Asymptotic Normality of the Nelson-Aalen and the Kaplan-Meier Estimators in Competing Risks

Didier Alain Njamen Njomen

Department of Mathematics and Computer Science, Faculty of Science, University of Maroua, Maroua, Cameroon
Email: didiernjamen1@gmail.com, didier-njamen@univ-maroua.cm

Abstract

This paper studies the asymptotic normality of the Nelson-Aalen and the Kaplan-Meier estimators in a competing risks context in presence of independent right-censorship. To prove our results, we use Robelledo’s theorem which makes it possible to apply the central limit theorem to certain types of particular martingales. From the results obtained, confidence bounds for the hazard and the survival functions are provided.

Keywords
Censored Data, Right-Censoring, Counting Process, Competing Risks, Nelson-Aalen and Kaplan-Meier Estimators, Asymptotic Properties of Estimators, Confidence Bands

1. Introduction and Background

The model of competing risks has been widely studied in the literature, see e.g., Heckman and Honoré [1], Commenges [2], Com-nougué [3], Fine and Gray [4], Crowder [5], Fermanian [6], Latouche, A. [7], Geffray [8], Belot [9], Njamen and Ngatchou ([10], [11]), Njamen ([12], [13]). In most approaches, the competing risks are assumed to be either all independent or all dependent. Here, the independent component of the potential risks constitutes an independent censoring variable while the other risks are kept as possibly dependent. This approach is used by Geffray [8]. Namely, we consider a population in which each subject is exposed to m mutually exclusive competing risks which may be dependent. For \( j \in \{1, \ldots, m\} \), the failure time from the \( j^{\text{th}} \) cause is a non-negative random variable (r.v.) \( \tau_j \). The competing risks model postulates
that only the smallest failure time is observable, it is given by the r.v. 
\( T = \min(\tau_1, \cdots, \tau_m) \) with distribution function (d.f.) denoted by \( F \). The cause of failure associated to \( T \) is then indicated by a r.v. \( \eta \) which takes value \( j \) if the failure is due to the \( j^{th} \) cause for \( j \in \{1, \cdots, m\} \), i.e. \( \eta = j \) if \( T = \tau_j \). The following modeling technique is extracted in Njamen and Ngatchou [10]: we assume that \( T \) is, in its turn, at risk of being independently right-censored by a non-negative r.v. \( C \) with d.f. \( G \). Consequently, the observable r.v. are
\[
\left( Z = \min(T, C), \xi = \eta \delta \right),
\]
where \( \delta = \mathbb{I}_{T < C} \) and where \( \mathbb{I}_{(\cdot)} \) denotes the indicator function. As \( T \) and \( C \) are independent, the r.v. \( Z \) has d.f. \( H \) given by \( 1 - H = (1 - F)(1 - G) \). Let \( \tau_H = \sup\{t : H(t) < 1\} \) denote the right-endpoint of \( H \) beyond which no observation is possible. The subdistribution functions \( F^{(j)} \) pertaining to the different risks or causes of failure are defined for \( j = 1, \cdots, m \) and \( t \geq 0 \) by
\[
F^{(j)}(t) = \mathbb{P}[T \leq t, \eta = j], \quad j = 1, \cdots, m
\]
(1)

When the independence of the different competing risks may not be assumed, the functions \( F^{(j)} \) for \( j = 1, \cdots, m \) are the basic estimable quantities.

The Kaplan-Meier estimator was developed for situations in which only one cause of failure and the independent right-censoring are considered. Aalen and Johansen [14] were the first to extend the Kaplan-Meier estimator to several causes of failure in the presence of independent censoring. In the present situation, the d.f. \( F \) may be consistently estimated by the Kaplan-Meier estimator denoted by \( \hat{F}_n \). For \( j = 1, \cdots, m \), the subdistribution functions \( F^{(j)} \) may be consistently estimated by means of the Aalen-Johansen estimators denoted respectively by \( \hat{F}^{(j)}_n \), for \( j = 1, \cdots, m \). Indeed, when the process of the states occupied by an individual in time is a time-inhomogeneous Markov process, Aalen and Johansen [14] introduced an estimator of the transition probabilities between states in presence of independent random right-censoring. The competing risks set-up corresponds to the case of a time-inhomogeneous Markov process with only one transient state and several absorbing states (that can be labeled \( 1, \cdots, m \)). Aalen and Johansen [14] obtained the joint consistency of \( \hat{F}^{(j)}_n \) to \( F^{(j)} \) for \( j = 1, \cdots, m \) uniformly over fixed compact intervals \( [0, \sigma] \) for \( \sigma < \tau_H \). They also obtained the joint weak convergence of the processes \( \sqrt{n}(\hat{F}^{(j)}_n - F^{(j)}) \) on fixed compact intervals \( [0, \sigma] \) for \( \sigma < \tau_H \).

The asymptotic properties of the Kaplan-Meier estimator on the distribution function have been studied by several authors (see Peterson [15], Andersen and al. [16], Shorack and Wellner [17], Breslow and Crowley [18]).

In this paper, in a region where there is at least one observation, we are interested in providing asymptotic properties of the Nelson-Aalen and Kaplan-Meier nonparametric estimators of the functions \( \Lambda^{(j)} \) and \( S^{(j)} \). For \( j = 1, \cdots, m \) in the presence of independent right-wing censorship in the context of competitive risks set out in Njamen and Ngatchou ([10], [11]).

The rest of the paper is organized as follows: Section 2 describes preliminary
results and rappels used in the paper. In Section 3, we obtain two laws: In Section 3.1, we give limit law of Nelson-Aalen’s nonparametric estimator for competing risks as defined in Njamen and Ngatchou [10] and Njamen [12]. In Sect. 3.2, we give limit law of Kaplan-Meier’s nonparametric estimator in competing risks as defined in Njamen and Ngatchou [10] and Njamen [13]. In Section 4, we give the trust Bands, including the Hall-Wellner trust Bands and the Nair precision equal bands.

2. Preliminary and Rappels

For \( t \geq 0 \), we introduce the following subdistribution functions \( H^{(0)} \) and \( H^{(i)} \) of \( H \) by:

\[
H^{(0)}(t) = \mathbb{P}[Z \leq t, \xi = 0],
\]

and

\[
H^{(i)}(t) = \mathbb{P}[Z \leq t, \xi \neq 0]
\]

and for \( j = 1, \ldots, m \)

\[
H^{(i,j)}(t) = \mathbb{P}[Z \leq t, \xi = j].
\]

The relations \( F(t) = \sum_{j=1}^{m} F^{(j)}(t) \) and \( H^{(0)}(t) = \sum_{j=1}^{m} H^{(i,j)}(t) \) hold for \( t \geq 0 \) since the different risks are mutually exclusive. The relation \( H(t) = H^{(0)}(t) + H^{(i)}(t) \) is also valid for \( t \geq 0 \). The relations that connect the observable distribution functions \( H^{(0)} \), \( H^{(i)} \) and \( H^{(i,j)} \) to the unobservable distributions \( F \), \( G \) and \( F^{(j)} \) are given by:

\[
H^{(0)}(t) = \int_0^t (1 - F) \, dG,
\]

\[
H^{(i)}(t) = \int_0^t (1 - G^-) \, dF,
\]

and

\[
H^{(i,j)}(t) = \int_0^t (1 - G^-) \, dF^{(j)}.
\]

The cumulative hazard function of \( T \) and the partial cumulative hazard function of \( T \) related to cause \( j \) for \( j \in \{1, \ldots, m\} \) are given for \( t \geq 0 \) respectively by the following expressions:

\[
\Lambda(t) = \int_0^t \frac{dF}{1 - F^-} = \int_0^t \frac{dH^{(i)}}{1 - H},
\]

(2)

\[
\Lambda^{(i,j)}(t) = \int_0^t \frac{dF^{(j)}}{1 - F^-} = \int_0^t \frac{dH^{(i,j)}}{1 - H^-}.
\]

(3)

Let us set estimators for the different quantities. Let \( (Z_i, \xi_i)_{i=1,\ldots,n} \) be \( n \) independent copies of the random vector \( (Z, \xi) \). We define the empirical counterparts of \( H^{(0)} \), \( H^{(i)} \), \( H^{(i,j)} \) and \( H \) for \( j \in \{1, \ldots, m\} \) by:

\[
H_n^{(0)}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(Z_i, \xi_i) \in \{Z \leq t, \xi = 0\}},
\]

\[
H_n^{(i)}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(Z_i, \xi_i) \in \{Z \leq t, \xi \neq 0\}},
\]

and

\[
H_n^{(i,j)}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(Z_i, \xi_i) \in \{Z \leq t, \xi = j\}}.
\]
The relations $H_n(t) = H_n^{(0)}(t) + H_n^{(1)}(t)$ and $H_n^{(1)}(t) = \sum_{j=1}^{m} H_{n,j}^{(1)}(t)$ are valid for $t \geq 0$. As $T$ is independently randomly right-censored by $C$, a well-known estimator for $F$ is the Kaplan-Meier estimator defined for $t \geq 0$ by:

$$\hat{F}_n(t) = 1 - \prod_{i=1}^{n} \left(1 - \frac{1 - \mathbb{I}_{[Z_i \leq \xi]}(Z_i)}{n(1-H_n(Z_i))}\right),$$

where the left-continuous modification of any d.f. $L$ is denoted by $L^-$. The Nelson-Aalen estimators of $\Lambda$ and of $\Lambda^{(1,j)}$ for $j=1,\ldots,m$ respectively are defined for $t \geq 0$ by:

$$\Lambda_n(t) = \int_0^t \frac{dH_n^{(1)}}{1-H_n},$$

$$\Lambda_n^{(1,j)}(t) = \int_0^t \frac{dH_n^{(1,j)}}{1-H_n}.$$  \hspace{1cm} (4)

The Aalen-Johansen estimator for $F^{(j)}$ is defined for $t \geq 0$ by:

$$\hat{F}_n^{(j)}(t) = \int_0^t \frac{1-\hat{F}_n}{1-H_n} dH_n^{(1,j)}.$$  \hspace{1cm} (5)

For all $t \geq 0$, the following equalities hold:

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

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

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

$$\hat{G}_n(t) = 1 - \prod_{i=1}^{n} \left(1 - \frac{1 - \mathbb{I}_{[Z_i \leq \xi]}(Z_i)}{n(1-H_n(Z_i))}\right).$$

3. Results

In this section, we continue the works of Njamen and Ngatchou [10], Njamen [12] and Njamen and Ngatchou [11]. In fact, Njamen and Ngatchou ([10], p. 9), studies the consistency of Nelson-Aalen’s non-parametric estimator in competing risks, while Njamen ([12], pp. 11-12) studies respectively the simple convergence and the uniform convergence in probability of Nelson-Aalen’s nonparametric estimator in competing risks; and Njamen and Ngatchou ([11], p. 13) study the bias and the uniform convergence of the non-parametric estimator survival function in a context of competing risks. It is also shown there that this
estimator is asymptotically unbiased. For this purpose, we use the martingale approach as the authors mentioned above.

3.1. Limit Law of Nelson-Aalen’s Nonparametric Estimator for Competing Risks

In what follows, we study the asymptotic normality of Nelson-Aalen’s non-parametric estimator in competitive risks. For that, considering, for all \( j \in \{1, \ldots, m\} \) and \( t \geq 0 \), one has the Nelson-Aalen type cumulative hazard function estimator (Nelson, [19]; Aalen, [20], Njamen and Ngatchou, [10]) defined by

\[
\hat{\Lambda}_n(t) = \int_0^t \frac{J(u)}{Y(u)} \, dN(u),
\]

where \( J(t) = \mathbb{1}_{[Y(t) > 0]} \).

The cumulative risk in a region where there is at least one observation is given for all \( j \in \{1, \ldots, m\} \), by (see Njamen, [12]. p. 9)

\[
\Lambda^{(j)} = \int_0^t L^{(j)} \Lambda^{(j)}(s) \, ds,
\]

with \( L^{(j)}(t) = \mathbb{1}_{[Z(t) > 0]} \) which indicates whether the individual \( j \) is still at risk just before time \( t \) (the individual has not yet undergone the event). Its estimator was defined in Njamen and Ngatchou ([10], p. 7).

The following theorem gives the limit law of the Nelson-Aalen estimator \( \hat{\Lambda}_n^{(j)} \) in competing risks of Njamen (2017, p. 9). This is the first fundamental result of this article.

**Theorem 1.**

In a region where there is at least one observation, it is assumed that \( F_i^{(j)}(t) < 1 \) for \( i \in \{1, \ldots, n\} \) and \( j \in \{1, \ldots, m\} \). Then, for all \( t \geq 0 \),

\[
\sqrt{n} \left( \hat{\Lambda}_n^{(j)}(t) - \Lambda^{(j)}(t) \right) \xrightarrow{d} U_t^{(j)}(t),
\]

where \( U_t^{(j)} \) is a centered Gaussian martingale of variance such that:

\[
\begin{align*}
\mathbb{E}(U_t^{(j)}(0)) &= 0, \\
\mathbb{V}(U_t^{(j)}(t)) &= \int_0^t \frac{\alpha_i^{(j)}(u)}{y_i^{(j)}(u)} \, du,
\end{align*}
\]

where for all \( s \geq 0 \),

\[
y_i^{(j)}(s) = \left[ 1 - F_i^{(j)}(s) \right] \left[ 1 - G_i^{(j)}(s) \right]
\]

with \( G_i^{(j)} \) standing for the distribution function of \( C_i^{(j)} \) and \( \alpha_i^{(j)} \) the instant risk function.

To prove this theorem, we need the Robeldeo theorem. In fact, the Rebolledo theorem below makes it possible to apply the central limit theorem for certain types of particular martingales.

**Theorem 2. (Rebolledo’s Theorem)**

Let \( M^n = \sum_{i=1}^n M_i \) a sequence of martingales where \( M_i = K_i - A_i \), \( K_i \).
denotes a counting process and \( A \) its compensator. Consider the processes \( I_n(t) = \int_0^t f_n(s) dM^n(s) \), and for all \( \varepsilon > 0 \), \( I_{n,\varepsilon}(t) = \int_0^t f_n(s) \mathbb{I}_{[f_n(s) \in \varepsilon]} dM^n(s) \).

Suppose that \( f_n \) and \( f \) are predictable and locally bounded \( \mathcal{F}_s \) processes such that

\[
\sup_s |f_n - f(s)| \to 0 \quad (n \to \infty).
\]

Suppose also that the processes \( K_i, A, f_n \) are bounded. Let’s for all \( t > 0 \),

1) \( \{I_n\}_{t} \overset{p}{\to} \alpha(t), (n \to \infty) \);
2) for all \( \varepsilon > 0 \), \( \{I_{n,\varepsilon}\}_{t} \overset{p}{\to} 0, (n \to \infty) \).

Then,

\[
(I_n(t), t > 0) \Rightarrow \left( \int_0^t f(s) dW(s), t > 0 \right), \quad (n \to \infty),
\]

where \( \Rightarrow \) denotes the weak convergence in the space of continuous functions on the right, having a left-hand boundary with the topology of Skorokhod and where \( W \) is a Brownian motion.

To prove Theorem 1, it is sufficient to check whether the previous conditions of Rebolledo’s Theorem are satisfied:

**Proof.** For all \( j \in \{1, \ldots, m\} \) and \( t \geq 0 \), \( M^{(j)}(t) \) also decomposes into

\[
M^{(j)}(t) = K^{(j)}(t) - \int_0^t d\Lambda^{(j)}(s) ds,
\]

which in turn can be written in terms of \( \alpha_j(t) \) by

\[
M^{(j)}(t) = K^{(j)}(t) - \int_0^t \alpha^{(j)}(s) L^{(j)}(s) ds,
\]

which finally, can be rewritten as

\[
dK^{(j)}(t) = \alpha^{(j)}(t) L^{(j)}(t) dt + dM^{(j)}(t),
\]

where \( dM^{(j)}(t) \) can be seen as a random noise process. The martingale \( M^{(j)}(t) \) above represents the difference between the number of failures due to a specific cause \( j \) observed in the time interval \( [0, t] \), i.e. \( K^{(j)}(t) \) (see Njamen, [12], p.6), and the number of failures predicted by the model for the \( j \)th cause. This definition fulfills the Doob-Meyer decomposition.

This martingale is used in Fleming and Harrington ([21], p. 26) and in Breuils ([22], p. 25).

Now, to explain the asymptotic nature of the results, we defined, for all \( t \geq 0 \), \( j \in \{1, \ldots, m\} \), to pose:

\[
N(t) = \sum_{i=1}^t K_i(t), Y(t) = \sum_{i=1}^t L_i(t), J(t) = \mathbb{I}_{[t \leq \tau]}.
\]

In a subgroup \( A^{(j)} \), where there is at least one observation, the survival function of \( Z_j = \min(T, C_i) \) is defined for all \( t \geq 0 \) by:

\[
S^{(j)}_z(t) = \left( 1 - F^{(j)}_z(t) \right) \left( 1 - G^{(j)}(t) \right).
\]
Recall also that $F^{(i)}_t$ is the distribution function of $T_i$, $G^{(i)}_t$ is that of $C_i$’s and $1 - (1 - F^{(i)}_t)^{\left[1 - G^{(i)}_t\right]}$ that of the $Z_i$’s. From the Glivenko-Cantelli theorem, one has:

$$\sup_{s \in [0,1]} \left| \frac{F^{(i)}_s (s)}{n} - \left[1 - F^{(i)}_s (s)\right] \left[1 - G^{(i)}_s (s)\right] \right| \xrightarrow{p} 0 \quad (n \to \infty). \quad (11)$$

Otherwise,

$$J^{(i)}(t) = \mathbb{I}_{\left[\left(1 - F^{(i)}_t\right)\left(1 - G^{(i)}_t\right)\right]},$$

one has:

$$1 - J^{(i)}(t) = \mathbb{I}_{\left[\left(1 - F^{(i)}_t\right)\left(1 - G^{(i)}_t\right)\right]} = \mathbb{I}_{\left[n \left[1 - F^{(i)}_t\right] \left[1 - G^{(i)}_t\right]\right]} \xrightarrow{p} 0 \quad (n \to \infty),$$

from which one obtains (see Theorem 3, p. 11 of Njamen, [12]),

$$J^{(i)}(t) \xrightarrow{p} 1 \quad (n \to \infty).$$

Differentiating the martingale $M^{(i)}_t = K^{(i)}_t - \int_0^t L^{(i)}_s (s) \alpha^{(i)}_t (s) ds$, one has:

$$dM^{(i)}_t (t) = dK^{(i)}_t (t) - L^{(i)}_t (t) \alpha^{(i)}_t (t) dt,$$

and from

$$d \left\{ M^{(i)}_t \right\}_t = \mathbb{V}ar \left\{ dM^{(i)}_t (t) / \mathcal{F}_t \right\},$$

one obtains

$$d \left\{ M^{(i)}_t \right\}_t = \mathbb{V}ar \left\{ dK^{(i)}_t (t) - \int_0^t L^{(i)}_s (s) \alpha^{(i)}_s (s) ds / \mathcal{F}_t \right\} = \mathbb{V}ar \left\{ dK^{(i)}_t (t) / \mathcal{F}_t \right\} = L^{(i)}_t (t) \alpha^{(i)}_t (t) dt.$$

Consequently, the increasing process of

$$D_t = \int_0^t \frac{J^{(i)}(u)}{\left(\frac{1}{y^{(i)}(u)}\right)} dM^{(i)}_t (u), \quad t \geq 0,$$

is given by

$$\left\langle D_t \right\rangle_t = \int_0^t \frac{\left(J^{(i)}(u)\right)^2}{\left(\frac{1}{y^{(i)}(u)}\right)^2} dM^{(i)}_t (u), \quad t \geq 0.$$

Next, for all $t \geq 0$ and $j = \{1, \ldots, m\}$, one has

$$\sqrt{n} \sum_{i=1}^n \int_0^t \frac{J^{(i)}(u)}{\left(\frac{1}{y^{(i)}(u)}\right)} dM^{(i)}_t (u) = \sqrt{n} \sum_{i=1}^n \int_0^t \frac{\left(J^{(i)}(u)\right)^2}{\left(\frac{1}{y^{(i)}(u)}\right)^2} L^{(i)}_t (u) \alpha^{(i)}_t (u) du$$

$$= \int_0^t \frac{\left(J^{(i)}(u)\right)^2}{\left(\frac{1}{y^{(i)}(u)}\right)^2} \sum_{i=1}^n L^{(i)}_t (u) \alpha^{(i)}_t (u) du$$

$$= \int_0^t \frac{\left(J^{(i)}(u)\right)^2}{\left(\frac{1}{y^{(i)}(u)}\right)^2} Y^{(i)}_t (u) \alpha^{(i)}_t (u) du$$

$$= \int_0^t \frac{J^{(u)}(u)}{\left(\frac{1}{y^{(i)}(u)}\right)} \alpha^{(i)}_t (u) du.$$
Also, for all \( t \geq 0 \) and for all \( j \in \{1, \ldots, m\} \), the process
\[
\sqrt{n} \left( \Lambda_n^{(j)}(t) - \Lambda^{(j)}(t) \right) = \sqrt{n} \sum_{i=1}^{n} J^{(a)}(u) \frac{dM_i^{(j)}(u)}{Y^{(a)}(u)} = R_n(t), \quad \forall i \in \{1, \ldots, n\},
\]
is a martingale. We apply the central limit theorem for the martingales (Rebolledo’s Theorem). In this purpose, we show that the condition of this theorem is satisfied by \( R_n(t) \).

One has, for all \( i \in \{1, \ldots, n\} \),
\[
\{R_n\}_i = \int_0^n \frac{J^{(a)}(u)}{Y^{(a)}(u)} \alpha_i^{(j)}(u) \, du, \quad \forall j \in \{1, \ldots, m\},
\]
and also by the proof of the Theorem 3 of Njamen ([12], p. 11), we have:
\[
\frac{Y^{(a)}(u)}{n} \xrightarrow{p} \left(1 - F_i^{(j)}(u)\right)\left(1 - G_i^{(j)}(u)\right), \quad J^{(a)}(u) \xrightarrow{p} 1, \quad (n \to \infty).
\]
So that, for all \( j \in \{1, \ldots, m\} \), when \( n \to \infty \),
\[
\{R_n\}_i = \int_0^n \frac{J^{(a)}(u)}{Y^{(a)}(u)} \alpha_i^{(j)}(u) \, du \xrightarrow{p} \int_0^n \frac{\alpha_i^{(j)}(u) \, du}{\left(1 - F_i^{(j)}(u)\right)\left(1 - G_i^{(j)}(u)\right)} = \beta(t), \quad (n \to \infty),
\]
which is determinist. Thus, the first condition of Robelledo Theorem holds.

To check the second condition, for all \( \epsilon > 0 \) and \( t \geq 0 \), define
\[
R_{n, \epsilon}(t) = \int_0^{t} \sqrt{n} \frac{J^{(a)}(u)}{Y^{(a)}(u)} \frac{1}{\left|\frac{dJ^{(a)}(u)}{dY^{(a)}(u)}\right|} \, dM^{(a)}(u),
\]
where for all \( j = 1, \ldots, m \), \( M^{(a)}(u) = \sum_{i=1}^{n} M_i^{(j)}(u) \).

We have to show that as \( n \to \infty \), \( \{Z_{n, \epsilon}\}_i \) converges to 0 in probability.

One has, for all \( t \geq 0 \),
\[
\{R_{n, \epsilon}\}_i = \int_0^n \frac{J^{(a)}(u)}{Y^{(a)}(u)} \frac{1}{\left|\frac{dJ^{(a)}(u)}{dY^{(a)}(u)}\right|} \frac{d\{M^{(a)}\}}{\frac{d\{M^{(a)}\}}{Y^{(a)}(u)}} \, du \rightarrow 0, \quad (n \to \infty),
\]
because
\[
\frac{n J^{(a)}(u)}{Y^{(a)}(u)} \xrightarrow{p} \frac{1}{\left(1 - F_i^{(j)}(u)\right)\left(1 - G_i^{(j)}(u)\right)}, \quad (n \to \infty).
\]
Then
\[ \sqrt{n} \frac{J^{(a)}(u)}{Y^{(a)}(u)} = \sqrt{n} \frac{J^{(a)}(u)}{Y^{(a)}(u)} \xrightarrow{p} 0, \quad (n \to \infty). \]

Thus, the second condition of Robelledo Theorem holds.

The conditions of the Rebolledo Theorem are verified and by consequently, for all \( t \geq 0 \),
\[ (R_n(t), t > 0) \Rightarrow \left( \int_0^t f(s) dW(s), t > 0 \right), \quad (n \to \infty), \]
with \( \gamma(t) = \int_0^t f^2(s) ds \).

Finally, for all \( t > 0 \),
\[ R_n(t) \Rightarrow R(t) = \int_0^t f(s) dW(s) \xrightarrow{d} \mathcal{N}(0, \gamma(t)), \quad (n \to \infty). \]

This ends the proof of the Theorem 1.

The following subsection gives the asymptotic law of nonparametric Kaplan-Meier’s estimator of the survival function in the competing risks of Njamen and Ngatchou ([10], p. 13).

### 3.2. Limit Law of Kaplan-Meier’s Nonparametric Estimator in Competing Risks

The Kaplan-Meier estimator of the survival function (Kaplan and Meier, [23]) is defined by
\[ \hat{S}_n(t) = \prod_{s \leq t} \left(1 - \Delta \hat{\Lambda}_n(s)\right) = \prod_{s \leq t} \left(1 - \frac{J^{(a)}(s) \Delta \Lambda^{(a)}(s)}{Y^{(a)}(s)}\right), \]
where \( \hat{\Lambda}_n(t) \) is the Nelson-Aalen estimator and where, for a process \( X(t) \) continuous to the right with a left limit such that
\[ \Delta X(t) = X(t) - X(t^-). \]

For all \( j = 1, \cdots, m \), an estimator of the variance of \( \hat{S}_n^{(j)}(t)/S^{(j)}(t) \), where \( S^{(j)} \) is the survival function associated with the subgroup \( A^{(j)} \) is given by
\[ \hat{\sigma}^{(j)^2}(t) = \int_0^t \frac{J^{(a)}(s)}{Y^{(a)}(s)^2} \sigma^{(j)}(s) ds. \]

The variance of \( \hat{S}_n^{(j)}(t)/S^{(j)}(t) \) approximated by that of \( \hat{S}_n^{(j)}(t)/S^{(j)}(t) \) is:
\[ \sqrt{\hat{V} \left[ \frac{\hat{S}_n^{(j)}(t)}{S^{(j)}(t)} \right] - 1} = \mathbb{E} \left[ \left( \frac{\hat{S}_n^{(j)}(t)}{S^{(j)}(t)} - 1 \right)(t) \right] \]
\[ = \int_0^t \left( \frac{\hat{S}_n^{(j)}(s)}{S^{(j)}(s)} \right)^2 \frac{J^{(a)}(s)}{Y^{(a)}(s)} \sigma^{(j)}(s) ds \quad \forall i \in \{1, \cdots, n\}. \]

The estimator of the corresponding variance of \( \hat{S}_n^{(j)}(t) \) is given by
\[ \hat{V} \left( \hat{S}_n^{(j)}(t) \right) = \left[ \hat{S}_n^{(j)}(t) \right]^2 \hat{\sigma}^{(j)^2}(t) \quad \forall i \in \{1, \cdots, n\}. \]
The following result concerning the asymptotic law of nonparametric Kaplan-Meier estimator and constituted the second fundamental result of this paper:

**Theorem 3.**

In an area where there is at least one observation, if we assume that for all \( j \in \{1, \ldots, m\} \) and \( i \in \{1, \ldots, n\} \),

1) For all \( s \in [0, t] \),

\[
\frac{1}{n} \int_0^s \frac{j^{(a)}(u)}{y^{(a)}(u)} \alpha_i^{(j)}(u) du \xrightarrow{p} \sigma_i^{(j)}(u) \quad (n \to \infty),
\]

2) For all \( \varepsilon > 0 \),

\[
\frac{1}{n} \int_0^s \frac{j^{(a)}(u)}{y^{(a)}(u)} \alpha_i^{(j)}(u) \left\{ \sqrt{s} \left[ j^{(a)}(u) \right] \right\} du \xrightarrow{p} 0 \quad (n \to \infty),
\]

3) For all \( t > 0 \),

\[
\sqrt{n} \int_0^t (1 - j^{(a)}(u)) \alpha_i^{(j)}(u) du \xrightarrow{p} 0 \quad (n \to \infty).
\]

Then, for all \( t > 0 \) and \( j \in \{1, \ldots, m\} \), the non-parametric estimator \( \hat{S}_{n}^{(j)}(u) \) checks

\[
\sqrt{n} \left( \hat{S}_{n}^{(j)}(t) - S^{(j)}(t) \right) \Rightarrow -U^{(j)}(t) \times S^{(j)}(t), \quad (n \to \infty),
\]

where \( U^{(j)}(t) \) is the center Gaussian martingale and where \( \Rightarrow \) denotes the weak convergence in the space of continuous functions on the right, having a left-hand boundary with the topology of Skorokhod.

**Proof.** To prove this theorem, it suffices to show that it satisfies the conditions of the Rebolledo Theorem.

In an area where there is at least one observation, by posing, for all \( j = 1, \ldots, m \), \( i = 1, \ldots, n \),

\[
\hat{S}_{n}^{(j)}(t) = \exp \left( -\hat{\lambda}_{n}^{(j)} \right)
\]

where \( \hat{\lambda}_{n}^{(j)} = \int_0^t j^{(a)}(u) \alpha_i^{(j)}(u) du \).

For \( t \in [0, T] \) and \( \tau > 0 \), we have for all \( j = 1, \ldots, m \) and \( i = 1, \ldots, n \),

\[
\sqrt{n} \left( \frac{\hat{S}_{n}^{(j)}(u)}{S_{n}^{(j)}(u)} - 1 \right) = n \int_0^u \frac{S_{n}^{(j)}(u)}{S_{n}^{(j)}(u)} \alpha_i^{(j)}(u) \frac{f^{(a)}(u)}{y^{(a)}(u)} du \xrightarrow{p} \sigma_j^{(j)}(u) \quad (n \to \infty).
\]

By the proof of Theorem 3 of Njamen ([12], p.11), we deduce that

\[
\frac{\hat{S}_{n}^{(j)}(u)}{S_{n}^{(j)}(u)} \xrightarrow{p} 1, \quad (n \to \infty).
\]

Hence the 1st condition of Robolledo’s Theorem.

For the second condition of Robolledo’s Theorem, condition B is similar to
the proof of Theorem 1 above, we find that for all \( \epsilon > 0 \),

\[
\lim_{n \to \infty} \int_0^\infty \frac{\tilde{S}_n^{(j)}(u)}{S_n^{(j)}(u)} \frac{J^{(n)}(u)}{Y^{(n)}(u)} \alpha_n^{(j)}(u) \, du = 0, \quad (n \to \infty).
\]

So, for each \( t > 0 \),

\[
\sqrt{n} \int_0^t \frac{\tilde{S}_n^{(j)}(u^{-})}{S_n^{(j)}(u)} \frac{J^{(n)}(u)}{Y^{(n)}(u)} \, dM^{(n)}(u) \Rightarrow U^{(j)}_{\epsilon}(t),
\]

where \( M^{(n)}(u) = \sum_{i=1}^n M_t^{(j)}(u) \) and where

\[
U^{(j)}_{\epsilon}(t) \Rightarrow N(0, \sigma_n^{(j)}(t)).
\]

Finally,

\[
\sqrt{n} \left( \frac{\tilde{S}_n^{(j)}(t)}{S_n^{(j)}(t)} - 1 \right) \Rightarrow -U^{(j)}_{\epsilon}(t).
\]

The fact that \( S^{(j)}(u) \leq \tilde{S}_n^{(j)}(u) \), for all \( u \in [0, s] \) and condition C implies:

\[
\sqrt{n} \left| \frac{S^{(j)}(s)}{S_n^{(j)}(u)} - 1 \right| \leq \sqrt{n} \int_0^s \frac{S^{(j)}(u)}{S_n^{(j)}(u)} \, d\left( \Lambda^{(j)} - \Lambda^{(j)} \right)(u)
\]

\[
\leq \sqrt{n} \int_0^s (1 - J(u)) \alpha^{(j)}(u) \, du \xrightarrow{p} 0 \quad (n \to \infty).
\]

As \( \tilde{S}_n^{(j)}(t) \to S^{(j)}(t) \) when \( n \to \infty \), we deduce that:

\[
\sqrt{n} \left( \tilde{S}_n^{(j)}(t) - S^{(j)}(t) \right) \xrightarrow{p} 0, \quad n \to \infty.
\]

It follows that:

\[
\sqrt{n} \left( \frac{\tilde{S}_n^{(j)}(t)}{S_n^{(j)}(t)} - S^{(j)}(t) \right)
\]

\[
= \sqrt{n} \left( \frac{\tilde{S}_n^{(j)}(t) - \tilde{S}_n^{(j)}(t)}{S_n^{(j)}(t)} \right) + \sqrt{n} \left( \frac{\tilde{S}_n^{(j)}(t) - S^{(j)}(t)}{S_n^{(j)}(t)} \right)
\]

\[
= \frac{\sqrt{n} \left( \tilde{S}_n^{(j)}(t) - \tilde{S}_n^{(j)}(t) \right)}{\tilde{S}_n^{(j)}(t)} + \frac{\sqrt{n} \left( \tilde{S}_n^{(j)}(t) - S^{(j)}(t) \right)}{S_n^{(j)}(t)}
\]

\[
\Rightarrow -U^{(j)}_{\epsilon}(t) S^{(j)}, \quad (n \to \infty).
\]

This ends the proof of the theorem.

4. Confidence Bands of Survival Function

4.1. Confidence Intervals

For \( \alpha \in (0,1) \), we wish to find two random functions \( b_L \) and \( b_U \) such that \( \forall t > 0 \),

\[
P \left( b_U(t) \geq S(t) \geq b_L(t) \right) = 1 - \alpha.
\]

Recall that from the previous sections, for all \( j \in \{1, \cdots, m\} \),

\[
\sqrt{n} \left( \tilde{S}_n^{(j)}(t) - S^{(j)}(t) \right) \bigg/ S_n^{(j)}(t)
\]

converges in distribution to a Gaussian
martingale centered (see Theorem 3 above). As a consequence, \( \hat{S}_n^{(j)}(t) \) is asymptotically Gaussian centered on \( S^{(j)} \). Given the above results, the estimated standard deviation of \( S^{(j)} \), noted \( \hat{\sigma}_S \) is given for all \( t \geq 0 \) by:

\[
\hat{\sigma}^2_S(t) = \frac{\hat{\gamma} \left( \hat{S}_n^{(j)}(t) \right)}{\left[ \hat{S}_n^{(j)}(t) \right]^2}.
\] (14)

Therefore a threshold confidence level \( 100(1-\alpha)\% \) can be built for all \( t \geq 0 \) and \( j \in \{1, \ldots, m\} \), by:

\[
\hat{S}_n^{(j)}(t) - Z_{\alpha/2} \sigma_{\hat{S}_n^{(j)}}(t) \leq S^{(j)}(t) \leq \hat{S}_n^{(j)}(t) + Z_{\alpha/2} \sigma_{\hat{S}_n^{(j)}}(t).
\] (15)

Here \( z_{\alpha/2} \) is the \( 1-\alpha/2 \) percentile of a standard normal distribution.

A threshold confidence interval \( 100(1-\alpha)\% \) can also be obtained for all \( j \in \{1, \ldots, m\} \), by:

\[
\hat{S}_n^{(j)}(t) \pm z_{\alpha/2} \hat{\sigma}_S^{(j)}(t),
\] (16)

where \( z_{\alpha/2} \) is the rank of fractile \( 100 \times \alpha/2 \) of the standardized normal distribution.

A disadvantage of the construction of the confidence interval (CI) with the previous formula is that the bound can be obtained external to the interval \([0,1]\). A solution is to consider a \( \{ g(S)^{(j)} \} \) transform via a continuous function \( g \), differentiable and invertible such that \( g \left( S^{(j)}(t) \right) \) belongs to a more wide space ideally unbounded and best approximate a Gaussian random variable. The delta method then allows for the estimation of the standard deviation of the object created by \( \hat{\sigma}^{(j)}_{g(S)} \) defined by:

\[
\hat{\sigma}^{(j)}_{g(S)}(t) = g \left( \hat{S}_n^{(j)}(t) \right) \hat{\sigma}_S^{(j)}(t).
\]

The confidence interval associated with the risk threshold \( \alpha \) is built as for all \( j \in \{1, \ldots, m\} \),

\[
g^{-1} \left( g \left( \hat{S}_n^{(j)}(t) \right) \pm z_{\alpha/2} g' \left( \hat{S}_n^{(j)}(t) \right) \hat{\sigma}_S^{(j)}(t) \right).
\]

Remark 1. It is also possible to use log, square-root or logit-type transformations in most software defined respectively by for all \( j \in \{1, \ldots, m\} \),

\[
g \left( S^{(j)}_t \right) = \log \left( \sqrt{S^{(j)}_t} \right), \quad g \left( S^{(j)}_t \right) = \sin^{-1} \left( \sqrt{S^{(j)}_t} \right), \quad g \left( S^{(j)}_t \right) = \log \left( \frac{S^{(j)}_t}{1-S^{(j)}_t} \right).
\]

4.2. The Confidence Bands

The challenge now is to find an area containing the survival function with
probability $1 - \alpha$, or a set of bounds $b_L(t)$ and $b_U(t)$ which, with probability $1 - \alpha$, contains $S^{(j)}(t)$ for all $t \in [t_L, t_U]$ and $j \in \{1, \ldots, m\}$. Among the proposed solutions, the two most commonly used are firstly Hall and Wellner ([24]) bands and secondly, strips Nair ([25]) (“equal precision bands”). If $t_k$ is the maximum time event observed in the sample, then for the Nair bands, we have the following restrictions $0 < t_L < t_U \leq t_k$, however, at other Hall-Wiener may authorize the nullity of $t_L$, let $0 \leq t_L < t_U \leq t_k$. Technically obtaining these bands is complex, and their practical utility in relation to the point intervals is not obvious.

Remark 2. The starting point uses the fact that for all $j \in \{1, \ldots, m\}$,

$$\sqrt{n} \left( \frac{\hat{S}^{(j)}(t)}{S^{(j)}(t)} - 1 \right)$$

converges to a centered Gaussian martingale. We then go through a transformation making appear a Brownian bridge $\{W^n(x), x \in [0,1]\}$, weighted by $\frac{1}{\sqrt{x(1-x)}}$ at Nair, to retrieve the suitable critical value.

In particular, because of the joined character, for a given $t$ their extent is wider than that of the corresponding point IC. In what follows we give the expressions obtained in the absence of transformation.

4.2.1. The Hall-Wellner Confidence Bands

Under the assumption of continuity of survival functions $S^{(j)}(t)$ and $C^{(j)}(t)$ respectively related to the event time and the time of censorship, Hall and Wellner show that for every $t \in [t_L, t_U]$, the IC joined the risk threshold $\alpha$ is given for all $j = 1, \ldots, m$ and $t \geq 0$ by:

$$\hat{S}^{(j)}(t) \pm b_u(x_L, x_U) n^{-\frac{1}{2}} \left[ 1 + n \hat{\sigma}^{(j)2}_{\hat{S}^{j}}(t) \right] \hat{S}^{(j)}(t),$$

where $x_L$ and $x_U$ are given by

$$x_i = \frac{n \hat{\sigma}^{(j)2}_{\hat{S}^{j}}(t)}{1 + n \hat{\sigma}^{(j)2}_{\hat{S}^{j}}(t)}, \text{ for } i = L, U$$

and $b_u(x_L, x_U)$ is bounds checking

$$\alpha = \mathbb{P} \left[ \sup_{x_L \leq x \leq x_U} \left| W^n(x) \right| > b_u(x_L, x_U) \right].$$

4.2.2. The Nair Precision Equal Bands

Using a weighted Brownian bridge will notably modify the bounds to IC. For $\alpha \in (0,1)$, $t \in [t_L, t_U]$ and all $j \in \{1, \ldots, m\}$, they are then given by:

$$\hat{S}^{(j)}(t) \pm e_u(x_L, x_U) \hat{\sigma}^{(j)}_{\hat{S}^{j}},$$

where $e_u(x_L, x_U)$ satisfies

$$\alpha = \mathbb{P} \left[ \sup_{x_L \leq x \leq x_U} \frac{\left| W^n(x) \right|}{\sqrt{x(1-x)}} > e_u(x_L, x_U) \right].$$
If we compare (12) and (14), we see that the bounds relating to Nair ([25]) bands are proportional to the bounds IC and simply correspond to a risk adjustment threshold used in the past.

5. Conclusions and Perspectives

In this paper we have studied the asymptotic normality of Nelson-Aalen and Kaplan-Meier type estimators in the presence of independent right-censorship as defined in Njamen and Ngatchou ([10], [11]) and Njamen [12] using Robeledo’s theorem that allows applying the central limit theorem to certain types of particular martingales. From the results obtained, confidence bounds for the hazard and the survival functions are provided.

As a perspective, obtaining actual data would allow us to perform numerical simulations to gauge the robustness of our obtained estimators.

Acknowledgements

We thank the publisher and the referees for their comments which allowed to raise considerably the level of this article.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Heckman, J.J. and Honoré, B.E. (1989) The Identifiability of the Competing Risks Models. *Biometrika, 77*, 325-330. [https://www.jstor.org/stable/2336666](https://www.jstor.org/stable/2336666)

[2] Commenges, D. (2017) Risques compétitifs et modèles multi-états en épidémiologie. *Revue d'épidémiologie et de santé publique Elsevier Masson, 77*, 605-611.

[3] Com-Nougué, C., Guérin, S. and Rey, A. (1999) Estimation des risques associés à des événements multiples. *Revue d'épidémiologie et de Santé Publique, 47*, 75-85.

[4] 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, 94*, 496-509. [https://www.jstor.org/stable/2670170](https://www.jstor.org/stable/2670170)

[5] Crowder, M. (2001) Classical Competing Risks. Chapman and Hall, London.

[6] Fermanian, J.D. (2003) Nonparametric Estimation of Competing Risks Models with Covariates. *Journal of Multivariate Analysis, 85*, 156-191. [https://doi.org/10.1016/S0047-259X(02)00069-6](https://doi.org/10.1016/S0047-259X(02)00069-6)

[7] Latouche, M. (2004) Modèles de régression en présence de compétition. Thèse de doctorat, Université de Paris, Paris, 6. [https://tel.archives-ouvertes.fr/tel-00129238](https://tel.archives-ouvertes.fr/tel-00129238)

[8] Geffray, S. (2009) Strong Approximations for Dependent Competing Risks with Independent Censoring with Statistical Applications. *Test, 18*, 76-95. [https://doi.org/10.1007/s11749-008-0113-y](https://doi.org/10.1007/s11749-008-0113-y)

[9] 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
[10] Njamen, N.D.A. and Ngatchou, W.J. (2014) Nelson-Aalen and Kaplan-Meier Estimators in Competing Risks. *Applied Mathematics, 5*, 765-776. [https://doi.org/10.4236/am.2014.54073](https://doi.org/10.4236/am.2014.54073)

[11] Njamen, N.D.A. and Ngatchou, W.J. (2018) Consistency of the Kaplan-Meier Estimator of the Survival Function in Competing Risks. *The Open Statistics and Probability Journal, 9*, 1-17. [https://benthamopen.com/TOSPJ/home](https://benthamopen.com/TOSPJ/home)

[12] Njamen, N.D.A. (2017) Convergence of the Nelson-Aalen Estimator in Competing Risks. *International Journal of Statistics and Probability, 6*, 9-23. [https://doi.org/10.5539/ijsp.v6n3p9](https://doi.org/10.5539/ijsp.v6n3p9)

[13] Njamen, N.D.A. (2018) Study of the Nonparametric Kaplan-Meier Estimator of the Cumulative Incidence Function in Competing Risks. *Journal of Advanced Statistics, 3*, 1-13. [https://doi.org/10.22606/jas.2018.31001](https://doi.org/10.22606/jas.2018.31001)

[14] 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. [https://www.jstor.org/stable/4615704](https://www.jstor.org/stable/4615704)

[15] Peterson, G.L. (1977) A Simplification of the Protein Assay Method of Lowry et al. Which Is More Generally Applicable. *Analytical Biochemistry, 83*, 346-356. [https://doi.org/10.1016/0003-2697(77)90043-4](https://doi.org/10.1016/0003-2697(77)90043-4)

[16] Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993) Statistical Models Based on Counting Processes. Springer Series in Statistics, Springer-Verlag, New York.

[17] Shorack, G.R. and Wellner, J.A. (1986) Empirical Processes with Applications to Statistics. John Wiley and Sons, Inc., New York.

[18] 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. [https://www.jstor.org/stable/2958131](https://www.jstor.org/stable/2958131)

[19] Nelson, W. (1972) A Short Life Test for Comparing a Sample with Previous Accelerated Test Results. *Technometrics, 14*, 175-185. [https://www.jstor.org/stable/1266929](https://www.jstor.org/stable/1266929)

[20] Aalen, O.O. (1978) Nonparametric Inference for a Family of Counting Processes. *The Annals of Statistics, 6*, 701-726. [https://www.jstor.org/stable/2958850](https://www.jstor.org/stable/2958850)

[21] Fleming, T.R. and Harrington, D.P. (1990) Counting Processes and Survival Analysis. John Wiley and Sons, Hoboken.

[22] 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. [https://tel.archives-ouvertes.fr/tel-00005524](https://tel.archives-ouvertes.fr/tel-00005524)

[23] Kaplan, E.L. and Meier, P. (1958) Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association, 53*, 457-481. [https://www.jstor.org/stable/2281868](https://www.jstor.org/stable/2281868)

[24] Hall, W.J. and Wellner, J.A. (1980) Confidence Bands for a Survival Curve. *Biometrika, 67*, 133-143. [https://www.jstor.org/stable/2335326](https://www.jstor.org/stable/2335326)
[25] Nair, V.N. (1984) Confidence Bands for Survival Functions with Censored Data: A Comparative Study. *Technometrics*, **26**, 265-275.

https://www.jstor.org/stable/1267553

https://doi.org/10.1080/00401706.1984.10487964