I N S T I T U T D E
S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

0115

NONPARAMETRIC COMPETING RISKS MODELS: IDENTIFICATION AND STRONG CONSISTENCY

JEAN-MARIE ROLIN

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

Nonparametric Competing Risks Models : Identification and Strong Consistency

Jean-Marie ROLIN *

April 2001

Abstract

In this paper, it is shown, in the general case, that a multiple causes death model is equivalent to a competing independent risks model. Uniqueness of the representation and identification conditions are discussed. Martingales arguments generalizing results of Stute and Wang (1993) are used to show the almost sure convergence of simple functionals of the predictable hazard measures and of the distributions of the latent or "fictitious" independent risks. These results entails the almost sure uniform convergence on the real line of the distributions of these independent risks.

AMS MOS primary: 62N01
secundary: 62N02, 60F15

Key words: Nonparametric Statistics, Competing risks, Multiple causes death model, Strong Uniform Consistency.

* Institut de statistique, Université catholique de Louvain.
Research support from the contract "Projet d'Actions de Recherche Concertées" (PARC No. 98/03-217) of the Belgian Government is gratefully acknowledged.

1 Introduction

In multiple causes death models, we observe a lifetime, Y, a finite positive random variable, and the cause of death, A, a categorical random variable with values in a finite set of possible causes : $K = \{r : 1 \leq r \leq g\}$. Typically, the model considers a finite number of precisely defined causes along with a "residual" cause (conventionally the last one) that gathers all other possible causes of death and/or a censoring state. In this paper, we will only consider the simple model in which $\{(Y_i\,,\,A_i) : 1 \leq i \leq n\}$, is an i.i.d. sample of size n of the random vector $(Y\,,\,A)$ whose distribution will be denoted by $\tilde{H}$. Section 2 will be devoted to some terminologies and notations.

In section 3, we will present the competing independent risks model. In such a model, independent latent lifetimes are introduced for each of the studied causes of death, $\{T_r : 1 \leq r \leq g\}$. These random variables are positive but may be infinite for $r < g$ because it is conceivable not to die from a well defined cause, but T_g is supposed to be finite because one cannot be immortal (or the observation has to be made in case of censoring). We will also suppose that causes of death have been ordered in such a way that, in cases of equality of two or several latent lifetimes, the death is due to the smallest cause. Under mild assumptions, we will show that a multiple causes death model is equivalent to a unique competing independent risks model.

Finally, in section 4, extending results of Stute and Wang (1993), we will show, using martingale arguments, that the "predictable" hazard functions of these independent "fictitious" lifetimes are almost surely uniformly consistent on the set where they are finite and that the distributions of these independent "fictitious" lifetimes are almost surely consistent on the real line.

2 Multiple Causes Death Model

As said in the Introduction, the sampling statistical model is an i.i.d. sample of size n, $\{\,(Y_i\,,\,A_i) : 1 \leq i \leq n\,\}$, of the random vector $(Y\,,\,A)$ where Y is a positive random variable defined on an arbitrary probability space $(\Omega\,,\,\mathcal{M})$ with values in $(\mathbb{R}^+\,,\,\mathcal{B}^+)$, ($\mathcal{B}^+$ are the Borel subsets of $\mathbb{R}^+ = [0,\infty)$) and A is a random variable on $(\Omega\,,\,\mathcal{M})$ with values in $(K\,,\,2^K)$ where $K = \{r : 1 \leq r \leq g\}$ and 2^K is the collection of all subsets of K. The joint distribution of $(Y\,,\,A)$, denoted by $\tilde{H}$, is then a probability measure on $(\mathbb{R}^+ \times K\,,\,\mathcal{B}^+ \bigotimes 2^K)$ and there exists a probability measure $P_{\tilde{H}}$ on $(\Omega\,,\,\mathcal{M})$ such that

$$(i) \qquad \underset{1 \leq i \leq n}{\perp\!\!\!\perp} (\,Y_i\,,\,A_i\,)\,;\,P_{\tilde{H}} \tag{2.1}$$

$$(ii) \qquad P_{\tilde{H}} \circ (Y_i\,,\,A_i\,)^{-1} \;=\; P_{\tilde{H}} \circ (\,Y\,,\,A\,)^{-1} \;=\; \tilde{H} \quad \forall\, 1 \,\leq\, i \,\leq\, n\,.$$

In nonparametric statistics, $\tilde{H}$ is the parameter and it is conveniently specified through its *sub-distributions.* For $r \in K$, they are defined as

$$H_r(\,B\,) \doteq P_{\tilde{H}}[\,Y \,\in\, B\,,\,A \,=\, r\,] \,=\, \tilde{H}(\,B \times \{r\}\,) \quad \forall \quad B \in \mathcal{B}^+\,. \tag{2.2}$$

Note that the marginal distribution of Y is given by

$$H\,(\,B\,) \doteq P_{\tilde{H}}[\,Y \,\in\, B\,] \,=\, \tilde{H}(\,B \times K\,) \quad \forall \quad B \in \mathcal{B}^+ \tag{2.3}$$

and therefore

$$H \;=\; \sum_{1 \leq r \leq g} H_r\,. \tag{2.4}$$

We will use "mutatis mutandis" the same notations for the empirical distribution of the sample, the "generalized" maximum likelihood estimator of the distribution $\tilde{H}$, defined as

$$\tilde{H}_n \;=\; \frac{1}{n} \sum_{1 \leq i \leq n} \varepsilon_{(\,Y_i\,,\,A_i\,)}\,. \tag{2.5}$$

By the strong law of large numbers, we know that, if f is any positive H-integrable Borel measurable function on $(\mathbb{R}^+\,,\,\mathcal{B}^+)$, $H_n\,(\,f\,)$ converges almost surely to $H\,(\,f\,)$ where $H\,(\,f\,)$ is a short notation for

$$H\,(\,f\,) \;=\; \int_{[\,0\,,\,\infty\,)} f(t)\,H\,(dt)$$

By the Glivenko -Cantelli Theorem, this implies the almost sure strong uniform convergence of H_n to H, i.e., the norm defined as

$$||\;H_n \;-\; H\;|| = \sup_{0 \leq t \leq \infty} |\;H_n\,(\,[\,0\,,\,t\,]\,) \;-\; H\,(\,[\,0\,,\,t\,]\,)\;|$$

tends to zero almost surely as n tends to infinity.
Evidently the same results are obtained for H_r replacing H.

For identification purposes, we will have to consider some features of the support of the subdistributions : the set of point masses will be denoted by

$$I_r = \{\,t : H_r\,(t) > 0\,\} \tag{2.6}$$

where $H_r\,(t)$ is a short notation for $H_r\,(\{t\})$ and the right end-point of the support will be denoted as

$$t_r \;=\; \inf\{\,t : H_r\,(\,(t\,,\,\infty]\,) \;>\; 0\,\}\,. \tag{2.7}$$

If, with the the same definitions, I and t_* denotes the set of point masses and the right end-point of the support of H, it is clear that

$$I = \bigcup_{1 \leq r \leq g} I_r \tag{2.8}$$

and

$$t_* = \max_{1 \leq r \leq g} t_r \,. \tag{2.9}$$

We now turn to the competing risks model.

3 Competing Risks Models

3.1 General Competing Risks Model

As said in the introduction, in a competing risks model, we introduce latent random variables T_r, $1 \leq r \leq g$, defined on $(\Omega, \mathcal{M})$ with values in $(\overline{I\!R}^+, \overline{\mathcal{B}}^+)$, representing the time of death due to cause r ($\overline{\mathcal{B}}^+$ denotes the Borel subsets of $\overline{I\!R}^+ = [0, \infty]$). $\{T_i : 1 \leq i \leq n\}$ is an i.i.d. sample of the random vector $T = \{T_r : 1 \leq r \leq g\}$ whose distribution is denoted by N, a probability measure on $(\overline{I\!R}^{+g}, \overline{\mathcal{B}}^{+g})$. Therefore, there exists a probability P_N on $(\Omega, \mathcal{M})$ such that

$$\begin{array}{lll} (i) & \displaystyle\coprod_{1 \leq i \leq n} T_i \quad ; \; P_N & \\ (ii) & P_N \circ T_i^{-1} \;=\; P_N \circ T^{-1} \;=\; N \quad \forall\, 1 \leq i \leq n \,. & \end{array} \tag{3.1}$$

The marginal distributions of the lifetimes of the causes are given by

$$\tilde{F}_r\,(dt) \doteq P_N[\,T_r \in dt\,] \qquad \forall\, t \in \overline{I\!R}^+ \qquad \forall\, r \in K \,. \tag{3.2}$$

Note that $\tilde{F}_r$ is a probability distribution on $(\overline{I\!R}^+, \overline{\mathcal{B}}^+)$ $\forall \; r < g$ while $\tilde{F}_g$ is a probability distribution on $(I\!R^+, \mathcal{B}^+)$ since we suppose that

$$\tilde{F}_g\,(\,[0, \infty)\,) = P_N\,[\,T_g \,<\, \infty\,] \,=\, 1 \,.$$

Now, the observed lifetime is given by

$$Y \,=\, \min_{1 \leq r \leq g} T_r \tag{3.3}$$

and the cause of death is given by

$$A = \min \{ r : T_r = Y \}, \tag{3.4}$$

i.e., the causes of death have been ordered in such a way that, in cases of equality of two or several lifetimes, the death is attributed to the smallest index of the cause.
In this competing model, the following equation shows that the sub-distribution of each cause is absolutely continuous with respect to the distribution of the corresponding lifetime.

$$H_r(B) = \int_B P_N[R = r \mid T_r = t] \, \tilde{F}_r(dt) \quad \forall B \in \mathcal{B}^+ \quad \forall r \in K. \tag{3.5}$$

From definition (3.3), we deduce that

$$H((t, \infty]) = P_N[Y > t] = P_N[\bigcap_{1 \leq r \leq g} \{T_r > t\}]$$

but thanks to definition (3.4), this may be generalized as follows:

Lemma 3.1 For any $r \in K$ and $t \in I\!R^+$,

$$\begin{aligned} &P_N[Y > t, R < r] + P_N[Y \geq t, R \geq r] \\ &\quad = P_N[\bigcap_{1 \leq \ell < r} \{T_\ell > t\} \cap \bigcap_{r \leq \ell \leq g} \{T_\ell \geq t\}]. \end{aligned}$$

■

The proof is a straightforward consequence of the following logical set equality implied by the definition of A :

$$\{Y > t, R < r\} \cup \{Y \geq t, R \geq r\} = \bigcap_{1 \leq \ell < r} \{T_\ell > t\} \cap \bigcap_{r \leq \ell \leq g} \{T_\ell \geq t\}.$$

For statistical inference, it is easily seen that the observation is a function of the times of death, i.e., $(Y, A) = a(T)$ and model (3.1) implies model (2.1) with $\tilde{H} = N \circ a^{-1}$.There is however a problem of identification in the sense that different N's may generate the same $\tilde{H}$. A way to avoid this problem is to assume identification restrictions. The simplest one is to assume that the random variables T_r, $1 \leq r \leq g$, are independent. Several authors (see, e.g., Williams, Lagakos (1977), Kalbfleisch, Mac Kay (1979) and Tsai (1986)) have shown that weaker assumptions provide identification of $\{\tilde{F}_r : 1 \leq r \leq g\}$. Note that these assumptions are untestable since the times of death of the different causes are never simultaneously observed.

However, under the assumptions provided by (3.3) and (3.4), it has been shown (see, e.g., Mouchart, Rolin (1995)) that whatever $\tilde{H}$ is, there exists a (almost) unique competing independent risks model generating $\tilde{H}$.

3.2 Competing Independent Risks Model

This model is defined as follows :

$$\begin{aligned} &(i) \qquad \coprod_{1\le r\le g} T_r \; ; \; P_N \\ &(ii) \qquad P_N \circ T_r^{-1} \;=\; \tilde{F}_r \quad \forall\, r \in K \end{aligned} \tag{3.6}$$

so that $N \;=\; \bigotimes_{1\le r\le g} \tilde{F}_r \,.$

The following lemma is a simple corollary of formula (3.5) and expresses $\tilde{H}$ as a functional of $\{\, \tilde{F}_r \,:\, 1 \le r \le g \,\}$.

Lemma 3.2 For any $r \in K$ and $B \in \mathcal{B}^+$,

$$H_r\,(B) \;=\; \int_B U_r(t)\, \tilde{L}_r^-\,(dt) \;=\; \int_B U_{r+1}(t)\, \tilde{L}_r^+\,(dt)$$

where

$$\tilde{L}_r^-\,(dt) \;\doteq\; \tilde{F}_r\,(\,[t,\infty]\,)^{-1}\, \tilde{F}_r\,(dt)$$

$$\tilde{L}_r^+\,(dt) \;\doteq\; \tilde{F}_r\,(\,(t,\infty]\,)^{-1}\, \tilde{F}_r\,(dt)$$

and

$$U_r\,(t) \doteq \prod_{1\le \ell<r} \tilde{F}_\ell(\,(t,\infty]\,) \prod_{r\le\ell\le g} \tilde{F}_\ell(\,[t,\infty]\,)\,.$$

■

Note that

$$U_1\,(t) \;=\; H\,(\,[t,\infty]\,) \quad \text{and} \quad U_{g+1}\,(t) \;=\; H\,(\,(t,\infty]\,)\,. \tag{3.7}$$

Now, $\tilde{L}_r^-$ and $\tilde{L}_r^+$ are two different definitions, in the general case, of the *usual* hazard measure defined by

$$\tilde{L}_r\,(\,[0,t]\,) \;=\; -\ln \tilde{F}_r\,(\,(t,\infty]\,)\,. \tag{3.8}$$

We will call $\tilde{L}_r^-$ the *predictable* hasard measure and $\tilde{L}_r^+$ the *unpredictable* hazard measure. For more details on these definitions and their interpretations see, e.g., Rolin (1998) or Florens, Mouchart, Rolin (1999). Recall that $\tilde{L}_r^-$ is also called the "compensator form" of the hazard measure in reference to the Doob-Meyer decomposition of the underlying process. It is known that

$$\tilde{L}_r^- \leq \tilde{L}_r \leq \tilde{L}_r^+ .$$

The continuous parts (denoted by a index c) of these three hazard measures are the same, i.e., if

$$J_r = \{ t : \tilde{F}_r(t) > 0 \} \tag{3.9}$$

then

$$\tilde{L}_r^- - \sum_{t \in J_r} \tilde{L}_r^-(t)\, \varepsilon_t \doteq \tilde{L}_{rc}^- = \tilde{L}_{rc} = \tilde{L}_{rc}^+$$

and the jumps are in bijection, i.e., $\forall \quad t \in J_r$,

$$\tilde{L}_r(t) = -\ln[\, 1 - \tilde{L}_r^-(t)\,] = \ln[\, 1 + \tilde{L}_r^+(t)\,] . \tag{3.10}$$

If we look at the supports of the distributions of the latent lifetimes, we may notice that, if

$$s_r \doteq \inf\{ t : \tilde{F}_r([0,t]) = 1 \} \tag{3.11}$$

then, by Lemma 3.2 and formula (3.7),

$$t_r \leq t_* = \min_{1 \leq r \leq g} s_r . \tag{3.12}$$

The following lemma is a simple corollary of Lemma 3.1 and is the clue for defining, for a general multiple causes death model, the equivalent competing independent risks model.

Lemma 3.3 For all $r \in K$ and for all $t \in \mathbb{R}^+$,

$$U_r(t) = V_r(t)$$

where

$$
\begin{aligned}
V_r(t) \quad &= P_{\tilde{H}}[\,Y > t\,,\, A < r\,] + P_{\tilde{H}}[\,Y \geq t\,,\, A \geq r\,] \\
&= \sum_{1 \leq \ell < r} H_\ell\,(\,(t,\infty]\,) + \sum_{r \leq \ell \leq g} H_\ell\,(\,[t,\infty]\,) \\
&= H\,(\,(t,\infty]\,) + \sum_{r \leq \ell \leq g} H_\ell\,(t)\,.
\end{aligned}
$$

■

3.3 Equivalent Competing Independent Risks Model

Considering a general multiple causes death model as defined in (2.1), we define the predictable hazard measures of the "fictitious" lifetimes of the causes by

$$L_r^-\,(dt) \doteq V_r\,(t)^{-1}\,H_r\,(dt) \qquad \forall\, r \in K\,. \tag{3.13}$$

According to formula (3.10), the usual hazard measures of the "fictitious" lifetimes of the causes are defined by

$$L_r\,(B) \doteq L_r^-\,(B) - \sum_{t \in B} L_r^-\,(t) - \sum_{t \in B} \ln[\,1 - L_r^-\,(t)\,] \tag{3.14}$$

and finally, the distributions of the "fictitious" lifetimes of the causes are defined, according to formula (3.8), by

$$F_r\,(\,(t,\infty]\,) \doteq \exp\{\,-L_r\,(\,[0,t]\,)\}\,. \tag{3.15}$$

Note that F_r may be defined directly from L_r^- using the theory of product integrals (see, e.g., Andersen, Bogdan, Gill, Keiding (1993), Fleming, Harrington (1991) or Gill (1994)). However, in this paper, we do not want to use sophisticated theories if we can easily avoid them.

Similarly, we define predictable and usual hazard measures of the marginal distribution of Y by :

$$L^-\,(dt) \doteq H\,(\,[t,\infty]\,)^{-1}\,H\,(\,dt\,)\,, \tag{3.16}$$

$$L\,(\,[0,t]\,) \doteq -\ln\,[\,H\,(\,(t,\infty]\,)\,]\,, \tag{3.17}$$

The next proposition shows that the multiple causes model may be represented as a competing independent risks model.

Proposition 3.4 For all $r \in K$ and for all $t \in I\!R^+$,

$$V_r(t) = \prod_{1 \leq \ell < r} F_\ell(\,(t, \infty]\,) \prod_{r \leq \ell \leq g} F_\ell(\,[t, \infty]\,)$$

and in particular

$$H(\,(t, \infty]\,) = \prod_{1 \leq r \leq g} F_r(\,(t, \infty]\,)\,.$$

■

A proof of this result in a particular case may be found in Mouchart and Rolin (1995). The general proof is given in Appendix (A.1).

Concerning the supports, it is easy to see that the set of point masses of L_r^- (and of L_r) is the same as the set of point masses of H_r, i.e., I_r. Note also that each point mass of L_r^- is a point mass of F_r but F_r may be defective, i.e., may have a supplementary mass at ∞. This happens if and only if $L_r(\,[0, t_r]\,) < \infty$ or equivalently if $L_r^-(\,[0, t_r)\,) < \infty$ and $L_r^-(t_r) < 1$. This is clearly the case if $t_r < t_*$. If $t_r = t_*$ and if $H(t_*) > 0$, then only F_{r_*} will be nondefective where $r_* = \max\{r : H_r(t_*) > 0\}$. Finally, if $H(t_*) = 0$, according to the above expression of H, at least one of the distributions must be nondefective but some of them may be defective.

3.4 Identification

In order to prove the unicity of the representation of a multiple causes death model as a competing independent risks model, it sufficies to prove that a competing independent risks model is identified.
Now, by the definition,

$$L_r^-(dt) = V_r(t)^{-1} H_r(dt)\,.$$

But in a competing independent risks model, we saw that

$$H_r(dt) = U_r(t)\,\tilde{L}_r^-(dt)$$

and that

$$U_r(t) = V_r(t) \qquad \forall\, t \in I\!R^+\,.$$

Therefore,

$$\begin{aligned} L_r^-(dt) &= \tilde{L}_r^-(dt) && \text{on} && \{t : U_r(t) > 0\} \\ &= 0 && \text{on} && \{t : U_r(t) = 0\}. \end{aligned}$$

Note that $[0, t_*) \subset \{t : U_r(t) > 0\} \subset [0, t_*]$.
This entails the following proposition :

Proposition 3.5 If the model is a competing independent risks model with probability distributions of the risks given by $\{\tilde{F}_r : r \in K\}$, then the equivalent competing independent risks model as defined in Section 3.3, has the following property : for all $r \in K$,

$$F_r([0,t]) = \tilde{F}_r([0,t]) \qquad \forall \quad t < t_*$$

■

It may be noted that Proposition 3.5 holds for $t = t_*$ if, either $\tilde{L}_r^-(t_*) = 0$ or $\tilde{L}_r^-(t_*) > 0$ and $U_r(t_*) > 0$.

Partial identification is thus achieved. To obtain full identification, we have to consider what is happening at t_*. Several situations must be considered :

a) If $t_* = \infty$, then $F_r = \tilde{F}_r$ for all $r \in K$.

b) If $t_* < \infty$ and $s_r > t_*$, then $\tilde{F}_r$ will not be fully identified.

c) If $t_* < \infty$ and $H(t_*) = 0$, then $U_r(t_*) = 0$ for all $r \in K$ and therefore $\tilde{F}_r$ will be fully identified if and only if $s_r = t_*$ and $\tilde{F}_r(t_*) = 0$. A mass of $\tilde{F}_r$ at t_* will become the mass of F_r at ∞.

d) If $t_* < \infty$ and $H(t_*) > 0$, then $\tilde{F}_r(t_*) > 0$ for all r such that $s_r = t_*$ and only the distribution corresponding to the smallest r in this set, i.e., r_* (defined at the end of Section 3.3) will be fully identified. The point masses of the distributions $\tilde{F}_r$ corresponding to the other r's will become point masses of F_r at ∞.

This gives the following corollary that generalizes the result of Crowder (1994):

Corollary 3.6 If a multiple causes death model is such that $t_* = \infty$, i.e., $H((t,\infty]) \doteq P_{\tilde{H}}[Y > t] > 0$ for all $t \in I\!R^+$, or if $t_* < \infty$ and

$\forall\ r \in K\,,\ t_r = t_*$ and $L_r^-\,(\,[0,t_*)\,) = \infty\,,$ then there exists a unique competing independent risks model with the same sub-distributions.

■

4 Estimators and Strong Consistency

In a competing independent risks model, the estimators of the hazard measures and of the distributions of the latent lifetimes are obtained by the plug-in principle, i.e., by replacing $\tilde{H}$ by $\tilde{H}_n$ and they are in fact the (generalized) maximum likelihood estimators. If the model is misspecified, i.e., in a general multiple causes death model, they are the estimators of the hazard measures and of the distributions of "fictitious" lifetimes of the equivalent competing independent risks model. We call them empirical hazard measures and empirical distributions of the (latent or fictitious) lifetimes.

4.1 Estimators

From now on, f will be a (finite) positive Borel measurable function on $(\,\overline{I\!R}^+\,,\,\overline{\mathcal{B}}^+\,)\,.$ Since $\tilde{H}_n$ is a purely discrete measure, the integral of f with respect to the empirical predictable hazard measures (for $r \in K$) may be written as

$$L_{nr}^-\,(\,f\,) \;=\; \sum_{0\le t<\infty} f(t)\,L_{nr}^-\,(\,t\,) \tag{4.1}$$

where

$$L_{nr}^-\,(\,t\,) \;=\; V_{nr}\,(\,t\,)^{-1}\,H_{nr}\,(\,t\,) \tag{4.2}$$

and

$$V_{nr}\,(\,t\,) \;=\; H_n\,(\,(t,\infty]\,) + \sum_{r\le\ell\le g} H_{n\ell}\,(\,t\,)\,. \tag{4.3}$$

Similarly, the integral of f with respect to the empirical distributions may be written as

$$F_{nr}\,(\,f\,) \;=\; \sum_{0\le t\le\infty} f(t)\,F_{nr}\,(\,t\,) \tag{4.4}$$

where, for $t < \infty\,,$

$$F_{nr}(t) = F_{nr}([t,\infty])\, L^-_{nr}(t) \tag{4.5}$$

and, for $t \leq \infty$,

$$F_{nr}([t,\infty]) = \prod_{0\leq s<t} [1 - L^-_{nr}(s)]. \tag{4.6}$$

In order to express these estimators in an easily computable form, we define the following statistics : $\{Z_j : 1 \leq j \leq M\}$ are the distinct values of the sample $\{Y_i : 1 \leq i \leq n\}$ in increasing order,

$$D_{rj} = \sum_{1\leq i\leq n} 1\!\!I_{\{Y_i = Z_j\,,\,A_i = r\}}$$

and

$$N_j = \sum_{1\leq i\leq n} 1\!\!I_{\{Y_i > Z_j\}}.$$

In other words, D_{rj} are the number of deaths at time Z_j due to cause r and N_j is the population at risk just after time Z_j. Now the estimators may be rewritten as

$$L^-_{nr}(f) = \sum_{1\leq j\leq M} f(Z_j)\, L^-_{nr}(Z_j) \tag{4.7}$$

and

$$F_{nr}(f) = \sum_{1\leq j\leq M} f(Z_j)\, F_{nr}(Z_j) + f(\infty)\, F_{nr}(\infty) \tag{4.8}$$

where

$$\begin{aligned} F_{nr}(Z_j) &= L^-_{nr}(Z_j) \prod_{1\leq \ell < j} [1 - L^-_{nr}(Z_\ell)] \\ F_{nr}(\infty) &= \prod_{1\leq \ell\leq M} [1 - L^-_{nr}(Z_\ell)] \end{aligned} \tag{4.9}$$

and

$$L^-_{nr}(Z_j) = \frac{D_{rj}}{N_j + \sum_{r\leq \ell\leq M} D_{\ell j}}. \tag{4.10}$$

The estimators of the distributions of the lifetimes are therefore product-limit estimators generalizing the Kaplan-Meier (1958) estimator.

4.2 Strong Consistency of the Predictable Hazard Measures

The aim of this section is to prove the following theorem:

Theorem 4.1 If $L_r^-(f) < \infty$, then $L_{nr}^-(f) \to L_r^-(f)$ almost surely and in L_1 with respect to $P_{\tilde{H}}$.

■

Theorem 4.1 is a particular case of a more general result. Let us define for $I \subset K$,

$$
\begin{aligned}
H_I &= \sum_{r \in I} H_r \\
t_I &= \max_{r \in I} t_r \\
V_I(t) &= H((t, \infty]) + H_I(t)
\end{aligned}
\tag{4.11}
$$

and for $J \subset I$, the following predictable hazard measure :

$$L^-(dt) = V_I(t)^{-1} H_J(dt). \tag{4.12}$$

Then, we will prove the following result:

Theorem 4.2 If $L^-(f) < \infty$, then $L_n^-(f) \to L^-(f)$ almost surely and in L_1 with respect to $P_{\tilde{H}}$.

■

Mimating the proof of the Glivenko-Cantelli Theorem will provide the following corollary :

Corollary 4.3 If either $t_J < t_*$, or $t_J = t_*$ and $H(t_*) > 0$ or $t_J = t_*$ and $H(t_*) = 0$ and $L^-([0, t_*)) < \infty$, then, almost surely $P_{\tilde{H}}$,

$$\| L_n^- - L^- \| = \sup_{0 \le t \le t_*} | L_n^-([0,t]) - L^-([0,t]) | \to 0.$$

If $t_J = t_*$ and $H(t_*) = 0$ and $L^-([0, t_*)) = \infty$, then, for all $t < t_*$, almost surely $P_{\tilde{H}}$,

$$\sup_{0 \le s \le t} | L_n^-([0,s]) - L^-([0,s]) | \to 0$$

and

$$L_n^- \, (\, [0, t_*) \,) \; \to \; \infty \, .$$

■

The proof closely follows the techniques of the proof for the Kaplan-Meier estimator and is a consequence of the following three lemmas:

Lemma 4.4 If $L^- \, (\, f \, 1\!\!I_{[0,t)} \,) < \infty$ and $H \, (\, [t, \infty] \,) \; > \; 0$ then, almost surely $P_{\tilde{H}}$,

$$L_n^- \, (\, f \, 1\!\!I_{[0,t)} \,) \; \to \; L^- \, (\, f \, 1\!\!I_{[0,t)} \,) \, .$$

■

Proof : Note first that V_{nI} converges to V_I uniformly on $[0, \infty]$. Now,

$$| \int_{[0,t)} f \, (s) \, V_{nI} \, (s)^{-1} \, H_{nJ} \, (ds) \; - \; \int_{[0,t)} f \, (s) \, V_I \, (s)^{-1} \, H_{nJ} \, (ds) \; |$$

$$\leq || \; V_{nI} \; - \; V_I \; || \; H_n \, (\, [t, \infty] \,)^{-1} \int_{[0,t)} f \, (s) \, V_I \, (s)^{-1} \, H_{nJ} \, (ds) \, .$$

The second member of the inequality tends to zero because

$$H_n \, (\, [t, \infty] \,) \; \to \; H \, (\, [t, \infty] \,)$$

and

$$\int_{[0,t)} f \, (s) \, V_I \, (s)^{-1} \, H_{nJ} \, (ds) \; \to \; L^- \, (\, f \, 1\!\!I_{[0,t)} \,)$$

by the strong law of large numbers.

■

This Lemma suffices to prove Theorem 4.2 and Corollary 4.3 in the case where, either $t_J \; < \; t_*$ or $t_J \; = \; t_*$ and $H \, (\, t_* \,) \; > \; 0$ (convergence at the point t_* in this latter case is a consequence of the strong law of large numbers). This Lemma also implies in letting t increase to t_* that for any f such that $L^- \, (\, f \, 1\!\!I_{[0,t)} \,) < \infty$ for all $t \; < \; t_*$,

$$L^- \, (\, f \, 1\!\!I_{[0,t_*)} \,) \; \leq \; \liminf_n \; L_n^- \, (\, f \, 1\!\!I_{[0,t_*)} \,)$$

and therefore the second part of Corollary 4.3 is also proved. As for the Kaplan-Meier estimator, the only difficult part of the proof concerns the case where $t_J = t_*$ and $H(t_*) = 0$ and $L^-([0, t_*)) < \infty$. In the Kaplan-Meier case, this problem has been finally settled in 1993 by Stute and Wang when they discover a reverse martingale argument to obtain a proof. Our second lemma generalizes this idea.

Lemma 4.5 Let us define the following σ-algebras :

$$\mathcal{F}_n = \sigma(\tilde{H}_n) \quad \text{and} \quad \mathcal{B}_n = \bigvee_{n \le m < \infty} \mathcal{F}_m .$$

Then, for all n,

$$E[L_n^-(f) \mid \mathcal{B}_{n+1}] \le L_{n+1}^-(f),$$

i.e., $L_n^-(f)$ is a reverse supermartingale.

■

The proof, a little bit technical, is given in Appendix A.2. Now as soon as the expectations of $L_n^-(f)$ are finite (it is an increasing sequence), $L_n^-(f)$ is $P_{\tilde{H}}$ - almost surely convergent. The limit is in the σ-algebra of asymptotic permutable events and is therefore a constant by the Hewit-Savage 0-1 law. This constant is equal to $\sup_n E[L_n^-(f)]$ if this last quantity is finite and the sequence $L_n^-(f)$ is L_1- convergent. Therefore, it suffices to show that $\sup_n E[L_n^-(f)]$ is equal to $L^-(f)$ to prove Theorem 4.2. This is a trivial consequence of the following lemma that gives the bias of the estimator :

Lemma 4.6

$$\begin{aligned} E[L_n^-(f)] &= \int_{[0,\infty]} f(t)\, P[V_{nI}(t) > 0]\, L^-(dt) \\ &= \int_{[0,\infty]} f(t)\, [1 - \{1 - V_I(t)\}^n]\, L^-(dt) \end{aligned}$$

■

As $n \to \infty$, the bracketted term in the last integral increases to $\mathbb{1}_{\{V_I(t) > 0\}}$ which is included in the support of L^- and therefore $E[L_n^-(f)]$ increases to $L^-(f)$.

In the Kaplan-Meier case, Gill (1994) shows that this result may be deduced from the general theory of stochastic processes of Dellacherie-Meyer and from the notion of the compensator of a counting process (see, e.g., Jacob (1975)) as presented in Andersen, Bogdan, Gill, Keiding (1993). Indeed, Lemma 4.6 would be a direct consequence of the fact that

$$M_{1t} \;=\; L_n^- \,(\,[0,t]\,) \;-\; \int_{[0,t]} \mathbb{1}_{\{\,V_{nI}\,(\,s\,)>0\,\}}\, L^-\,(\,ds\,)$$

is a martingale. This would be a corollary of the fact that

$$M_{2t} \;=\; H_{nJ}\,(\,[0,t]\,) \;-\; \int_{[0,t]} V_{nI}\,(\,s\,)\, L^-\,(\,ds\,)$$

is a martingale and that $V_{nI}\,(\,t\,)$ is a "predictable" process. But, this approach would require a new definition of predictable processes and this is far beyond the scope of this paper. However, a direct proof of the fact that M_{1t} is a martingale may be deduced from elementary properties of the multinomial process and from discrete approximations of right and left continuous processes. This proof is given in Appendix A.3.

4.3 Strong Consistency of the Distributions

We first look at the strong consistency of the usual hazard measures. For the general case, we have the following theorem :

Theorem 4.7 Let us suppose that $L\,(\,f\,) \;<\; \infty$. If $L^-\,(\,[0,t_*)\,) \;<\; \infty$, then, almost surely $P_{\tilde{H}}$,

$$L_n\,(\,f\,) \;\to\; L\,(\,f\,)\,.$$

If $L^-\,(\,[0,t_*)\,) \;=\; \infty$, then, for all $t \;<\; t_*$ almost surely $P_{\tilde{H}}$

$$L_n\,(\,f\;\mathbb{1}_{[0,t)}\,) \;\to\; L\,(\,f\;\mathbb{1}_{[0,t)}\,)$$

and this implies that

$$L\,(\,f\,) \;\leq\; \liminf_n\; L_n\,(\,f\,)\,.$$

∎

Proof : Since $L\,(\,f\,) \;<\; \infty$ implies that $L^-\,(\,f\,) \;<\; \infty$, in order to prove that $L_n\,(\,f\;\mathbb{1}_{[0,t)}\,) \;\to\; L\,(\,f\;\mathbb{1}_{[0,t)}\,)$ for all $t \;\leq\; t_*$, it suffices to prove, by the definition of L, formula (3.14), that

$$\sum_{0\,\leq\,s\,<\,t} f(s)\,[\,L_n\,(s) \;-\; L_n^-\,(s) \;-\; L\,(s) \;+\; L^-\,(s)\,] \;\to\; 0\,.$$

But,

$$
\begin{aligned}
L_n(s) - L_n^-(s) - L(s) + L^-(s) &= \ln\Big[\frac{1 - L^-(s)}{1 - L_n^-(s)}\Big] - [\, L_n^-(s) - L^-(s)\,] \\
\leq [\, L_n^-(s) - L^-(s)\,] \frac{L_n^-(s)}{1 - L_n^-(s)} &\leq \frac{\sup\limits_{0 \leq s < t} |\, L_n^-(s) - L^-(s)\,|}{1 - \sup\limits_{0 \leq s < t} L_n^-(s)} L_n^-(s) \\
\geq [\, L_n^-(s) - L^-(s)\,] \frac{L^-(s)}{1 - L^-(s)} &\geq - \frac{\sup\limits_{0 \leq s < t} |\, L_n^-(s) - L^-(s)\,|}{1 - \sup\limits_{0 \leq s < t} L^-(s)} L^-(s)\,.
\end{aligned}
$$

Now, for $t < t_*$ or $t = t_*$ and $L^-(\,[0, t_*)\,) < \infty$, the uniform convergence of the empirical predictable hazard measure on the interval $[0, t)$ implies that

$$\sup_{0 \leq s < t} |\, L_n^-(s) - L^-(s)\,| \to 0$$

and that

$$\sup_{0 \leq s < t} L_n^-(s) \to \sup_{0 \leq s < t} L^-(s)\,.$$

This last term is strictly smaller than 1. Furthermore, $\sum_{0 \leq s < t} f(s)\, L_n^-(s) = L_n^-(\, f\, 1\!\!I_{[0,t)}\,)$ converges almost surely $P_{\tilde{H}}$ to $L^-(\, f\, 1\!\!I_{[0,t)}\,)$ which is finite. Finally, if $H(\,t_*\,) > 0$, $L_n(\,t_*\,)$ converges to $L(\,t_*\,)$ almost surely $P_{\tilde{H}}$. This last term is finite if $L^-(\,t_*\,) < 1$. If $L^-(\,t_*\,) = 1$ (this corresponds to the case $H_{I-J}(\,t_*\,) = 0$ and $H_J(\,t_*\,) > 0$), $L(\,f\,) < \infty$ implies that $f(\,t_*\,) = 0$ and there is nothing to prove.

∎

This entails a Glivenko-Cantelli theorem for the usual empirical hazard measure.

Corollary 4.8 If $L^-(\,[0, t_*)\,) < \infty$, then, almost surely $P_{\tilde{H}}$,

$$\sup_{0 \leq s < t_*} |\, L_n(\,[0, s]\,) - L(\,[0, s]\,)\,| \to 0$$

and

$$L_n(\,t_*\,) \to L(\,t_*\,)\,.$$

If $L^-([0,t_*)) = \infty$, then, almost surely $P_{\tilde{H}}$,

$$\sup_{0 \le s \le t} \mid L_n([0,s]) - L([0,s]) \mid \to 0$$

for all $t < t_*$ and

$$L_n([0,t_*)) \to \infty .$$

■

This directly implies the uniform convergence on $[0,\infty]$ of the empirical distribution.

Corollary 4.9 Almost surely $P_{\tilde{H}}$,

$$\| F_n - F \| = \sup_{0 \le t \le \infty} \mid F_n([0,t]) - F([0,t]) \mid \to 0 .$$

■

We now consider the convergence of simple functionals or moments. Recall that

$$F(f) = \int_{[0,\infty]} f(s)\, F([s,\infty])L^-(ds) .$$

The aim is to prove the following theorem :

Theorem 4.10 If $F(f) < \infty$, then, almost surely $P_{\tilde{H}}$,

$$F_n(f) \to F(f) .$$

■

Theorem 4.10 is again a particular case of a more general result. Let us define for $J_k \subset I_k \subset K$, $k = 1, 2$, the two predictable hazard measures :

$$L_k^-(dt) = V_{I_k}(t)^{-1} H_{J_k}(dt) \tag{4.13}$$

and consider the following measure

$$G(dt) = F_2([t,\infty])\, L_1^-(dt) . \tag{4.14}$$

We have the following result :

Theorem 4.11 If $G(f) < \infty$, then, almost surely and in L_1 with respect to $P_{\tilde{H}}$,

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

■

Theorem 4.11 clearly implies Theorem 4.10 because convergence at the point ∞ if F is degenerate is trivial.
The proof seems to be rather delicate. Indeed, if $G(f) < \infty$, from Theorem 4.2 we obtain that

$$\int_{[0,\infty]} f(s)\, F_2([s,\infty])L_{n1}^-(ds) \to G(f).$$

Now, for $t \leq t_*$,

$$\begin{aligned} &| \; G_n(f\, 1\!\!1_{[0,t)}) - \int_{[0,t)} f(s)\, F_2([s,\infty])L_{n1}^-(ds) \; | \\ &\leq || \; F_{n2} - F_2 \; || \int_{[0,t)} f(s)\, L_{n1}^-(ds) \end{aligned}$$

and

$$G(f\, 1\!\!1_{[0,t)}) \geq F_2([t,\infty])\, L_1^-(f\, 1\!\!1_{[0,t)}).$$

Now, if $t < t_*$ or $t = t_*$ and $L_2^-([0,t_*)) < \infty$, then $L_1^-(f\, 1\!\!1_{[0,t)}) < \infty$ and almost surely $P_{\tilde{H}}$,

$$\int_{[0,t)} f(s)\, L_{n1}^-(ds) \to L_1^-(f\, 1\!\!1_{[0,t)}).$$

Therefore, since convergence at the point t_* is trivial, if $L_2^-([0,t_*)) < \infty$, we obtain the desired result.

Now, if f is bounded,

$$| \; G_n(f\, 1\!\!1_{[t,t_*)}) \; | \leq || \; f \; || \; L_{n1}^-([t,t_*)).$$

If $L_1^-([0,t_*)) < \infty$, the last term in the above inequality tends to $|| \; f \; ||$ $L_1([t,t_*))$ which can be made arbitrarily small by taking t close to t_* and we again obtain the desired result.

Note that, contrary to the case of predictable hazard measures, the difficulty arises here when $L_2^-([0,t_*)) = \infty$. Once again, it seems that the general case must rely on a reverse supermartingale argument. The proof of Theorem 4.11 is a direct consequence of the following two lemmas :

Lemma 4.12 The sequence $G_n(f)$ is a reversed supermartingale, i.e.,

$$E[\,G_n(f) \mid \mathcal{B}_{n+1}\,] \leq G_{n+1}(f)\,.$$

■

The proof of this Lemma, a litle bit technical, is given in Appendix A.4.

Lemma 4.13

$$\sup_n E[\,G_n(f)\,] = G(f)\,.$$

■

Proof : If $H(t_*) = 0$, let $f_{ct} = (f \wedge c)\, 1\!\!I_{[0,t)}$ with $t < t_*$, $G_n(f_{ct})$ converges to $G(f_{ct})$ almost surely $P_{\tilde{H}}$. But $G_n(f_{ct}) \leq c\, L^-_{n1}([0,t))$ and by Theorem 4.2, $L^-_{n1}([0,t))$ converges to $L^-_1([0,t))$ almost surely and in L_1 with respect to $P_{\tilde{H}}$. Therefore, $G_n(f_{ct})$ is uniformly integrable and converges to $G(f_{ct})$ in L_1 with respect to $P_{\tilde{H}}$. Now, by the reverse supermartingale property, the expectations are increasing. Hence,

$$\begin{aligned} &\sup_n \sup_c \sup_t E[\,G_n(f_{ct})\,] = \sup_n E[\,G_n(f)\,] \\ &= \sup_c \sup_t \sup_n E[\,G_n(f_{ct})\,] = \sup_c \sup_t G(f_{ct}) = G(f)\,. \end{aligned}$$

If $H(t_*) > 0$, then $L^-_1([0,t_*]) < \infty$ and $L^-_2([0,t_*]) < \infty$ and the same argument as above may be used in replacing f_{ct} by $f_c = (f \wedge c)\, 1\!\!I_{[0,t_*]}$.

■

Theorem 4.11 implies other interesting results than Theorem 4.10. As an example, let us consider, in the general multiple causes death model, for $1 \leq p \leq r \leq q \leq g$, the following measure :

$$G_{prq}(dt) = \prod_{p \leq \ell < r} F_\ell((t,\infty])\prod_{r < \ell \leq q} F_\ell([t,\infty])\,F_r(dt)\,.$$

This may be rewritten as

$$G_{prq}\,(\,dt\,)\;=\;\exp[\,-\sum_{p\le\ell\le q}L_\ell\,(\,[0,t)\,)\,]\prod_{p\le\ell<r}[\,1-L_\ell^-\,(\,t\,)\,]\,L_r^-\,(\,dt\,)\,.$$

Mimating the proof of Proposition 3.4, shows that this may be rewritten as

$$G_{prq}\,(\,dt\,)\;=\;F_2\,(\,[t,\infty]\,)\,L_1^-\,(\,dt\,)$$

where $I_1\,=\,I_2\,=\,\{\,\ell\,:\,p\le\ell\le g\,\}$, $J_1\,=\,\{\,r\,\}$ and $J_2\,=\,\{\,\ell\,:\,p\le\ell\le q\,\}$ and therefore, we obtain directly that, if $G_{prq}\,(\,f\,)\,<\,\infty$, $G_{nprq}\,(\,f\,)\,\to\,G_{prq}\,(\,f\,)$ almost surely and in L_1 with respect to $P_{\tilde H}$.

Finally, in reliability theory, it may happen that $(\,Y\,,\,A\,)$ is defined as

$$Y\;=\;\max_{1\le r\le g}\,T_r$$

and

$$A\;=\;\min\{\,r\,:\,T_r\;=\;Y\,\}\,.$$

A special case is provided by left-censored observations. We may duplicate this paper and obtain the same results. It suffices to reverse the time, i.e., to consider reverse hazard measures defined as:

$$M_r^-\,(dt)\;=\;\frac{F_r\,(dt)}{F_r\,(\,[0,t]\,)}\;=\;\frac{H_r\,(dt)}{H\,(\,[0,t)\,)\,+\,\sum_{r\le\ell\le g}H_\ell\,(\,t\,)}$$

where

$$F_r\,(\,[0,t]\,)\;=\;e^{\,-\,M_r\,(\,(t,\infty]\,)}\,.$$

Appendix

A.1 Proof of Proposition 3.4

At first, let us prove that for all $t \in \mathbb{R}^+$,

$$H((t,\infty]) = \prod_{1\le r\le g} F_r((t,\infty])$$

or equivalently, that for all $t \in \mathbb{R}^+$,

$$L([0,t]) = \sum_{1\le r\le g} L_r([0,t]).$$

On one hand, for the continuous part, since the set of point masses $I_r \subset I$ are countable sets, we have according to formulae (3.13) and (3.16),

$$\begin{aligned}
\sum_{1\le r\le g} L_{rc}([0,t]) &= \sum_{1\le r\le g} \int_0^t \mathbb{1}_{I_r^c}(s)\, V_r(s)^{-1} H_r(ds) \\
&= \sum_{1\le r\le g} \int_0^t \mathbb{1}_{I^c}(s)\, H([s,\infty])^{-1} H_r(ds) \\
&= \int_0^t \mathbb{1}_{I^c}(s)\, H([s,\infty])^{-1} H(ds) \\
&= L_c([0,t]).
\end{aligned}$$

On the other hand, for $t \in I$ and $t < t_*$,

$$\begin{aligned}
\sum_{1\le r\le g} L_r(t) &= \sum_{1\le r\le g} -\ln[1 - L_r^-(t)] \\
&= \sum_{1\le r\le g} -\ln \frac{V_{r+1}(t)}{V_r(t)} = -\ln \frac{V_{g+1}(t)}{V_1(t)} \\
&= -\ln \frac{H((t,\infty])}{H([t,\infty])} = -\ln[1 - L^-(t)] = L(t).
\end{aligned}$$

Now, if $t_* \in I$, $L^-(t_*) = 1$, i.e., $L(t_*) = \infty$. But, if $r_* = \max\{r : H_r(t_*) > 0\}$, $L_{r_*}^-(t_*) = 1$, i.e., $L_{r_*}(t_*) = \infty$, and the equality still holds. Note also that $F_{r_*}((t,\infty]) = 0$ for all $t \ge t_*$.
Similarly, we have for $t \in I$

$$\sum_{1\le \ell<r} L_\ell(t) = -\ln \frac{V_r(t)}{V_1(t)}$$

and therefore, for $t < t_*$ and for $t = t_*$ and $r \leq r_*$ if $H(t_*) > 0$,

$$
\begin{aligned}
-\ln V_r(t) &= -\ln H([t,\infty]) + \sum_{1 \leq \ell < r} L_\ell(t) \\
&= \sum_{1 \leq \ell \leq g} L_\ell([0,t)) + \sum_{1 \leq \ell < r} L_\ell(t) \\
&= \sum_{1 \leq \ell < r} L_\ell([0,t]) + \sum_{r \leq \ell \leq g} L_\ell([0,t))
\end{aligned}
$$

and this is equivalent to,

$$V_r(t) = \prod_{1 \leq \ell < r} F_\ell((t,\infty]) \prod_{r \leq \ell \leq g} F_\ell([t,\infty]).$$

Finally, for $t > t_*$ and for $t = t_*$ and $r > r_*$ if $H(t_*) > 0$, both members of the last equality are equal to zero.

A.2 Proof of Lemma 4.5

Let $N_n = n\,\tilde{H}_n$ be the empirical counting measure of the sample $\{(Y_i, A_i) : 1 \leq i \leq n\}$ and

$$C_t^I = (t,\infty] \times I^c \cup [t,\infty] \times I.$$

Then, $L_n^-(f)$ may be written as

$$L_n^-(f) = \sum_{1 \leq i \leq n} f(Y_i)\, \mathbb{1}_{\{A_i \in J\}}\, N_n(C_{Y_i}^I)^{-1}.$$

Now,

$$\mathcal{B}_{n+1} = \sigma(\tilde{H}_{n+1}) \vee \bigvee_{n+1 < m < \infty} \sigma(Y_m, A_m)$$

and therefore, by the i.i.d. property of the sample and elementary properties of conditional independence (see, e.g., Florens, Mouchart, Rolin (1990)),

$$E[L_n^-(f) \mid \mathcal{B}_{n+1}] = n\,E[f(Y_1)\, \mathbb{1}_{\{A_1 \in J\}}\, N_n(C_{Y_1}^I)^{-1} \mid \tilde{H}_{n+1}].$$

The σ-algebra generated by $\tilde{H}_{n+1}$ is the sub-σ -algebra of permutable events in the σ-algebra generated by $\{(Y_i, A_i) : 1 \leq i \leq n+1\}$. Hence,

$$E[L_n^-(f) \mid \mathcal{B}_{n+1}] = \frac{n}{(n+1)!} \sum_{\sigma \in P_{n+1}} f(Y_{\sigma(1)})\, \mathbb{1}_{\{A_{\sigma(1)} \in J\}}\, N_n^\sigma(C_{Y_{\sigma(1)}}^I)^{-1}$$

where P_{n+1} is the set of permutations of $n+1$ elements and

$$N_n^\sigma\,(\,C^I_{Y_{\sigma(1)}}\,)\;=\;N_{n+1}\,(\,C^I_{Y_{\sigma(1)}}\,)\;-\;\mathbb{1}_{\{\,(\,Y_{\sigma(n+1)}\,,\,A_{\sigma(n+1)}\,)\,\in\,C^I_{Y_{\sigma(1)}}\,\}}\,.$$

There are $(n-1)!$ permutations such that $\sigma(1)\;=\;i$ and $\sigma(n+1)\;=\;j$ and therefore,

$$E\,[\,L_n^-\,(\,f\,)\;|\;\mathcal{B}_{n+1}\,]\;=\;\frac{1}{n+1}\sum_{1\,\le\,i\,\le\,n+1}f\,(\,Y_i\,)\;\mathbb{1}_{\{\,A_i\,\in\,J\,\}}$$

$$\times\sum_{1\,\le\,j\,\ne\,i\,\le\,n+1}[\,N_{n+1}\,(\,C^I_{Y_i}\,)\;-\;\mathbb{1}_{\{\,(\,Y_j\,,\,A_j\,)\,\in\,C^I_{Y_i}\,\}}\,]^{-1}\,.$$

Now,

$$\sum_{1\,\le\,j\,\ne\,i\,\le\,n+1}[\,N_{n+1}\,(\,C^I_{Y_i}\,)\;-\;\mathbb{1}_{\{\,(\,Y_j\,,\,A_j\,)\,\in\,C^I_{Y_i}\,\}}\,]^{-1}$$

$$=\;\sum_{1\,\le\,j\,\ne\,i\,\le\,n+1}\{\frac{\mathbb{1}_{\{\,(\,Y_j\,,\,A_j\,)\,\notin\,C^I_{Y_i}\,\}}}{N_{n+1}\,(\,C^I_{Y_i}\,)}\;+\;\frac{\mathbb{1}_{\{\,(\,Y_j\,,\,A_j\,)\,\in\,C^I_{Y_i}\,\}}}{N_{n+1}\,(\,C^I_{Y_i}\,)\;-\;1}\}$$

$$=\;\frac{(n+1)\;-\;N_{n+1}\,(\,C^I_{Y_i}\,)}{N_{n+1}\,(\,C^I_{Y_i}\,)}\;+\;\mathbb{1}_{\{\,N_{n+1}\,(\,C^I_{Y_i}\,)\,>\,1\,\}}$$

$$=\;\frac{n+1}{N_{n+1}\,(\,C^I_{Y_i}\,)}\;-\;\mathbb{1}_{\{\,N_{n+1}\,(\,C^I_{Y_i}\,)\,=\,1\,\}}\,.$$

Note that $(\,Y_i\,,\,A_i\,)\;\in\;C^I_{Y_i}$, since $A_i\;\in\;J\;\subset\;I$. Therefore,

$$E\,[\,L_n^-\,(\,f\,)\;|\;\mathcal{B}_{n+1}\,]$$

$$=\;\sum_{1\,\le\,i\,\le\,n+1}f\,(\,Y_i\,)\;\mathbb{1}_{\{\,A_i\,\in\,B\,\}}\;N_{n+1}\,(\,C^I_{Y_i}\,)^{-1}\;-\;R\;\le\;L^-_{n+1}\,(\,f\,)$$

where

$$R\;=\;\frac{f\,(\,Z_M\,)}{n+1}\;\mathbb{1}_{\{\,N_{n+1}\,(\,\{Z_M\}\,\times\,J\,)\,=\,1\,,\,N_{n+1}\,(\,\{Z_M\}\,\times\,(\,I\,-\,J\,)\,)\,=\,0\,\}}$$

and

$$Z_M\;=\;\sup_{1\,\le\,i\,\le\,n+1}\;Y_i\,.$$

A.3 Proof of Lemma 4.6

The empirical predictable hazard measure has the following integral form :

$$L_n^-((t,t+s]) = \int_{(t,t+s]} N_n(C_u^I)^{-1} N_n(du \times J)$$

that may be approximated as follows :

Let $t_{m,k} = t + \frac{k}{2^m} s$, for $0 \leq k \leq 2^m$,

$$C_{m,k}^I = (t_{m,k}, \infty] \times I^c \cup (t_{m,k-1}, \infty] \times I$$

and

$$S_m(u) = \sum_{1\leq k\leq 2^m} \frac{1\!\!I_{\{N_n(C_{m,k}^I)>0\}}}{N_n(C_{m,k}^I)} 1\!\!I_{(t_{m,k-1},t_{m,k}]}(u).$$

Clearly,

$$\lim_{m\to\infty} S_m(u) = \frac{1\!\!I_{\{N_n(C_u^I)>0\}}}{N_n(C_u^I)} 1\!\!I_{(t,t+s]}(u)$$

and since $S_m(u) \leq 1$, by the dominated convergence theorem,

$$L_n^-((t,t+s]) = \lim_{m\to\infty} \int_{(t,t+s]} S_m(u) N_n(du \times J)$$

$$= \lim_{m\to\infty} \sum_{1\leq k\leq 2^m} \frac{N_n((t_{m,k-1}, t_{m,k}] \times J)}{N_n(C_{m,k}^I)} 1\!\!I_{\{N_n(C_{m,k}^I)>0\}}.$$

By a well-known property of the multinomial distribution,

$$N_n((t_{m,k-1}, t_{m,k}] \times J) \mid N_n(C_{m,k}^I)$$
$$\sim Bi[N_n(C_{m,k}^I), \tilde{H}((t_{m,k-1}, t_{m,k}] \times J) \tilde{H}(C_{m,k}^I)^{-1}].$$

Therefore,

$$E[N_n((t_{m,k-1}, t_{m,k}] \times J) N_n(C_{m,k}^I)^{-1} \mid N_n(C_{m,k}^I)]$$
$$= \tilde{H}((t_{m,k-1}, t_{m,k}] \times J) \tilde{H}(C_{m,k}^I)^{-1} \times 1\!\!I_{\{N_n(C_{m,k}^I)>0\}}.$$

Now, if we define

$$T_m(u) = \sum_{1\leq k\leq 2^m} \frac{1\!\!I_{\{N_n(C_{m,k}^I)>0\}}}{\tilde{H}(C_{m,k}^I)} 1\!\!I_{(t_{m,k-1},t_{m,k}]}(u)$$

then, clearly,

$$\lim_{m\to\infty} T_m(u) = \frac{1\!\!1_{\{N_n(C_u^I)>0\}}}{\tilde H(C_u^I)}\, 1\!\!1_{(t,t+s]}(u)$$

and since $C_{t+s}^I \subset C_{m,k}^I$, $T_m(u) \le \tilde H(C_{t+s}^I)^{-1}$. Therefore, for $t+s < t_*$ or $t+s = t_*$ and $H_I(t_*) > 0$, by the dominated convergence theorem,

$$\begin{aligned}\lim_{m\to\infty} \sum_{1\le k\le 2^m} &\frac{\tilde H(\,(t_{m,k-1},t_{m,k}]\times J\,)}{\tilde H(C_{m,k}^I)}\, 1\!\!1_{\{N_n(C_{m,k}^I)>0\}} \\ &= \lim_{m\to\infty}\int_{(t,t+s]} T_m(u)\,\tilde H(du\times J) \\ = \int_{(t,t+s]} &1\!\!1_{\{N_n(C_u^I)>0\}}\, L^-(du) \doteq A_n(\,(t,t+s]\,)\,.\end{aligned}$$

Therefore, for $t+s < t_*$ or $t+s = t_*$ and $H_I(t_*) > 0$, by the dominated convergence theorem with respect to $P_{\tilde H}$,

$$E[\,L_n^-(\,(t,t+s]\,)\,] = E[\,A_n(\,(t,t+s]\,)\,]$$

Now for $t+s \ge t_*$, $L_n^-(\,(t,t+s]\,) = L_n^-(\,(t,t_*]\,)$ if $H_J(t_*) > 0$ and $L_n^-(\,(t,t+s]\,) = L_n^-(\,(t,t_*)\,)$ if $H_J(t_*) = 0$ and the same is true for A_n. Hence by the monotone convergence theorem (in letting $t+s$ increase to t_*) and the monotone class theorem, we obtain the desired result, i.e.,

$$E[\,L_n^-(f)\,] = E[\,A_n(f)\,]\,.$$

It is not too difficult to obtain in a similar way an interesting martingale property. Let us recall that the counting measure N_n is a multinomial measure with parameters n and $\tilde H$. Now, for any $D \in \mathcal{B}^+ \bigotimes 2^K$, we denote by N_{nD} the trace of N_n on D and by $\mathcal{N}_n(D)$ the σ-algebra generated by N_{nD}. Clearly, if $D_1 \subset D_2$, $\mathcal{N}_n(D_1) \subset \mathcal{N}_n(D_2)$. Now, for the multinomial measure, we have the following conditional property : conditionally on $N_n(D)$, $\mathcal{N}_n(D)$ and $\mathcal{N}_n(D^c)$ are independent σ-algebras and N_{nD} is a multinomial measure with parameters $N_n(D)$ and $\tilde H(\,.\mid D\,)$.
Now, let

$$\mathcal{M}_t^I = \mathcal{N}_n(D_t^I)$$

where

$$D_t^I = [0,\infty)\times K - C_t^I = [0,t]\times I^c \cup [0,t)\times I\,.$$

Clearly, $\{ \mathcal{M}_t^I \, : \, t \, \in \, \mathbb{R}^+ \}$ is a filtration and

$$\mathcal{M}_{t+}^I \, = \, \mathcal{M}_{t+} \, = \, \mathcal{N}_n \, (\, [0,t] \times K \,) \, .$$

Now, for $1 \leq k \leq 2^m$, $[0,t] \times K \, \subset \, D_{m,k}^I \, = \, [0, t_{m,k}] \times I^c \cup [0, t_{m,k-1}] \times I$ and $N_n \, (\, (t_{m,k-1}, t_{m,k}] \times J \,) \, N_n \, (\, C_{m,k}^I \,)^{-1}$ is $\mathcal{N}_n \, (\, C_{m,k}^I \,)$-measurable. Therefore, by the conditional property of the multinomial measure,

$$\begin{aligned}
& E \, [\, N_n \, (\, (t_{m,k-1}, t_{m,k}] \times J \,) \, N_n \, (\, C_{m,k}^I \,)^{-1} \, \mid \, \mathcal{M}_{t+} \,] \\
= \; & E \, [E \, [\, N_n \, (\, (t_{m,k-1}, t_{m,k}] \times J \,) \, N_n \, (\, C_{m,k}^I \,)^{-1} \, \mid \, \mathcal{N}_n \, (\, D_{m,k}^I \,) \,] \, \mid \, \mathcal{M}_{t+} \,] \\
= \; & E \, [\, E \, [\, N_n \, (\, (t_{m,k-1}, t_{m,k}] \times J \,) \, N_n \, (\, C_{m,k}^I \,)^{-1} \, \mid \, N_n \, (\, C_{m,k}^I \,) \,] \, \mid \, \mathcal{M}_{t+} \,] \\
= \; & E \, [\, \tilde{H} \, (\, t_{m,k-1}, t_{m,k}] \times J \,) \, \tilde{H} \, (\, C_{m,k}^I \,)^{-1} \, \times \, \mathbb{1}_{\{ \, N_n \, (\, C_{m,k}^I \,) > 0 \, \}} \, \mid \, \mathcal{M}_{t+} \,] \, .
\end{aligned}$$

Using the same arguments as above, we obtain

$$E \, [\, L_n^- \, (\, f \, \mathbb{1}_{(t,\infty \,]} \,) \, \mid \, \mathcal{M}_{t+} \,] \, = \, E \, [\, A_n \, (\, f \, \mathbb{1}_{(t,\infty \,]} \,) \, \mid \, \mathcal{M}_{t+} \,] \, .$$

Now, it is easy to see that $L_n^- \, (\, [0,t] \,)$ is $\mathcal{M}_{t+}$-measurable and that $A_n \, (\, [0,t] \,)$ is $\mathcal{M}_t^I$-measurable. and therefore, $L_n^- - A_n$ is a martingale for the filtration $\{ \, \mathcal{M}_{t+} \, , \, t \, \in \, \mathbb{R}^+ \, \}$. Pursuing the martingale analysis would then require a definition of "predictability" with respect to the filtration $\{ \, \mathcal{M}_t^I \, , \, t \, \in \, \mathbb{R}^+ \, \}$. Heuristically, the above martingale property may be infinitesimally expressed as

$$E \, [\, L_n^- \, (\, dt \,) \, \mid \, \mathcal{M}_t^I \,] \, = \, A_n \, (\, dt \,) \, .$$

A.4 Proof of Lemma 4.12

The integral of f with respect to G_n may be written as

$$G_n \, (\, f \,) \, = \, \sum_{1 \, \leq \, i \, \leq \, n} \, f \, (\, Y_i \,) \, \mathbb{1}_{\{ \, A_i \, \in \, J_1 \, \}} \, F_{n,2} \, (\, [Y_i, \infty] \,) \, N_n \, (\, C_{Y_i}^{I_1} \,)^{-1}$$

where

$$F_{n,2} \, (\, [Y_i, \infty] \,) \, = \, \prod_{0 \leq u < Y_i} \, \{ 1 \, - \, L_{n,2}^- \, (u) \, \} \, = \, \prod_{0 \leq u < Y_i} \, \frac{N_n \, (\, C_u^{I_2 \, - \, J_2} \,)}{N_n \, (\, C_u^{I_2} \,)} \, .$$

By the same arguments as in the proof of Lemma 4.5, we obtain

$$E\,[\,G_n\,(\,f\,)\;|\;\mathcal{B}_{n+1}\,] \;=\; \frac{1}{n+1}\sum_{1\,\le\, i\,\le\, n+1} f\,(\,Y_i\,)\;\mathbb{1}_{\{\,A_i\,\in\, J_1\,\}}$$

$$\times \sum_{1\,\le\, j\,\ne\, i\,\le\, n+1} N_n^j\,(\,C_{Y_i}^{I_1}\,)^{-1}\;F_{n,2}^j\,(\,[Y_i,\infty]\,)$$

where

$$F_{n,2}^j\,(\,[Y_i,\infty]\,) \;=\; \prod_{0\le u<Y_i}\{1\;-\;L_{n,2}^{j\,-}\,(u)\,\} \;=\; \prod_{0\le u<Y_i}\frac{N_n^j\,(\,C_u^{I_2\,-\,J_2}\,)}{N_n^j\,(\,C_u^{I_2}\,)}$$

and

$$N_n^j\,(\,C_u^I\,) \;=\; N_{n+1}\,(\,C_u^I\,) \;-\; \mathbb{1}_{\{\,(Y_j,A_j)\,\in\, C_u^I\,\}}\,.$$

Now, if $Y_j \ge Y_i$,

$$N_n^j\,(\,C_{Y_i}^{I_1}\,) \;=\; N_{n+1}\,(\,C_{Y_i}^{I_1}\,) \;-\; \mathbb{1}_{\{\,Y_j\,>\,Y_i\,\}} \;-\; \mathbb{1}_{\{\,Y_j\,=\,Y_i\,,\,A_j\,\in\, I_1\,\}}$$

$$F_{n,2}^j\,(\,[Y_i,\infty]\,) \;=\; \prod_{0\le u<Y_i}\frac{N_{n+1}\,(\,C_u^{I_2\,-\,J_2}\,)\;-\;1}{N_{n+1}\,(\,C_u^{I_2}\,)\;-\;1} \;\doteq\; \tilde{F}_{n,2}\,(\,[Y_i,\infty]\,)$$

and if $Y_j < Y_i$,

$$N_n^j\,(\,C_{Y_i}^{I_1}\,) \;=\; N_{n+1}\,(\,C_{Y_i}^{I_1}\,)$$

$$F_{n,2}^j\,(\,[Y_i,\infty]\,) \;=\; \tilde{F}_{n,2}\,(\,[Y_j,\infty]\,)\;\times\;\prod_{Y_j<u<Y_i}\frac{N_{n+1}\,(\,C_u^{I_2\,-\,J_2}\,)}{N_{n+1}\,(\,C_u^{I_2}\,)}$$

$$\times\;\frac{N_{n+1}\,(\,C_{Y_j}^{I_2\,-\,J_2}\,)\;-\;\mathbb{1}_{\{\,A_j\,\in\, I_2\,-\,J_2\,\}}}{N_{n+1}\,(\,C_{Y_j}^{I_2}\,)\;-\;\mathbb{1}_{\{\,A_j\,\in\, I_2\,\}}}$$

Therefore, on one hand,

$$\sum_{1\,\le\, j\,\ne\, i\,\le\, n+1} N_n^j\,(\,C_{Y_i}^{I_1}\,)^{-1}\;F_{n,2}^j\,(\,[Y_i,\infty]\,)\;\mathbb{1}_{\{\,Y_j\,\ge\, Y_i\,\}}$$

$$=\;\tilde{F}_{n,2}\,(\,[Y_i,\infty]\,)\sum_{1\,\le\, j\,\ne\, i\,\le\, n+1}\frac{\mathbb{1}_{\{\,(\,Y_j\,,\,A_j\,)\,\in\, C_{Y_i}^{I_1}\,\}}}{N_{n+1}\,(\,C_{Y_i}^{I_1}\,)\;-\;1}\;+\;\frac{\mathbb{1}_{\{\,Y_j\,=\,Y_i\,,\,A_j\,\in\, I_1^c\,\}}}{N_{n+1}\,(\,C_{Y_i}^{I_1}\,)}$$

$$= \tilde{F}_{n,2}([Y_i,\infty]) \times \left(1\!\!1_{\{N_{n+1}(C_{Y_i}^{I_1})>1\}} + \frac{N_{n+1}(\{Y_i\} \times I_1^c)}{N_{n+1}(C_{Y_i}^{I_1})} \right)$$

$$= \tilde{F}_{n,2}([Y_i,\infty]) \frac{N_{n+1}([Y_i,\infty] \times K) - 1\!\!1_{\{N_{n+1}(C_{Y_i}^{I_1})=1\}}}{N_{n+1}(C_{Y_i}^{I_1})}$$

because $A_i \in J_1 \subset I_1$.

On the other hand, if $Y_i = Z_k$ and $Y_j = Z_\ell$ with $\ell < k$ (recall that $\{Z_k, 1 \le k \le M\}$ is the order statistics of the sample),

$$\sum_{1 \le j \le n+1} \frac{N_{n+1}(C_{Y_j}^{I_2 - J_2}) - 1\!\!1_{\{A_j \in I_2 - J_2\}}}{N_{n+1}(C_{Y_j}^{I_2}) - 1\!\!1_{\{A_j \in I_2\}}} 1\!\!1_{\{Y_j = Z_\ell\}}$$

$$= \frac{N_{n+1}(C_{Z_\ell}^{I_2 - J_2})}{N_{n+1}(C_{Z_\ell}^{I_2})} N_{n+1}(\{Z_\ell\} \times I_2^c) + \frac{N_{n+1}(C_{Z_\ell}^{I_2 - J_2})}{N_{n+1}(C_{Z_\ell}^{I_2}) - 1}$$

$$\times N_{n+1}(\{Z_\ell\} \times J_2) + \frac{N_{n+1}(C_{Z_\ell}^{I_2 - J_2}) - 1}{N_{n+1}(C_{Z_\ell}^{I_2}) - 1} N_{n+1}(\{Z_\ell\} \times (I_2 - J_2)).$$

Noticing that

$$N_{n+1}(\{Z_\ell\} \times I_2^c) = N_{n+1}([Z_\ell, \infty] \times K) - N_{n+1}(C_{Z_\ell}^{I_2}),$$

$$N_{n+1}(\{Z_\ell\} \times J_2) = N_{n+1}(C_{Z_\ell}^{I_2}) - N_{n+1}(C_{Z_\ell}^{I_2 - J_2}),$$

$$N_{n+1}(\{Z_\ell\} \times (I_2 - J_2)) = N_{n+1}(C_{Z_\ell}^{I_2 - J_2}) - N_{n+1}((Z_\ell, \infty] \times K),$$

we obtain after some simplifications,

$$\sum_{1 \le j \le n+1} \frac{N_{n+1}(C_{Y_j}^{I_2 - J_2}) - 1\!\!1_{\{A_j \in I_2 - J_2\}}}{N_{n+1}(C_{Y_j}^{I_2}) - 1\!\!1_{\{A_j \in I_2\}}} 1\!\!1_{\{Y_j = Z_\ell\}}$$

$$= \frac{N_{n+1}(C_{Z_\ell}^{I_2 - J_2})}{N_{n+1}(C_{Z_\ell}^{I_2})} N_{n+1}([Z_\ell, \infty] \times K)$$

$$- \frac{N_{n+1}(C_{Z_\ell}^{I_2 - J_2}) - 1}{N_{n+1}(C_{Z_\ell}^{I_2}) - 1} N_{n+1}([Z_{\ell+1}, \infty] \times K).$$

Therefore, if

$$G_\ell = N_{n+1}([Z_\ell, \infty] \times K) \; \tilde{F}_{n,2}([Z_\ell,\infty]) \prod_{Z_\ell \le u < Z_k} \frac{N_{n+1}(C_u^{I_2 - J_2})}{N_{n+1}(C_u^{I_2})},$$

we can see that

$$\sum_{1 \le j \le n+1} N_n^j \left(C_{Y_i}^{I_1} \right)^{-1} F_{n,2}^j \left([Y_i, \infty] \right) 1\!\!I_{\{ Y_j < Y_i \}}$$

$$= \frac{1}{N_{n+1} \left(C_{Y_i}^{I_1} \right)} \sum_{1 \le \ell < k} \left(G_\ell - G_{\ell+1} \right)$$

$$= \frac{n+1}{N_{n+1} \left(C_{Y_i}^{I_1} \right)} F_{n+1,2} \left([Y_i, \infty] \right) - \frac{N_{n+1} \left([Y_i, \infty] \times K \right)}{N_{n+1} \left(C_{Y_i}^{I_1} \right)} \tilde{F}_{n,2} \left([Y_i, \infty] \right).$$

Hence,

$$\sum_{1 \le j \ne i \le n+1} N_n^j \left(C_{Y_i}^{I_1} \right)^{-1} F_{n2}^j \left([Y_i, \infty] \right)$$

$$= (n+1) \frac{F_{n+1,2} \left([Y_i, \infty] \right)}{N_{n+1} \left(C_{Y_i}^{I_1} \right)} - \tilde{F}_{n,2} \left([Y_i, \infty] \right) 1\!\!I_{\{ N_{n+1} (C_{Y_i}^{I_1}) = 1 \}}.$$

Finally,

$$E \left[G_n(f) \mid \mathcal{B}_{n+1} \right] = G_{n+1}(f) - R \le G_{n+1}(f)$$

where

$$R = \frac{f(Z_M)}{n+1} \tilde{F}_{n,2} \left([Z_M, \infty] \right) 1\!\!I_{\{ N_{n+1} (\{ Z_M \} \times J_1) = 1 \, , \, N_{n+1} (\{ Z_M \} \times (I_1 - J_1)) = 0 \}}.$$

References

[1] Andersen P.K., O. Borgan, R.D. Gill, N. Keiding (1993), *Statistical Models Based on Counting Processes*, Springer, New York.

[2] Crowder, M. (1994), Identifiability Crises in Competing Risks, *International Statistical Review*, **62**, 379-391.

[3] Dellacherie C., P.A. Meyer (1980), *Probabilité et potentiel (B), Théorie des martingales*, Hermann, Paris.

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

[5] Florens J.P., M. Mouchart, J.M. Rolin (1990), *Elements of Bayesian Statistics*, Marcel Dekker, New York.

[6] Florens J.P., M. Mouchart, J.M. Rolin (1999), Semi- and Nonparametric Bayesian analysis of Duration Models : A Survey, *International Statistical Review*, **67**, 187-210.

[7] Gill R. D. (1993) Lectures on Survival Analysis in *Ecole d'Eté de Probabilités de Saint-Flour XXII*, ed. Bernard P., Lecture Notes in Mathematics, Springer.

[8] Jacod J. (1975), Multivariate point processes: Predictable projection, Radon-Nikodym derivatives, representation of martingales, *Z. Wahrsch. verw. Geb.* **31**, 235-253.

[9] Kalbfleisch J. D., R. J. Mac Kay (1979), On constant-sum models for censored survival data, *Biometrika*, **66**, 87-90.

[10] Kaplan E. L., P. Meier (1958), Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, **53**, 457-481.

[11] Mouchart M., J.M. Rolin (1995), Competing risks models: problems of modelling and of identification, Discussion Paper 9507, Institut de Statistique, Université catholique de Louvain, Belgium.

[12] Rolin J.M. (1998), Bayesian Survival Analysis, in *Encyclopedia of Biostatistics*, Wiley, **1**, 271-286.

[13] Stute W., J. L. Wang (1993), The strong law under random censorship, *Annals of Statistics*, **21**, 1591-1607.

[14] Tsai W.Y., (1986), Estimation of survival curves from dependent censorship models via a generalized self-consistent property with nonparametric Bayesian estimation application, *Annals of Statistics*, **14**, 238-249.

[15] Williams J.S., S.W. Lagakos (1977), Models for censored survival analysis : Constant-sum and variable-sum models, *Biometrika*, **64**, 215-224.