Applied Mathematics, 2014, 5, 765-776
Published Online March 2014 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2014.54073

Scientific Research

Nelson-Aalen and Kaplan-Meier Estimators in Competing Risks

Didier Alain Njamen-Njomen[1,2], Joseph Ngatchou-Wandji[3]

[1]Department of Mathematics and Computer Sciences, Faculty of Sciences, University of Maroua, Maroua, Cameroon
[2]Department of Mathematics, Faculty of Sciences, University of Yaounde 1, Yaounde, Cameroon
[3]University of Lorraine, Lorraine, France
Email: dangaza@yahoo.fr, didiernjamen1@gmail.com, Joseph.ngatchou-wandji@univ-lorraine.fr

Received 11 December 2013; revised 11 January 2014; accepted 18 January 2014

Abstract

In this paper, stochastic processes developed by Aalen [1] [2] are adapted to the Nelson-Aalen and Kaplan-Meier [3] estimators in a context of competing risks. We focus only on the probability distributions of complete downtime individuals whose causes are known and which bring us to consider a partition of individuals into sub-groups for each cause. We then study the asymptotic properties of nonparametric estimators obtained.

Keywords

Censored Data; Counting Process; Competitive Risk; Non-Parametric Estimators; The Cumulative Incidence Function; Risk Function Specific Cause; Conditional Distribution Function

1. Introduction

Let us consider a data model which lives time where the event of interest is a failure (or death) due to the j^{th} event, $j \in J = \{1, 2, \cdots, m\}$, and the non-zero integer *m*, the number of possible causes. By convention, $j = 0$ corresponds to the state of functioning (or of life) of the observed individual. It is assumed that the observation is stopped when a failure (or death) occurs, but this observation may be right-censored in a non-informative way. Some examples of this situation corresponds to the case where the event of interest is due to another cause, or withdrawal of the individual from the study or at the end of the study. In the case of right censoring time, the time of failure of year for individuals and their causes are not known to the experimenter. A data model as described above is commonly called "competing risks model" (or competitors) and is studied in fields such as medical control, demography, actuarial science, economics or industrial reliability. In Andersen *et al*. [4], an illustration and details of mathematics techniques on competing risks in biomedical applications are developed. For example in the study of AIDS, the different competitive risks can be 1) death due to AIDS, 2) death due to tuberculosis or 3) death due to other causes and in this case $m = 3$ (see **Figure 1**).

It is important to note that in most data models in competing risks, the functions that characterize the probability distribution of the variable of interest and the marginal are not always observable (see Tsiatis [5], Heckman and Honoré [6]). Issues to be resolved include virtually the underlying functions for different causes and effects of covariates on the rate of occurrence of competing risks. One of the problems we may face is that the information on the cause of failure of the individual observation can only be known after the autopsy, while we don't know anything about individuals censored in monitoring. In addition, the incident distributions (due to specific causes) do not allow to describe satisfactorily the probabilities of the various marginal (failures τ case j) in competing risks models. Assumptions of independence of competing risks can help ensure observability in some cases, but they are not reasonable only in such models.

1.1. Related Works

The estimators of Nelson-Aalen and Kaplan-Meier [3] are generally studied in the literature following two approaches: firstly, the method of martingale (Aalen [1] [2]; Andersen *et al.* [4]; Fleming and Harrington [7], Prentice *et al.* [8]) and secondly the law of the iterated logarithm (Breslow and Crowley [9], Földes and Rejtö [10] or Major and Rejtö [11], Földes and Rejtö [12], Gill [13], Csörgö and Horváth [14], Ying [15] and Chen and Lo [16]). Recently, applications have been made in the context of competing risks (Latouche [17]; Belot [18]). Latouche [17] states that during the planification of clinical trials, the evaluation of the number of patients to be included is a critical issue because such a formulation does not exist in the Fine and Gray's [19] model. For this purpose, he therefore computes the number of patients within the context of competition for an inference based function on cumulative incidence and then, he studies the properties of the model of Fine and Gray when it is wrongly specified. Belot [18] presents the data got from randomized clinical tests on prostate cancer patients who died for several reasons.

1.2. Contributions

In this paper, the stochastic processes developed by Aalen [1] [2] are adapted to Nelson-Aalen and Kaplan-Meier estimators [3] in a context of competing risks (e.g. Aalen and Johansen [20], Andersen *et al.* [4]). We focus only on the complete probability distributions of downtime individuals whose causes are known and which bring us to consider a partition of individuals sub-groups for each cause. We provide a new proof of the consistency of the Nelson-Aalen estimator in the context of competing risks by using the method of martingale. Under the regularity assumptions for the sequence (k_n) ((k_n) is a sequence of integers such that $1 \le k_n < n$ and n is the number of observable samples) we obtain an almost-safe speed estimator of Kaplan-Meier [3], which is the same as that obtained by Giné and Guillou [21] which is $\mathcal{O}\left(\left(\log\log n/k_n\right)^{1/2}\right)$.

The rest of the paper is organized as follows: Section 2 describes preliminary results and notations used in the paper and Section 3 evaluates the conditional functions of distribution to the specific causes. Section 4 contains the main results of the paper as well as some properties of our estimators obtained. The last section concludes the paper.

2. Preliminary Results

Lifetime analysis (also referred to as survival analysis) is the area of statistics that focuses on analyzing the time

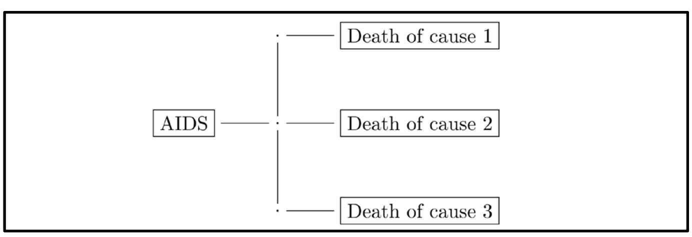

Figure 1. Example of 3 risks competing model.

duration between a given starting point and a specific event. This endpoint is often called failure and the corresponding length of time is called the failure time or survival time or lifetime.

Formally, a failure time is a nonnegative random variable (r.v.) X that describes the length of time from a time origin until an event of interest occurs. We will suppose throughout that $\mathbb{P}[X<\infty]=1$.

The most basic quantities used to summarize and describe the time elapsed from a starting point until the occurrence of an event of interest are the distribution function and the hazard function. The cumulative distribution function at time t, also called lifetime distribution or the failure distribution, is the probability that the failure time of an individual is less or equal than the value t. It is given for $t\geq 0$ by: $F(t)=\mathbb{P}(X\leq t)$.

The function F is right-continuous, nondecreasing and satisfies $F(0)=0$ and $F(\infty)=1$. We denote by F^- the left-continuous function obtained from F in the following way:

$$F^-(t)=\lim_{u\uparrow t}F(u).$$

The distribution of X may equivalently be dealt with in terms of the survival function which is given, for $t\geq 0$, by:

$$1-F(t)=\mathbb{P}[X>t].$$

The cumulative hazard function is defined for $t\geq 0$ by:

$$\Lambda(t)=\int_0^t \frac{\mathrm{d}F}{1-F^-}.$$

When F is continuous, the relation $1-F(t)=\exp(-\Lambda(t))$ is valid for all $t\geq 0$. We can then call Λ the log-survival function.

If F admits a derivative with respect to Lebesgue measure on $\mathbb{R}$, the probability density function exists and is defined for $t\geq 0$ by:

$$f(t)=\frac{\mathrm{d}F(t)}{\mathrm{d}t}=\lim_{h\to 0}\frac{\mathbb{P}[t\leq X<t+h]}{h}.$$

Heuristically, the function f may be seen as the instantaneous probability of experiencing the event.

With the same hypothesis of differentiability, the hazard function exists and is defined for $t\geq 0$ by:

$$\lambda(t)=\frac{f(t)}{1-F^-(t)}=\lim_{h\to 0}\frac{1}{h}\frac{\mathbb{P}[t\leq X<t+h]}{\mathbb{P}[X\geq t]}=\lim_{h\to 0}\mathbb{P}[t\leq X<t+h\mid X\geq t].$$

The quantity $\lambda(t)$ can be interpreted as the instantaneous probability that an individual dies at time t, conditionally on he or she having survived until that time.

For an extensive introduction to lifetime analysis, the reader is referred e.g. to the books of Cox and Oakes [22] and Kalbfleisch and Prentice [23].

The main difficulty in the analysis of lifetime data lies in the fact that the actual failure times of some individuals may not be observed. An observation is right-censored if it is known to be greater than a certain value, provided the exact time is unknown. Let C be the nonnegative r.v. with distribution function G that stands for the censoring time of the individual. As before, the nonnegative r.v. X with distribution function F denotes the failure time of the individual. If X is censored, instead of X, we observe C which gives the information that X is greater than C. In any case, the observable r.v. consists of $T=\min(X,C)$, $D=\mathbb{1}_{\{X\leq C\}}$, where $\mathbb{1}_{\{.\}}$ denotes the indicator function. The nonnegative r.v. T stands for the observed duration of time which may correspond either to the event of interest $(D=1)$ or to a censoring time $(D=0)$.

As a sequel to above, it is assumed that X and C are independent. Consequently, the random variable T has the distribution function H given by

$$1-H=(1-F)(1-G).$$

The following subdistribution functions of H will be needed:

$$H^{(0)}(t)=\mathbb{P}[T\leq t,D=0]$$

and

$$H^{(1)}(t)=\mathbb{P}[T\leq t,D=1]$$

The relation

$$H(t) = H^{(0)}(t) + H^{(1)}(t)$$

is valid for any $t \geq 0$.

The relations that connect the subdistribution functions $H^{(0)}$, $H^{(1)}$ and to the distribution functions F and G are given by:

$$H^{(0)}(t) = \int_0^1 (1-F)\,\mathrm{d}G$$

and

$$H^{(1)}(t) = \int_0^1 (1-G^-)\,\mathrm{d}F.$$

The cumulative hazard function of X can be expressed as:

$$\Lambda(t) = \int_0^t \frac{\mathrm{d}F}{1-F^-} = \int_0^t \frac{\mathrm{d}H^{(1)}}{1-H^-}.$$

Kaplan and Meier [3] introduced the product-limit estimator for the survival distribution function. The estimator of the cumulative hazard function is the Nelson-Aalen estimator introduced by Nelson [24] [25] and generalized by Aalen [1] [2].

Let (T_i, D_i) for $i = 1, \cdots, n$ be n independent copies of the random vector (T, D). Let $T_{1,n} \leq T_{2,n} \leq \cdots \leq T_{n,n}$ be the order statistics associated to the sample $T_1, \cdots, T_n$. If there are ties between a failure time (or several failure times) and a censoring time, then the failure time(s) is (are) ranked ahead of the censoring time(s).

We define the empirical counterparts of $H^{(0)}$, $H^{(1)}$ and H, by:

$$H_n^{(0)}(t) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{T_i \leq t,\ D_i = 0\}},$$

$$H_n^{(1)}(t) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{T_i \leq t,\ D_i = 1\}},$$

$$H_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{T_i \leq t,\}}.$$

The Kaplan-Meier product-limit estimator is defined for $t \geq 0$ by:

$$\widehat{F}_n(t) = 1 - \prod_{i=1}^n \left(1 - \frac{\mathbb{1}_{\{T_i \leq t,\ D_i = 1\}}}{n(1 - H_n^-(T_i))}\right).$$

The Nelson-Aalen estimator for Λ is then defined for $t \geq 0$ by:

$$\Lambda_n(t) = \int_0^n \frac{\mathrm{d}H_n^{(1)}}{1-H_n^-}.$$

The following relations are valid for $t \geq 0$:

$$H_n(t) = H_n^{(1)}(t) + H_n^{(0)}(t),$$

$$1 - H_n(t) = \left(1 - \hat{F}_n(t)\right)\left(1 - \hat{G}_n(t)\right),$$

$$\Lambda_n(t) = \int_0^t \frac{\mathrm{d}\hat{F}_n}{1-\hat{F}_n^-},$$

where $\hat{G}_n$ the Kaplan-Meier estimator of G, is defined for $t \geq 0$ by:

$$\widehat{G}_n(t) = 1 - \prod_{i=1}^n \left(1 - \frac{\mathbb{1}_{\{T_i \leq t,\ D_i = 0\}}}{n(1 - H_n^-(T_i))}\right).$$

Let (k_n) be a sequence of integers between 1 and $n-1$. In order to always have asymptotical results, we suppose that the sequence (k_n) satisfies the following hypothesis:

$(\mathcal{H}\cdot 1)$: for n large enough, the sequence (k_n/n) is non-increasing and $k_n \ge \log n$,

$(\mathcal{H}\cdot 2)$: for n large enough, the sequence (k_n/n) is non-increasing and there exists a constant $R>0$ such that $k_n \ge R d_n \log n$, with (d_n) is a non-increasing sequence such that:

$$\sum \frac{1}{k d_{2^k} \log k} < \infty.$$

$$\left(\text{e.g.} d_n = (\log\log\log n)^{1+\varepsilon}, d_n = (\log\log\log n)(\log\log\log\log n)^{1+\varepsilon}, \text{etc}\cdots\right).$$

Condition $(\mathcal{H}\cdot 1)$ is required when applying the results of Gin? and Guillou [21] while Condition $(\mathcal{H}\cdot 2)$ is required when applying the results of Cs ö rgö [26].

The following result formulates the laws of the iterated logarithm-type (LIL-type) result on the mentioned increasing intervals.

Theorem 1 (Csörgö [26]; Giné and Guillou [21]) *Let* (k_n) *be a sequence of integers such that* $1 \le k_n < n$ *and, for the almost sure results, satisfying* $(\mathcal{H}\cdot 2)$. *We have*[1]:

$$\sup_{t\le T_{n-k_n,n}} \left|\Lambda_n(t)-\Lambda(t)\right| = \begin{cases} \mathcal{O}\left(\sqrt{\dfrac{\log\log n}{k_{2n}}}\right), \\ \mathcal{O}_{\mathbb{P}}\left(\dfrac{1}{\sqrt{k_n}}\right). \end{cases}$$

If, in addition, F is assumed continuous, then we also have:

$$\sup_{t\le T_{n-k_n,n}} \left|\frac{\hat{F}_n(t)-F(t)}{1-F(t)}\right| = \begin{cases} \mathcal{O}\left(\sqrt{\dfrac{\text{loglog}\, n}{k_{2n}}}\right), \\ \mathcal{O}_{\mathbb{P}}\left(\dfrac{1}{\sqrt{k_n}}\right). \end{cases}$$

Proof. See **Csörgö [26]; Giné and Guillou [21]**. □

The continuity of F is required to linearize the Kaplan-Meier process. Indeed, if F is continuous, then $(\hat{F}_n - F)/(1-F)$ can be approximated by $\Lambda_n - \Lambda$ on the random interval $\left[0, T_{n-k,n}\right]$. Precisely, we have the following result.

Proposition 1 (Giné and Guillou [21]) *Let* (k_n) *be a sequence of integers satisfying* $1 \le k_n < n$ *and Hypothesis* $(\mathcal{H}\cdot 1)$. *If* F *is continuous, then*

$$\sup_{t\le T_{n-k_n,n}} \left|\frac{\hat{F}_n(t)-F(t)}{1-F(t)} - \left(\Lambda_n(t)-\Lambda(t)\right)\right| = \mathcal{O}\left(\frac{\log\log n}{k_n}\right).$$

Proof. See **Giné and Guillou [21]**. □

3. Evaluation of the Conditional Functions of Distribution to the Specific Causes

Let $\tau_1, \tau_2, \cdots, \tau_m$ be a continuous random variables representing respectively the lifetimes in each of the m risks competing, $J = \{1,2,\cdots,m\} \cup \{0\}$ be the set of index cause, where 0 corresponds to the condition of the individual observed, $T = \min(\tau_1, \tau_2, \cdots, \tau_m)$ the random variable of the event of interest and $\eta \in J$ the random variable case, where $\eta = j$ if $T = \tau_j$, for all $j = 1,2,\cdots,m$, F is the distribution function of T, $S = 1-F$, the survival function such that $S(t) = \mathbb{P}[T > t]$, the random variable C of the event censoring right, $\delta = 1\!\!1_{\{T \le C\}}$ and for technical reasons, $\xi = \eta\delta$ such that $\xi = j$ if ($T \le C$ and $\eta = j$) and $\xi = 0$ if $T > C$.

We notice that δ and ξ are observable and η is so only for T uncensored.

We assume that censorship is not informative. The joint law (T, η) is completely specified by the specific

[1] $\mathcal{O}$ is the Landau in almost sure sense and $\mathcal{O}_{\mathbb{P}}$ is the Landau in probability.

incident distributions cause j, $F_j(t)$ defined by

$$F_j(t) = \mathbb{P}[T \le t, \eta = j],\ j = 1,\cdots,m, \tag{1}$$

which are none other than the sub-distributions of the specific cause of failure $j = 1,\cdots,m$.

The cumulative hazard rate of specific-cause $j\,(j = 1,\cdots,m)$ corresponding to (1) is given by

$$\Lambda_j(t) = \int_0^t \frac{\mathrm{d}F_j(s)}{1 - F(s-)}. \tag{2}$$

Let $(Z_1,\delta_1,\xi_1),\cdots,(Z_n,\delta_n,\xi_n)$ be n-sample of observable triplet (Z_i,δ_i,ξ_i) where $Z_i = \min(T_i, C_i)$ and $\delta_i = 1\!\!1_{\{T_i \le C_i\}}$, with $T_i = \min(\tau_1^i,\cdots,\tau_m^i)$ and where τ_j^i represent the time that an individual i is subject to the cause j. If T_i and C_i are independent, the random variable Z_i admits distribution function H_i defined by $1 - H_i = (1 - F_i)(1 - G_i)$. Then the Nelson-Aalen estimator of Λ_j is given for $j = 1,\cdots,m$ by (see e.g. in Andersen *et al.* [4])

$$\widehat{\Lambda}_{jn}(t) = \int_0^t \frac{J(s)dN_j(s)}{Y(s-)} = \sum_{Z_i \le t} \frac{1\!\!1_{\{\xi_i = j\}}}{Y(Z_i-)}, \tag{3}$$

with

$$J(t) = 1\!\!1_{\{Y(t) > 0\}}$$

and where

$$N_j(t) = \sum_{i=1}^n 1\!\!1_{\{Z_i \le t,\ \xi_i = j\}} \tag{4}$$

is the counting of the number of failures observed in case of j the time interval $[0,t]$ and

$$Y(t) = \sum_{i=1}^n 1\!\!1_{\{Z_i \ge t\}} \tag{5}$$

is the number of individuals in the sample observation that survive beyond time t. Thus, for any $j \in \{1,\cdots,m\}$,

$$Y_j(t) = \sum_{i=1}^n 1\!\!1_{\{Z_i \ge t,\ \eta_i = j\}} \tag{6}$$

represents the number of individuals who may fall down specific cause j or be censored.

Estimator similar $\Lambda_j^*(t) = \int_0^t 1\!\!1_{\{Y_j(s) > 0\}} \lambda_j(s) ds$ analogue to (2) and on the sub-group A_j individuals crashing case j is given by

$$\hat{\Lambda}_{jn}^*(t) = \int_0^t \frac{J_j(s)}{Y_j(s-)} dN_j(s) = \sum_{Z_i \le t} \frac{1\!\!1_{\{\xi_i = j\}}}{Y_j(Z_i-)}, \tag{7}$$

and with $J_j(t) = \mathbf{1}_{\{Y_j(t) > 0\}}$ and

$$Y_j(Z_i) = \sum_{k=1}^n 1\!\!1_{\{Z_k \ge Z_i, \eta_k = j\}}.$$

The relation between the cumulative hazard rate Λ_j^* and survival $S_j^* = 1 - F_j^*$ in the subgroup A_j is given by[2]

$$S_j^*(t) = \prod_{s \le t} \left(1 - d\Lambda_j^*(s)\right). \tag{8}$$

A nonparametric estimator of the distribution function $F_j^*(t)$ of time life in subgroups A_j is defined by

[2] $\prod$ denote the product integral (see Gill & Johansen [29]).

$$F_j^*(t) = \mathbb{P}\left[Z \le t \middle| \eta = j\right], 1 \le j \le m \tag{9}$$

is given by

$$1 - \hat{F}_{j_n}^*(t) = \prod_{i=1}^{n}\left(1 - \frac{\mathbb{1}_{\{\xi_i = j\}}\mathbb{1}_{\{Z_i \le t\}}}{Y_j^-(Z_i)}\right). \tag{10}$$

The size $Y_j(t)$ of the subgroup A_j individuals is not observable due to the inaccessibility of all subgroups of specific causes $j \in J$. Nevertheless, we can assign a probability α_{ij} to each of the individuals belonging to one of the m subgroups. Thus, one can estimate the size $Y_j(t)$ by $\hat{Y}_j(t)$ given by (see e.g. in Satten and Datta [27] or Datta and Satten [28]) $\hat{Y}_j(t) = \sum_{i=1}^{n}\hat{\alpha}_{ij}L_i(t)$ where $\hat{\alpha}_{ij}$ is the estimator of the probability that the individual n° i in the sample subgroup A_j, subset of risk of specific-cause j. Thus, the final estimators for the cumulative hazard rate $\Lambda_j^*(t)$ due to the specific cause j and the corresponding distribution function $F_j^*(t)$ have the respective expressions

$$\hat{\Lambda}_{jn}^*(t) = \int_0^t \frac{\mathbb{1}_{\{Y_j(s)>0\}}}{\hat{Y}_j^-(s)} dN_j(s) \tag{11}$$

and for $j = 1, \cdots, m$,

$$\hat{F}_{jn}^*(t) = 1 - \prod_{s \le t}\left(1 - \frac{dN_j(s)}{\hat{Y}_j^-(s)}\right) = 1 - \prod_{Z_i \le t}\left(1 - \frac{\mathbb{1}_{\{\xi_i = j\}}}{\hat{Y}_j(Z_i)}\right). \tag{12}$$

4. Main Results

Let T be a positive random variable and C be a censoring variable such that $Z = T \Lambda C$ and $\delta = \mathbb{1}_{\{T \le C\}}$. In this model of random censorship, for a sample $i = 1, \cdots, n$, subject to a specific causes $j(j = 1, \cdots, m)$, we can observe the couple (Z_i, δ_i) where $Z_i = \min(T_i, C_i)$ and $\delta_i = \mathbb{1}_{\{T_i \le C_i\}}$ with $T_i = \min(\tau_1^i, \cdots, \tau_m^i)$ and where τ_j^i is the time that an individual i is subject to the cause j.

For a given $t \ge 0$ and an individual i with $i = 1, \cdots, n$, the counting process is defined by:

$$K_i(t) = \mathbb{1}_{\{Z_i \le t,\ \delta_i = 1\}}.$$

Therefore, if an individual i undergoes event before time t, then $K_i(t) = 1$; otherwise $K_i(t) = 0$. We can also define the counting process

$$K_i^C(t) = \mathbb{1}_{\{Z_i \le t,\ \delta_i = 0\}}.$$

Naturally, it appears that we considered the information provided over time as a filter, which is used to describe the fact that past information is contained in the current information, hence we have the natural filtration $\{\mathcal{F}_t, t \ge 0\}$ where

$$\mathcal{F}_t = \sigma\left\{K_i(u), K_i^C(u), 0 \le u \le t; 1 \le i \le n\right\}.$$

For $\ll$ samll $\gg$ $\mathrm{d}t$ and for $i \in \{1, \cdots, n\}$, we have

$$\mathbb{P}\left(t \le T_i < t + \mathrm{d}t / T_i \ge t, C_i \ge t\right) = \lambda_i(t)\mathrm{d}t + \mathcal{O}(\mathrm{d}t).$$

If $K_i(t^-)$ denotes the left boundary at t of $K_i(t)$, we have

$$\lambda_i(t)\mathrm{d}t \approx \mathbb{E}\left(K_i\left((t+\mathrm{d}t)^-\right) - K_i(t^-) \middle| T_i \ge t, C_i \ge t\right)$$

since, the quantity $K_i\left((t+\mathrm{d}t)^-\right) - K_i(t^-)$ takes only the values 0 and 1.

For a given $i \in \{1, \cdots, n\}$, we define the function

$$L_i(t) = \mathbb{1}_{\{Z_i \ge t\}}$$

which indicates whether the individual i is still at risk just before time t (the individual has not yet undergone the event). Therefore,

- if $L_i(t) = 0$ then, $\mathbb{P}\left[\mathrm{d}K_i(t) = 1 \middle| \mathcal{F}_{t-}\right] = 0$, and

- if $L_i(t)=1$ then,

$$\begin{aligned}
&\mathbb{P}\left[\mathrm{d}K_i(t)=1\middle|\mathcal{F}_{t-}\right]=\mathbb{P}\left[t<T_i\le t+\mathrm{d}t\middle|T_i>t,C_i>t\right]\\
&=\mathbb{P}\left[t<T_i\le t+\mathrm{d}t\middle|T_i>t\right]\text{ by independence of } T_i \text{ and } C_i\\
&=\lambda_i(t)\mathrm{d}t
\end{aligned}$$

where $\mathcal{F}_t$ is the natural filtration (all information available at time t), where the notation $\mathrm{d}K_i(t)$ refers to formal writing of the stochastic integral

$$K_i(t)=\int_0^t \mathrm{d}K_i(s)=\int_0^t K_i(\mathrm{d}s),$$

writing made possible because $K_i(t)$ is a growing process. The expression of $\lambda_i(t)$ in function of the counting process $K_i(t)$ is given by

$$\lambda_i(t)=\lim_{dt\to 0}\left[\frac{\mathbb{P}\left[K_i(t+\mathrm{d}t)-K_i(t)=1\right]\middle|\mathcal{F}_{t^-}}{\mathrm{d}t}\right]\text{ for all } i\in\{1,\cdots,n\}.$$

Thus, we have $\mathbb{E}\left[\mathrm{d}K_i(t)\middle|\mathcal{F}_{t-}\right]=L_i(t)\lambda_i(t)\mathrm{d}t$.

The stochastic process defined for $t\ge 0$ and $i\in\{1,\cdots,n\}$ by

$$M_i(t)=K_i(t)-\int_0^t L_i(s)\mathrm{d}\Lambda_i(s)$$

is the martingale associated with the subject at risk i. Thereafter Λ_i is the compensating process K_i because it is the integral of the product of two predictable processes.

Theorem 2 *Let T_i be an absolutely continuous lifetime and C_i be a censoring variable for any arbitrary distribution $i\in\{1,\cdots,n\}$. Let λ_i be the risk function associated with T_i. Let's put $Z_i=T_i\Lambda C_i$ and $\delta_i=\mathbb{1}_{\{T_i<C_i\}}$.*

For $t\ge 0,$ the process defined by

$$M_i(t)=K_i(t)-\int_0^t \mathbb{1}_{\{Z_i\ge s\}}\lambda_i(s)ds,$$

is a $\mathcal{F}_{t-}$ martingale if and only if

$$\lambda_i(t)=\frac{-\frac{\partial}{\partial u}\mathbb{P}\left(T_i\ge u,C_i\ge t\right)\Big|_{u=t}}{\mathbb{P}\left(T_i\ge t,C_i\ge t\right)},$$

for t such that $P(Z_i>t)>0.$

Proof. See Breuils ([30], p. 25) and Fleming and Harrington ([7], p. 26). □

For a given $t\ge 0$ and a given $j\in\{1,\cdots,m\}$, the expressions of $N_j(t)$, $Y(t)$ and $\Lambda_j(t)$ are those of formulas (4), (5) and (2) respectively. Using these notations, we can directly obtain the following preliminary result:

Proposition 2 *For a given $t\ge 0$ and a given $j\in\{1,\cdots,m\}$, the stochastic processes defined by*

$$M_j(t)=N_j(t)-\int_0^t Y(s)\mathrm{d}\Lambda_j(s) \tag{13}$$

is the martingale associated with the subject specific cause j.

Proof.

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{d}M_j(t)\middle|\mathcal{F}_{t^-}\right]=\mathbb{E}\left[\mathrm{d}N_j(t)-Y(t)\lambda_j(t)\mathrm{d}t\middle|\mathcal{F}_{t^-}\right]\text{ since } \mathrm{d}\Lambda_j(t)=\lambda_j(t)\mathrm{d}t\\
&=\mathbb{E}\left[\mathrm{d}N_j(t)\middle|\mathcal{F}_{t^-}\right]-Y(t)\lambda_j(t)\mathrm{d}t\text{ since } Y \text{ is measurable with respect to } \mathcal{F}_{t^-}\\
&=Y(t)\lambda_j(t)\mathrm{d}t-Y(t)\lambda_j(t)\mathrm{d}t\\
&=0.
\end{aligned}$$
□

The martingale $M_j(t)$ represents the difference between the number of failures due to a specific cause j observed in the time interval $[0,t]$, *i.e.* $N_j(t)$, and the number of failures predicted by the model for the j^{th}

cause. This definition fulfills the Doob-Meyer decomposition.

The first result of this paper concerns the consistency of the Nelson-Aalen estimator for the competing risks based on martingale approach.

Theorem 3 *For* $t \geq 0$ *such that* $\Lambda_j(t) < \infty$, *we have*

$$\mathbb{E}\left[\hat{\Lambda}^*_{jn}(t) - \Lambda^*_j(t)\right] = 0, \text{ for all } j \in \{1, \cdots, m\}.$$

Proof.

$$\begin{aligned}
\widehat{\Lambda}^*_{jn}(t) - \Lambda^*_j(t) &= \int_0^t \frac{\mathbb{1}_{\{Y_j(s)>0\}}}{\widehat{Y}_j(s-)} dN_j(s) - \int_0^t \mathbb{1}_{\{Y_j(s)>0\}} \lambda_j(s) ds \\
&= \int_0^t \frac{\mathbb{1}_{\{Y_j(s)>0\}}}{\widehat{Y}_j(s-)} [dN_j(s) - \widehat{Y}_j(s-)\lambda_j(s) ds] \\
&= \sum_{j=1}^m \int_0^t \frac{\mathbb{1}_{\{Y_j(s)>0\}}}{\widehat{Y}_j(s-)} dM_j(s) \\
&= \sum_{j=1}^m \int_0^t \frac{J^*_j(s)}{\widehat{Y}_j(s-)} dM_j(s)
\end{aligned}$$

where the expectation of the martingale $M_j(s) = N_j(s) - \int_0^s Y_j(u) d\Lambda_j(u)$ (specific for j^{th} cause) is equal to zero and where $\Lambda_j(t) = \int_0^t \lambda_j(s) ds, t \geq 0$. Indeed,

$$\begin{aligned}
&\mathbb{E}\left[dM_j(t) \middle| \mathcal{F}_{t-}\right] = \mathbb{E}\left[dN_j(t) - Y_j(t)\lambda_j(t)dt \middle| \mathcal{F}_{t-}\right] \\
&= \mathbb{E}\left[dN_j(t) \middle| \mathcal{F}_{t-}\right] - Y_j(t)\lambda_j(t)dt \left(\text{since } Y_j \text{ is measurable with respect to } \mathcal{F}_{t^-}\right) \\
&= Y_j(t)\lambda_j(t)dt - Y_j(t)\lambda_j(t)dt \\
&= 0.
\end{aligned}$$

Hence, we arrive at result.

Using the fact that

$$\hat{\Lambda}^*_{jn}(t) - \Lambda_j(t) = \left[\hat{\Lambda}^*_{jn}(t) - \Lambda^*_j(t)\right] + \left[\Lambda^*_j(t) - \Lambda_j(t)\right]$$

we have:

$$\begin{aligned}
\mathbb{E}[\widehat{\Lambda}^*_{jn}(t) - \Lambda_j(t)] &= \mathbb{E}[\Lambda^*_j(t) - \Lambda_j(t)] \\
&= \mathbb{E}\left[\int_0^t \mathbb{1}_{\{Y_j(s)>0\}} \lambda_j(s) ds - \Lambda_j(t)\right] \\
&= \mathbb{E}\left[\int_0^t \mathbb{1}_{\{Y_j(s)>0\}} \lambda_j(s) ds - \int_0^t \lambda_j(s) ds\right] \quad \text{since } \Lambda_j(t) = \int_0^t \lambda_j(s) ds \\
&= \mathbb{E}\left[\int_0^t (\mathbb{1}_{\{Y_j(s)>0\}} - 1)\lambda_j(s) ds\right] \quad \text{since the integral is a linear form} \\
&= \mathbb{E}\left[\int_0^t (\mathbb{1}_{\{Y_j(s)>0\}} - 1) d\Lambda_j(s)\right] \quad \text{since } d\Lambda_j(t) = \lambda_j(s) ds \\
&= \mathbb{E}\left[-\int_0^t (1 - \mathbb{1}_{\{Y_j(s)>0\}}) d\Lambda_j(s)\right] \\
&= \mathbb{E}\left[-\int_0^t \mathbb{1}_{\{Y_j(s)=0\}} d\Lambda_j(s)\right] \quad \text{since } \mathbf{1} - \mathbb{1}_{\{\mathbb{A}\}} = \mathbb{1}_{\{\mathbb{A}^c\}} \\
&= \mathbb{E}\left[-\int_0^t \mathbb{1}_{\{Y_j(s)=0\}} d\Lambda_j(s)\right] = -\int_0^t \mathbb{E}[\mathbb{1}_{\{Y_j(s)=0\}}] d\Lambda_j(s) \\
&= -\int_0^t \mathbb{P}(Y_j(s) = 0) d\Lambda_j(s) \rightarrow 0 \text{ when } n \rightarrow \infty.
\end{aligned}$$

It follows that $\hat{\Lambda}^*_{jn}(t)$ is an asymptotically unbiased estimator of $\Lambda_j(t)$. Hence, we arrived at result. □

Our second LIL-type result provides almost sure and in probability rates of convergence of $\hat{F}^*_{jn}$ to F^*_j, for $j=1,\cdots,m,$ uniformly over the random increasing intervals $\left[0, Z_{n-k_n,n}\right]$. (See is Deheuvels and Einmahl [31] [32] for very fine results of the model law iterated logarithm functional and available in a point or on a compact strictly included in the support of H). This result is consistent with that of Stute [33] which constitutes a compromise between the results of Breslow and Crowley [9], Földes and Rejtö [10] or Major and Rejtö [11], and those of Földes and Rejtö [12], Gill [13], Csörgö and Horváth [14], Ying [15] and Chen and Lo [16].

Following Giné and Guillou [34], we say that a non-increasing sequence (a_n) of numbers is regular if there exists a constant $Q>0$ such that for all n, $a_{2n}\ge Qa_n$. We denote by $(\mathcal{H})$ the following hypothesis:

$(\mathcal{H})$: for n large enough, the sequence (k_n/n) is regular non-increasing and there exists a constant $R>0$ such that $k_n \ge Rd_n \log n$ with (d_n) is a non-increasing sequence such that

$$\sum \frac{1}{kd_{2^k}\log k}<\infty.$$

$$\left(\text{e.g. } d_n=(\text{logloglog}\, n)^{1+\varepsilon}, d_n=(\text{logloglog}\, n)(\text{loglogloglog}\, n)^{1+\varepsilon}, etc\cdots\right).$$

Theorem 4 *Let* (k_n) *be a sequence of integers such that* $1\le k_n<n$ *for all* n *and which satisfies hypothesis* $(\mathcal{H})$ *for the almost-sure part. For all* $j=1,\cdots,m,$ *we assume that* F^*_j *is alway continuous. Therefore,*

$$\sup_{t\le Z_{n-k_n,n}}\left|\hat{F}^*_{jn}(t)-F^*_j(t)\right|=\mathcal{O}\left(\sqrt{\frac{\text{loglog}\, n}{k_n}}\right),$$

where $\mathcal{O}$ is the Landau in almost sure sense, and

$$\sup_{t\le Z_{n-k_n,n}}\left|\hat{F}^*_{jn}(t)-F^*_j(t)\right|=\mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right),$$

where $\mathcal{O}_{\mathbb{P}}$ is the Landau in probability.

Both results of Theorem above always provides a rate in probability of uniform convergence of $\hat{F}^*_{jn}$ to F^*_j for all $j=1,\cdots,m,$ through a random growing intervals $\left[0,Z_{n-k_n,n}\right]$.

To prove Theorem 4, we have drawn from results based on the inference of empirical processes, given that in order to linearize the Kaplan-Meier process, it is necessary to impose continuity condition on F. Firstly, under the Hypothesis $(\mathcal{H}\cdot 1)$, we have the following result:

Lemma 1 *Let* (k_n) *be a sequence of integers such that* $1\le k_n<n$ *and, for the almost-sure results, such that* $(\mathcal{H}\cdot 1)$ *is satisfied. The rate of convergence of* Λ^*_{jn} *to* Λ^*_j *is given by*

$$\sup_{t\le Z_{n-k_n,n}}\left|\Lambda^*_{jn}(t)-\Lambda^*_j(t)\right|=\begin{cases}\mathcal{O}\left(\sqrt{\frac{\text{loglog}\, n}{k_n}}\right),\\ \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right).\end{cases}$$

Proof. The proof of this result follows straightforwardly from the proof of the first part of Theorem 1 concerning the supremum of $\Lambda_n-\Lambda$. □

Proof of Theorem 4. The following decomposition is obtained for $t\ge 0$ by means of integration by parts:

$$\begin{aligned}\hat{F}^*_{jn}(t)-F^*_j(t)&=\int_0^t\left(1-\hat{F}^*_n(s-)\right)\mathrm{d}\Lambda^*_{jn}(s)-\int_0^t\left(1-F^*(s-)\right)\mathrm{d}\Lambda^*_j(s)\\&=\int_0^t\left(1-\hat{F}^*_n(s-)\right)\mathrm{d}\left(\Lambda^*_{jn}(s)-\Lambda^*_j(s)\right)-\int_0^t\left(\hat{F}^*_n(s-)-F^*(s-)\right)\mathrm{d}\Lambda^*_j(s)\\&=\left(1-\hat{F}^*_n(t-)\right)\left(\Lambda^*_{jn}(t)-\Lambda^*_j(t)\right)+\int_0^t\left(\Lambda^*_{jn}(s)-\Lambda^*_j(s)\right)\mathrm{d}\hat{F}^*_n(s)-\int_0^t\frac{\hat{F}^*_n(s-)-F^*(s-)}{1-F^*(s-)}\mathrm{d}F^*_j(s).\end{aligned}\tag{14}$$

Equality (14) entails that:

$$\sup_{t \le Z_{n-k_n,n}} \left| \hat{F}^*_{jn}(t) - F^*_j(t) \right| \le 2 \sup_{s \le Z_{n-k_n,n}} \left| \Lambda^*_{jn}(s) - \Lambda^*_j(s) \right| + \sup_{s \le Z_{n-k_n,n}} \left| \frac{\hat{F}^*_n(s) - \hat{F}^*(s)}{1 - F^*(s)} \right|.$$

Notice that the assumption of continuity of F_j for $j = 1, \cdots, m$ ensures that F is continuous according to proposition 1. We then conclude with Theorem 1 and Lemma 1. □

5. Conclusion

In this paper, we have adapted the stochastic processes of Aalen [1] [2] to the Nelson-Aalen and Kaplan-Meier [3] estimators in a context of competing risks. We have focused particularly on the probability distributions of complete downtime individuals whose causes are known and which bring us to consider a partition of individuals into sub-groups for each cause. We have also provided some asymptotic properties of nonparametric estimators obtained.

Acknowledgements

I would like to thank Prof. Nicolas Gabriel ANDJIGA, Prof. Celestin NEMBUA CHAMENI, Prof. Eugene Kouassi for their support and their advices. I would also like to thank specially Prof. Kossi Essona GNEYOU for his collaboration and his cooperation during the preparation of this paper.

References

[1] Aalen, O.O. (1978) Nonparametric Estimation of Partial Transition Probabilities in Multiple Decrement Models. *The Annals of Statistics*, **6**, 534-545. http://dx.doi.org/10.1214/aos/1176344198

[2] Aalen, O.O. (1978) Nonparametric Inference for a Family of Counting Processes. *The Annals of Statistics*, **6**, 701-726. http://dx.doi.org/10.1214/aos/1176344247

[3] Kaplan, E.L. and Meier, P. (1958) Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, **53**, 457-481. http://dx.doi.org/10.1080/01621459.1958.10501452

[4] Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993) Statistical Models Based on Counting Processes. Springer Series in Statistics, Spring-Verlag, New York,. http://dx.doi.org/10.1007/978-1-4612-4348-9

[5] Tsiatis, A. (1975) A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceeding of the National Academy of Sciences of the United States of America*, **72**, 20-22. http://dx.doi.org/10.1073/pnas.72.1.20

[6] Heckman, J. and Honoré, B. (1989) The Identifiability of the Competing Risks Models. *Biometrika*, **76**, 325-330. http://dx.doi.org/10.1093/biomet/76.2.325

[7] Fleming, T. and Harrington, D. (1990) Counting Processes and Survival Analysis. John Wiley & Sons, Inc, Hoboken.

[8] Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V., Flournoy, N., Farewell, V.T. and Breslow, N.E. (1978) The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics*, **34**, 541-554. http://dx.doi.org/10.2307/2530374

[9] Breslow, N. and Crowley, J. (1974) A Large Sample Study of the Life Table and Product-Limit Estimates under Random Censorship. *The Annals of Statistics*, **2**, 437-453. http://dx.doi.org/10.1214/aos/1176342705

[10] Földes, A. and Rejtö, L. (1981) Strong Uniform Consistency for Nonparametric Survival Curve Estimators from Randomly Censored Data. *The Annals of Statistics*, **9**, 122-129. http://dx.doi.org/10.1214/aos/1176345337

[11] Major, P. and Rejtö, L. (1998) Strong Embedding of the Estimator of the Distribution Function under Random Censorship. *The Annals of Statistics*, **16**, 1113-1132. http://dx.doi.org/10.1214/aos/1176350949

[12] Földes, A. and Rejtö, L. (1981) A LIL-Type Result for the Product-Limit Estimator. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **56**, 75-86.

[13] Gill, R. (1983) Large Sample Behavior of the Product-Limit Estimator on the Whole Line. *The Annals of Statistics*, **11**, 49-58. http://dx.doi.org/10.1214/aos/1176346055

[14] Csörgö, S. and Horváth, L. (1981) On the Koziol-Green Model for Random Censorship. *Biometrika*, **68**, 391-401.

[15] Ying, Z. (1989) A Note on the Asymptotic Properties of the Product-Limit Estimator on the Whole Line. *Statistics & Probability Letters*, **7**, 311-314. http://dx.doi.org/10.1016/0167-7152(89)90113-2

[16] Chen, K. and Lo, S.-H. (1997) On the Rate of Uniform Convergence of the Product-Limit Estimator: Strong and Weak Laws. *The Annals of Statistics*, **25**, 1050-1087. http://dx.doi.org/10.1214/aos/1069362738

[17] Latouche, A. (2004) Modèles de Régression en Présence de Compétition. Thèse de Doctorat, Université de Paris, Paris.

[18] Belot, A. (2009) Modélisation Flexible des Données de Survie en Présence de Risques Concurrents et Apports de la Méthode du Taux en Excès. Thèse de Doctorat, Université de la Méditerranée, Marseille.

[19] Fine, J.P. and Gray, R.J. (1999) A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, **99**, 496-509. http://dx.doi.org/10.1080/01621459.1999.10474144

[20] Aalen, O.O. and Johansen, S. (1978) An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations. *Scandinavian Journal of Statistics*, **5**, 141-150.

[21] Giné, E. and Guillou, E. (1999) Laws of the Iterated Logarithm for Censored Data. *The Annals of Probability*, **27**, 2042-2067. http://dx.doi.org/10.1214/aop/1022874828

[22] Cox, D. and Oakes, D. (1984) Analysis of Survival Data. Chapman and Hall, London.

[23] Kalbfleisch, J. and Prentice, R. (1980) The Statistical Analysis of Failure Time Data. John Wiley, New York.

[24] Nelson, W. (1969) Hazard Plotting for Incomplete Observations. *Journal of Quality Technology*, **1**, 27-52.

[25] Nelson, W. (1972) A Short Life Test for Comparing a Sample with Previous Accelerated Test Results. *Technometrics*, **14**, 175-185. http://dx.doi.org/10.1080/00401706.1972.10488894

[26] Csörgö, S. (1996) Universal Gaussian Approximations under Random Censorship. *The Annals of Statistics*, **24**, 2744-2778. http://dx.doi.org/10.1214/aos/1032181178

[27] Satten, G.A. and Datta, S. (1999) Kaplan-Meier Representation of Competing Risk Estimates. *Statistics & Probability Letters*, **42**, 299-304. http://dx.doi.org/10.1016/S0167-7152(98)00220-X

[28] Datta, S. and Satten, G.A. (2000) Estimating Future Stage Entry and Occupation Probabilities in a Multistage Model Based on Randomly Right-Censored Data. *Statistics & Probability Letters*, **50**, 89-95. http://dx.doi.org/10.1016/S0167-7152(00)00086-9

[29] Gill, R. and Johansen, S. (1990) A Survey of Product-Integration with a View toward Application in Survival Analysis. *The Annals of Statistics*, **18**, 1501-1555. http://dx.doi.org/10.1214/aos/1176347865

[30] Breuils, C. (2003) Analyse de Durées de Vie: Analyse Séquentielle du Modèle des Risques Proportionnels et Tests d'Homogénéité. Thèse de Doctorat, Université de Technologie de Compiègne, Compiègne.

[31] Deheuvels, P. and Einmahl, J. (1996) On the Strong Limiting Behavior of Local Functionals of Empirical Processes Based upon Censored Data. *The Annals of Statistics*, **24**, 504-525. http://dx.doi.org/10.1214/aop/1042644729

[32] Deheuvels, P. and Einmahl, J. (2000) Functional Limit Laws for the Increments of Kaplan-Meier Product-Limit Processes and Applications. *The Annals of Statistics*, **28**, 1301-1335. http://dx.doi.org/10.1214/aop/1019160336

[33] Stute, W. (1994) Strong and Weak Representations of Cumulative Hazard Function and Kaplan-Meier Estimators on Increasing Sets. *Journal of Statistical Planning and Inference*, **42**, 315-329. http://dx.doi.org/10.1016/0378-3758(94)00032-8

[34] Giné, E. and Guillou, A. (2001) On Consistency of Kernel Density Estimators for Randomly Censored Data. *Annales de l'Institut Henri Poincare* (*B*) *Probability and Statistics*, **37**, 503-522.