(F_T)$,

$$F_2^1([0,t)) - H([0,t)) = F_{L_1}([0,t)) - H([0,t)) = F_T([t,\infty])\, P[L_1 < t \leq R_1]$$

and

$$H_0(dt) = P[L_1 < t \leq R_1]\, F_T(dt).$$

Therefore, clearly $L_T^{1-}(dt) = 0$ for $t \in I_1^c$, and $L_T^{1-}(dt) = L_T^{-}(dt)$ for $t \in I_1$. ■

A.2.2 Proof of Proposition 2.2

Under the assumption, we obtain

$$H_1(dt) = \{1 \,-\, F_T\,(\,[0,t]\,)\,P\,[L_2 \leq t \mid R_2 > t]\}\,F_{R_2}(\,dt\,)\,.$$

Moreover,

$$H(\,(t,\infty]\,) = F_{R_2}(\,(t,\infty]\,)\,\{1 \,-\, F_T(\,[0,t]\,)\,P\,[L_2 \leq t \mid R_2 > t]\}\,.$$

Therefore, $L_1^+(dt) \;=\; L_{R_2}^+(dt)$ for all t such that

$$1 - F_T(\,[0,t]\,)\,P\,[L_2 \leq t \mid R_2 > t] > 0.$$

In particular, this will be true for all t such that $F_T(\,(t,\infty]\,) > 0$, i.e., for $t < s_{1T}$ [see section 2.5 for the definition of s_{1T}]. Hence, for $t \leq s_{1T}$,

$$F_1^2(\,[t,\infty]\,) \;=\; F_{R_2}(\,[t,\infty]\,)\,.$$

Now, for $t \in supp(F_T)$ (this implies that $t \leq s_{1T}$),

$$F_1^2([t,\infty]) - H([t,\infty]) = F_{R_2}([t,\infty]) - H([t,\infty]) = F_T(\,[0,t)\,)\;P\,[L_2 < t \leq R_2]$$

and

$$H_0(dt) \;=\; P\,[L_2 < t \leq R_2]\,F_T(dt).$$

Therefore, clearly $M_T^{2+}(dt) \;=\; 0$ for $t \in I_2^c$, and $M_T^{2+}(dt) \;=\; M_T^+(dt)$ for $t \in I_2$. Given the relationship between the predictable and the unpredictable reverse hazard measures we deduce that $M_T^{2-}(dt) \;=\; M_T^-(dt)$ on I_2 and $M_T^{2-}(dt) \;= 0$ on I_2^c. ∎

A.3 Expressions of F_T

According to Proposition 2.1 and formula 2.26, in Model 1,

$$F_T^1(\,(t,\infty]\,) \;=\; \exp\{\,-\,L_T(\,[0,t] \cap (s_{0L_1}, s_{1R_1})\,)\,\}$$

if $F_{R_1}(s_{1R_1}) = 0$ and if $F_{R_1}(s_{1R_1}) > 0$,

$$F_T^1(\,(t,\infty]\,) \;=\; \exp\{\,-\,L_T(\,[0,t] \cap (s_{0L_1}, s_{1R_1}]\,)\,\}\,.$$

Therefore, in Model 1, if $F_{R_1}(s_{1R_1}) = 0$

$$F_T^1([0,t]) = \begin{cases} 0 & \text{for} \quad 0 \le t \le s_{0L_1} \\[2ex] \dfrac{F_T(\,(s_{0L_1}, t]\,)}{F_T(\,(s_{0L_1}, \infty]\,)} & \text{for} \quad s_{0L_1} < t < s_{1R_1} \\[2ex] \dfrac{F_T(\,(s_{0L_1}, s_{1R_1})\,)}{F_T(\,(s_{0L_1}, \infty]\,)} & \text{for} \quad s_{1R_1} \le t < \infty \\[2ex] 1 & \text{for} \quad t = \infty, \end{cases}$$

and if $F_{R_1}(s_{1R_1}) > 0$

$$F_T^1([0,t]) = \begin{cases} 0 & \text{for} \quad 0 \le t \le s_{0L_1} \\[2ex] \dfrac{F_T(\,(s_{0L_1}, t]\,)}{F_T(\,(s_{0L_1}, \infty]\,)} & \text{for} \quad s_{0L_1} < t < s_{1R_1} \\[2ex] \dfrac{F_T(\,(s_{0L_1}, s_{1R_1}]\,)}{F_T(\,(s_{0L_1}, \infty]\,)} & \text{for} \quad s_{1R_1} \le t < \infty \\[2ex] 1 & \text{for} \quad t = \infty. \end{cases}$$

Similarly, in Model 2,

$$F_T^2(\,[0,t]\,) \;=\; \exp\{\,-M_T(\,(t,\infty] \cap (s_{0L_2}, s_{1R_2})\,)\,\}$$

if $F_{R_2}(s_{1R_2}) = 0$ and if $F_{R_1}(s_{1R_2}) > 0$,

$$F_T^2(\,[0,t]\,) \;=\; \exp\{\,-M_T(\,(t,\infty] \cap (s_{0L_2}, s_{1R_2}]\,)\,\}\,.$$

Therefore, in Model 2, if $F_{R_2}(s_{1R_2}) = 0$

$$F_T^2([0,t]) = \begin{cases} \dfrac{F_T(\,[0, s_{0L_2}]\,)}{F_T(\,[0, s_{1R_2})\,)} & \text{for} \quad 0 \le t \le s_{0L_2} \\[2ex] \dfrac{F_T(\,[0, t]\,)}{F_T(\,[0, s_{1R_2})\,)} & \text{for} \quad s_{0L_2} < t < s_{1R_2} \\[2ex] 1 & \text{for} \quad s_{1R_2} \le t \le \infty, \end{cases}$$

and if $F_{R_2}(s_{1R_2}) > 0$

$$F_T^2([0,t]) = \begin{cases} \dfrac{F_T(\,[0, s_{0L_2}]\,)}{F_T(\,[0, s_{1R_2}]\,)} & \text{for} \quad 0 \le t \le s_{0L_2} \\[2ex] \dfrac{F_T(\,[0, t]\,)}{F_T(\,[0, s_{1R_2}]\,)} & \text{for} \quad s_{0L_2} < t < s_{1R_2} \\[2ex] 1 & \text{for} \quad s_{1R_2} \le t \le \infty. \end{cases}$$

A.4 Hadamard differentiability and asymptotic normality

Lemma A.1 *i) Let $C > 0$. The product-integral map $\phi : BV_C([a,b]) \to D([a,b])$,*

$$\phi(A)_t = \prod_{(t,b]} (1 + A(ds))$$

is Hadamard differentiable with derivative

$$\phi'_A(\alpha)_t = \int_{(t,b]} \prod_{(t,u)} (1 + A(ds)) \prod_{(u,b]} (1 + A(ds))\, d\alpha(u),$$

If α is of unbounded variation, then the integral with respect to α is defined by integration by parts.

ii) Let $C, \varepsilon > 0$ and $BV_{C,\varepsilon}([a,b]) = \{A \in BV_C([a,b]),\ A \ge \varepsilon\}$. Then, the map $\psi : BV_{C,\varepsilon}([a,b]) \times BV_C([a,b]) \to D([a,b])$ (the product space is equipped with the product norm)

$$\psi(A,B)_t = \int_{(t,b]} \frac{1}{A} dB$$

is Hadamard differentiable and

$$\psi'_{A,B}(\alpha,\beta)_t = \int_{(t,b]} \frac{1}{A} d\beta - \int_{(t,b]} \frac{\alpha}{A^2} dB,$$

where $\int_{(t,b]} (1/A) d\beta$ is defined through integration by parts if β is not of bounded variation.

Proof. i) The proof is based on the Duhamel equation and is the same as for Theorem 8 of Gill and Johansen (1990) [see also Lemma 3.9.30 of van der Vaart and Wellner (1996)].

ii) This is a consequence of the Hadamard differentiability of two maps giving ψ, that is $A \to 1/A$ and $(A,B) \to \int_{(\cdot,b]} A dB$ [see Lemma 3 of Gill (1989) or Lemma 3.9.17 of van der Vaart and Wellner (1996); see also their section 3.9.4.3]. Note that the range of ψ is included in $BV_{C'}([a,b])$ for some $C' > 0$. ■

A.4.1 Proof of Lemma 5.13

(for the sake of shorter formulae, in this and the following proof H^* denotes $H^*([0,u])$ where H^* is any of H, H_0 or H_2; moreover, M_{2s}^- stands for $M_2^-((s,\infty]))$. The covariance function for (G, G_0) is obvious. For the covariance function of the process G_M we write $E(G_{Ms}G_{Mt}) = E(A_sA_t) - E(A_sB_t) - E(B_sA_t) + E(B_sB_t)$ and we make use of the recepies

$$E[\int h_1(s)dG_{Ms}\int h_2(t)dG_{Mt}] = \int\int h_1(s)h_2(t)dE[G_{Ms}G_{Mt}]$$

and

$$E[\int h_1(s)dG_{Ms}\int h_2(t)G_{Mt}dh_3(t)] = \int\int h_1(s)h_2(t)d_sE[G_{Ms}G_{Mt}]dh_3(t)$$

[see, for instance, Wang *et al.* (1986), p 1603 for similar computations]. We obtain

$$E(A_sA_t) = \int\limits_{(s,\infty]}\int\limits_{(t,\infty]} \frac{d\left[H_2(u\wedge v) - H_2(u)H(v)\right]}{H(u)H(v)} = \int\limits_{(s\vee t,\infty]} \frac{H_2(du)}{H^2(u)} - M_{2s}^-M_{2t}^-,$$

$$\begin{aligned}
E(A_sB_t) &= \int\limits_{(s,\infty]}\int\limits_{(t,\infty]} \frac{H_2(dv)d_u\left[H_2(u\wedge v) - H_2(u)H(v)\right]}{H(u)H^2(v)} \\
&= \int\limits_{(t,\infty]} \frac{1}{H^2(v)} \int\limits_{(s,v\vee s]} \frac{H_2(du)}{H(u)} H_2(dv) - M_{2s}^-M_{2t}^- \\
&= \int\limits_{(s\vee t,\infty]} \frac{M_2^-((s,v])}{H^2(v)} H_2(dv) - M_{2s}^-M_{2t}^-,
\end{aligned}$$

$$\begin{aligned}
E(B_sA_t) &= \int\limits_{(s,\infty]}\int\limits_{(t,\infty]} \frac{d_v\left[H_2(u\wedge v) - H(u)H_2(v)\right]H_2(du)}{H^2(u)H(v)} \\
&= \int\limits_{(s,\infty]} \frac{1}{H^2(u)} \int\limits_{(t,u\vee t]} \frac{H_2(dv)}{H(v)} H_2(du) - M_{2s}^-M_{2t}^- \\
&= \int\limits_{(s\vee t,\infty]} \frac{M_2^-((t,u])}{H^2(u)} H_2(du) - M_{2s}^-M_{2t}^-,
\end{aligned}$$

$$E\left(B_sB_t\right)=\int\limits_{(s,\infty]}\int\limits_{(t,\infty]}\frac{H\left(u\wedge v\right)-H\left(u\right)H(v)}{H^2\left(u\right)H^2\left(v\right)}H_2(dv)H_2(du)$$

$$=\int\limits_{(s,\infty]}\int\limits_{(t,u\vee t]}\frac{H\left(v\right)}{H^2\left(u\right)H^2\left(v\right)}H_2(dv)H_2(du)$$

$$+\int\limits_{(s,\infty]}\int\limits_{(u\vee t,\infty]}\frac{H\left(u\right)}{H^2\left(u\right)H^2\left(v\right)}H_2(dv)H_2(du)-M_{2s}^-M_{2t}^-$$

$$=\int\limits_{(s\vee t,\infty]}\frac{M_2^-((t,u])}{H^2\left(u\right)}H_2(du)+\int\limits_{(s\vee t,\infty]}\frac{1}{H^2\left(v\right)}\int\limits_{(s,v)}\frac{H_2(du)}{H\left(u\right)}H_2(dv)-M_{2s}^-M_{2t}^-$$

$$=\int\limits_{(s\vee t,\infty]}\frac{M_2^-((t,u])}{H^2\left(u\right)}H_2(du)+\int\limits_{(s\vee t,\infty]}\frac{M_2^-((s,v))}{H^2\left(v\right)}H_2(dv)-M_{2s}^-M_{2t}^-.$$

Consequently,

$$E\left(G_{Ms}G_{Mt}\right)=\int_{(s\vee t,\infty]}\frac{1-M_2^-(u)}{H(u)}M_2^-(du).$$

In order to obtain the last two covariances we can write

$$E\left(G_sG_{Mt}\right)=E[G_s\int\limits_{(t,\infty]}\frac{dG_2(u)}{H\left(u\right)}]-E[G_s\int\limits_{(t,\infty]}\frac{G_u}{H^2\left(u\right)}H_2(du)]$$

$$=\int\limits_{(t,\infty]}\frac{d_u\left[H_2\left(u\wedge s\right)-H_2(u)H(s)\right]}{H(u)}-\int\limits_{(t,\infty]}\frac{H\left(u\wedge s\right)-H\left(u\right)H(s)}{H^2\left(u\right)}H_2(du)$$

$$=\int\limits_{(s\wedge t,s]}\frac{H_2(du)}{H\left(u\right)}-\int\limits_{(t,\infty]}\frac{H\left(u\wedge s\right)}{H^2\left(u\right)}H_2(du)$$

$$=\int\limits_{(s\wedge t,s]}\frac{H_2(du)}{H\left(u\right)}-\int\limits_{(t,s\vee t]}\frac{H_2(du)}{H\left(u\right)}-H(s)\int\limits_{(s\vee t,\infty]}\frac{H_2(du)}{H^2\left(u\right)}$$

$$=-H(s)\int\limits_{(s\vee t,\infty]}\frac{M_2^-(du)}{H([0,u])}.$$

Finally,

$$E\left(G_{0s}G_{Mt}\right) = E[G_{0s}\int\limits_{(t,\infty]}\frac{dG_2(u)}{H\left(u\right)}] - E[G_{0s}\int\limits_{(t,\infty]}\frac{G_u}{H^2\left(u\right)}H_2(du)]$$

$$= -\int\limits_{(t,\infty]}\frac{d_u\left[H_0\left(s\right)H_2(u)\right]}{H\left(u\right)} - \int\limits_{(t,\infty]}\frac{H_0\left(u\wedge s\right) - H_0\left(s\right)H(u)}{H^2\left(u\right)}H_2(du)$$

$$= -\int\limits_{(t,\infty]}\frac{H_0\left(u\wedge s\right)}{H^2\left(u\right)}H_2(du).$$

A.4.2 Proof of Lemma 5.14

Use Lemma 5.13 and deduce

$$E\left(G_sG_{3t}\right) = -F_2^1\left([0,t]\right)E[G_s\int\limits_{(t,\infty]}\frac{dG_M(u)}{1-M_2^-(u)}]$$

$$= F_2^1\left([0,t]\right)\int_{(s\vee t,\infty]}\frac{1}{\left(1-M_2^-(u)\right)}\frac{H(s)}{H(u)}M_2^-(du).$$

Moreover,

$$E\left(G_{0s}G_{3t}\right) = -F_2^1\left([0,t]\right)E[G_{0s}\int\limits_{(t,\infty]}\frac{dG_M(u)}{1-M_2^-(u)}]$$

$$= F_2^1\left([0,t]\right)\int_{(t,\infty]}\frac{1}{\left(1-M_2^-(u)\right)}\frac{H_0(s\wedge u)}{H(u)}M_2^-(du).$$

Finally, use again Lemma 5.13 and deduce

$$E\left(G_{3s}G_{3t}\right) = F_2^1\left([0,s]\right)F_2^1\left([0,t]\right)E[\int\limits_{(s,\infty]}\frac{dG_M(u)}{1-M_2^-(u)}\int\limits_{(t,\infty]}\frac{dG_M(v)}{1-M_2^-(v)}]$$

$$= F_2^1\left([0,s]\right)F_2^1\left([0,t]\right)\int\limits_{(s\vee t,\infty]}\frac{M_2^-(du)}{\left(1-M_2^-(u)\right)H(u)}.$$

References

[1] Chang, M.N.(1990). Weak convergence of a self-consistent estimator of the survival function with doubly censored data, *Ann. Statis.* **18**, 391-404.

[2] Chang, M.N. and Yang, G. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data, *Ann. Statis.* **15**, 1536-1547.

[3] Ferguson, T.S.and Phadia, E. (1979). Bayesian nonparametric estimation based on censored data, *Ann. of Statis.* **7**, 163-186.

[4] Fernandez, A.J., Bravo, J.I. and De Fuentes, I. (1997). Lifetime Estimation from Doubly Censored Data and Dirichlet Process Prior Knowledge on the Observable Random Vector, *Communications in Statistics - Theory and Methods* **26(6)**, 1541-1558.

[5] Fleming, T.R. and Harrington, P. (1991). *Counting Processes and Survival Analysis*, Wiley, New-York.

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

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

[8] Gill, R. (1994). Lectures on Survival Analysis (Ecole d'été de Probabilités de Saint-Flour XXII 1992), *Lecture Notes in Mathematics* (ed. P. Bernard), **1581**, 115-241, Springer Verlag, New-York.

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

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

[11] Hamburg, B.A., Kraemer H.C. and Jahnke, W. (1975). A Hierarchy of drug use in adolescence: behavioral and attitudinal correlates of substantial drug use, *American J. Psychiatry* **132**, 1155-1163.

[12] Huang, J. (1999). Asymptotic Properties of Nonparametric Estimation Based on Partly Interval-Censored Data, *Statistica Sinica* **9**, 501-519.

[13] Kaplan, E.L. and Meier, P. (1958). Non-parametric estimation from incomplete observations, *J. Amer. Statis. Assoc.* **53**, 457-481.

[14] Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many nuisance parameters, *Ann. Math. Statis.* **27**, 887-906.

[15] Liederman, P.H., Barbu, D., Kagia J., Kraemer H.C. and Leiderman, G.F. (1974). African infant precocity and some social influences during the first year, *Nature* **242**, 247-249.

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

[17] 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.

[18] Samuelsen, S. O. (1989). Asymptotic Theory for Non-parametric Estimators from Doubly Censored Data, *Scand. J. Statis.* **16**, 1-21.

[19] Shorack, G. and Wellner, J.A. (1986) *Empirical Processes with Applications to Statistics.* Wiley, New-York.

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

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

[22] Van der Laan, M.J. and Gill, R.D. (1999). Efficiency of the sieved-NPMLE in CAR-Censored Data Models, *Math. Methods Statis.* **8**, 251-276.

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

[24] Wang, M.-C., Jewell, N.P. and Tsai, W.-Y (1986). Asymptotic properties of the product limit estimate under random truncation, *Ann. Statis.* **14**, 1597-1605.

[25] Wellner, J.A. and Zhang, Y. (1997). A Hybrid Algorithm for Computation of the Nonparametric Maximum Likelihood Estimator From Censored Data, *J. Amer. Statis. Assoc.* **92**, 945-959.

[26] Wellner, J.A. and Zhang, Y. (1995). Double Censoring: Characterization and Computation of the Nonparametric Maximum Likelihood Estimator, Tech. Rep. no. 292, Dept of Stat, University of Washington.

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