

# Remerciements

Je tiens à remercier sincèrement le Professeur Mourid pour tout le temps qu'il m'a consacré, pour sa patience et sa présence tout au long de la réalisation de ce mémoire ainsi que pour son aide et ses précieux conseils et ce depuis le premier cours qu'il nous a donné.

Merci à Mr Labbas qui, non seulement fut le premier à me faire aimer les statistiques, mais m'a aussi fait le grand honneur de présider ce jury.

Je remercie Mme Dali Youcef pour son aide, sa gentillesse et son soutien tout au long de ces années ainsi que d'avoir accepté d'être membre de mon jury.

J'adresse des remerciements particuliers à Mr Boukhari pour avoir accepté de faire partie de mon jury. Je lui témoigne ma profonde reconnaissance pour son aide et sa grande disponibilité.

Je remercie Mme Sahi d'avoir bien voulu m'aider dans la réalisation de ce mémoire et d'avoir accepté d'assister à ma soutenance.

J'exprime également ma gratitude à tous les enseignants qui m'ont accompagnée et aidée à chaque étape de mon cursus de formation.

Je me permets de remercier mes amis, en particulier Ilham pour son aide si précieuse et mon fiancé Mehdi pour toute la joie qu'il a apporté dans ma vie.

Mes dernières pensées et non les moindres vont vers ma famille, ma petite soeur, mes frères et surtout mes parents qui ont toujours cru en moi. J'espère rester à jamais leur fierté.

# Table des matières

<b>Remerciement</b>	<b>1</b>
<b>Introduction</b>	<b>4</b>
<b>1 Convergence faible</b>	<b>6</b>
1.1 Convergence faible dans $\mathbb{R}^k$	6
1.2 Convergence faible dans un espace métrique	7
1.2.1 Notations et définitions	7
1.2.2 Théorème de Portmanteau	9
1.2.3 Théorème de l'application continue	9
1.2.4 Théorème de Prohorov	11
1.3 Convergence faible dans $C[0,1]$	15
1.3.1 Définitions	15
1.3.2 Théorème de Donsker	18
1.4 Convergence faible dans $D[0,1]$	22
1.4.1 Définitions	22
1.4.2 Topologie de Skorohod	23
1.4.3 Théorème de Donsker	25
<b>2 Modèles de survie</b>	<b>26</b>
2.1 Généralités	26
2.1.1 Fonctions associées aux distributions de survie	26
2.1.2 Censures	27
2.1.3 Processus ponctuels associés	28
2.2 Théorème de Rebolledo	30
2.3 Estimation du taux de risque cumulé	36
2.3.1 Introduction	36
2.3.2 Estimateur de Nelson-Aalen	37
1) Construction	37
2) Exemple :	38
3) Consistance et loi asymptotique	40
2.4 Estimation de la fonction de survie	44
2.4.1 Introduction	44
2.4.2 Estimateur de Fleming-Harington	44
1) Construction	44
2) Exemple	45
3) Consistance et loi asymptotique	45
2.4.3 Estimateur de Kaplan Meier	48
1) Construction	48

2)	Exemple . . . . .	48
3)	Consistence . . . . .	49
2.5	Estimation de la densité des durées de survie . . . . .	50
2.5.1	Données non censurées . . . . .	50
1)	Introduction . . . . .	50
2)	Comportement asymptotique de $f_n$ . . . . .	50
3)	Choix optimaux du noyau K et de la fenêtre $b_n$ . . . . .	51
2.5.2	Données censurées . . . . .	52
1)	Introduction . . . . .	52
2)	Etude du biais de l'estimateur $\hat{f}_n$ . . . . .	53
3)	Consistence et normalité asymptotique de $\hat{f}_n$ . . . . .	55
4)	Erreur quadratique intégrée et fenêtre optimale . . . . .	60
<b>3</b>	<b>Simulation</b>	<b>61</b>
3.1	Estimation de la fonction de survie S . . . . .	64
3.2	Estimation du taux de hasard cumulé H . . . . .	65
3.3	Estimation de la fonction de densité . . . . .	66
3.3.1	Données non censurées . . . . .	66
3.3.2	Données censurées . . . . .	67
	<b>Annexe</b>	<b>68</b>
	<b>Conclusion</b>	<b>69</b>
	<b>Bibliographie</b>	<b>70</b>

# Introduction

L'analyse des données de survie constitue un domaine de la statistique qui s'intéresse à mesurer le temps appelé "temps de survie" jusqu'à un évènement particulier.

## Exemples :

En fiabilité industrielle : durée de fonctionnement de machines avant une défaillance

En économie : durée de grèves ou de périodes de non-emploi (chômage)

En éducation : durée d'apprentissage d'une langue

En médecine : durée de vie de patients après traitement

Ce type d'analyse est aussi particulièrement utile dans la recherche biomédicale. Par exemple, lors d'un essai thérapeutique, on teste l'efficacité d'un nouveau traitement en le comparant aux durées de survie du traitement habituel ou d'un placebo.

Le temps de survie est représenté par une variable aléatoire positive  $T$ . Une notion essentielle liée à  $T$  est la fonction de survie, définie par :

$$S(t) = P(T > t)$$

On utilise également la fonction de risque, définie par :

$$h(t) = \frac{f(t)}{S(t)}$$

où  $f$  désigne la densité de  $T$ .

Dans ce mémoire nous présentons une synthèse sur les convergences faibles sur les espaces  $C_{[0,1]}$  et  $D_{[0,1]}$  en se basant sur le livre : Billingsley P. "*Convergence of Probability Measures*" Wiley (1968).

Nous présentons aussi une étude sur les modèles de survie grâce aux livres suivants :

1. Fleming, T.R. & D.P. Harrington. "*Counting processes and survival analysis*". Wiley Series in Probability and Statistics. John Wiley & Sons (1991).
2. Bosq, D. et Lecoutre, J. "*Théorie de l'estimation fonctionnelle*". Economica, Paris (1987).
3. Zhang Biao "*Some asymptotic results for kernel estimation under random censorship*". Bernoulli 2 (2) (1996), p. 183-198.
4. Sahi, N. "*Modèles de survie Estimation et Applications*" Mémoire de magister en probabilités statistiques, (octobre 2011).

Dans le chapitre 1 nous allons introduire la convergence faible dans un espace métrique quelconque en démontrant le théorème de Prohorov. Ensuite, nous allons étudier la convergence faible dans  $C_{[0,1]}$  l'espace des fonctions continues sur  $[0, 1]$  et démontrer le théorème de Donsker. Nous terminons par la convergence faible dans  $D_{[0,1]}$  l'espace des fonctions continues à droite avec limite à gauche en étudions la topologie de Skorohod et en citant le théorème de Donsker.

Dans le chapitre 2 nous présentons une étude sur les modèles de survie en étudiant l'estimation de la fonction de survie et la fonction de risque et leurs comportements asymptotiques. Puis nous abordons l'estimation de la densité des durées de survie et son comportement asymptotique.

Dans le chapitre 3 nous présentons des simulations sur l'estimation de la fonction de survie, la fonction de risque et la densité des durées de survie. Nous ferons cela sur des données réelles qui représentent des données de survie de malades atteints de CGD (chronic granulomatous disease). Le but est de comparer l'efficacité du traitement avec l'effet Placebo.

En Annexe, nous donnons quelques énoncés des théorèmes utilisés dans ce mémoire.

# Chapitre 1

## Convergence faible

Dans ce chapitre nous présentons les principales notions de la convergence faible dans  $\mathbb{R}^k$ , dans un espace métrique quelconque, dans  $C[0, 1]$  l'espace des fonctions continues sur  $[0, 1]$  puis dans  $D[0, 1]$  l'espace des fonctions continues à droite avec une limite à gauche dans  $[0, 1]$ . On se basera sur le livre de Billingsley [1].

### 1.1 Convergence faible dans $\mathbb{R}^k$

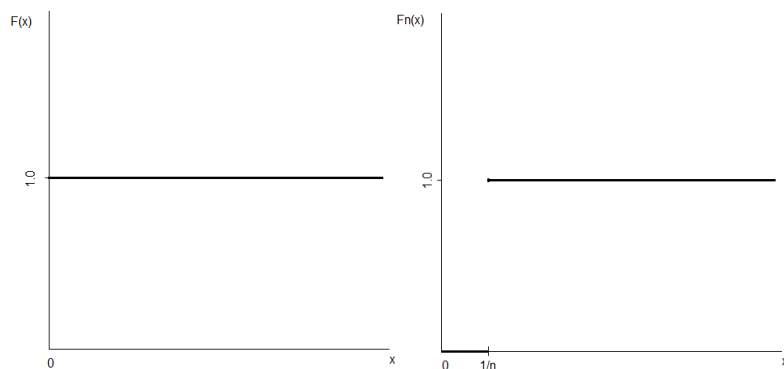
Soit  $(\Omega, A, P)$  un espace de probabilité. Toutes les v.a seront définies sur cet espace.

**Définition 1.1.1** Soit  $(X_n)$  une suite de v.a. réelles et  $X$  une v.a.r. Soit  $F_n$  et  $F$  leurs fonctions de répartition respectives sur  $\mathbb{R}$ .  $(X_n)$  converge en loi vers  $X$  ou encore  $F_n$  converge faiblement vers  $F$  ssi  $F_n(x) \rightarrow F(x)$  en tout point de continuité de  $F$ .

On note  $F_n \Rightarrow F$  ou  $X_n \xrightarrow{\mathcal{L}} X$ .

On utilise aussi :  $P_{X_n} \Rightarrow P_X$ .

**Exemple 1.1.2** Supposons  $F_n(x)$  et  $F(x)$  sont définies comme suit :



on a  $F_n \Rightarrow F$  (même si  $F_n(0) = 0 \neq 1 = F(0)$ ).

La proposition suivante précise cette notion.

**Propriété 1.1.3** Soit  $\mathbf{B}$  la tribu de Borel de  $\mathbb{R}^k$ ,  $P$  et  $P_n$  des mesures de probabilités sur  $(\mathbb{R}^k, \mathbf{B})$ . Si  $B$  est un borélien on note  $\partial B$  sa frontière. Nous avons l'équivalence suivante :

$$(P_n \Rightarrow P) \Leftrightarrow (P_n(B) \rightarrow P(B) \text{ pour tout } B \in \mathbf{B} \text{ tel que } P(\partial B) = 0)$$

## 1.2 Convergence faible dans un espace métrique

### 1.2.1 Notations et définitions

Soit  $(S, d)$  un espace métrique. Notons :

- $\mathbf{B}$  sa tribu de Borel
- $C(S)$  l'espace des fonctions réelles continues définies sur  $S$
- $C_b(S)$  l'espace des fonctions continues et bornées sur  $S$
- $\Gamma$  la tribu engendrée par les fermés

**Définition 1.2.1.1** Soit  $(P_n)$  et  $P$  des mesures de probabilité sur  $(S, \Gamma)$ . On dit que  $(P_n)$  converge faiblement vers  $P$  si pour toute fonction réelle, continue et bornée  $f$  on a :

$$\int_S f dP_n \rightarrow \int_S f dP$$

on note cette convergence par  $P_n \Rightarrow P$ .

Pour une suite de v.a.  $(X_n)$  et une v.a.  $X$  à valeurs dans  $S$  on dit que  $X_n \Rightarrow X$  si  $P_{X_n} \Rightarrow P_X$ .

On a le résultat suivant

**Propriété 1.2.1.2** Toute mesure de probabilité sur  $(S, \Gamma)$  est régulière i.e

Si  $A \in \Gamma$  et  $\varepsilon > 0$  alors il existe un fermé  $F$  et un ouvert  $O$  tel que  $F \subset A \subset O$  et  $P(O - F) < \varepsilon$

**Propriété 1.2.1.3** Soit  $P$  et  $Q$  deux probabilités définies sur  $(S, d)$ . Alors

$$(P \equiv Q) \Leftrightarrow \left( \int f dP = \int f dQ \quad \forall f \in C_b(S) \right)$$

#### Preuve

Soit  $\int f dP = \int f dQ \quad \forall f \in C_b(S)$ . Montrons que  $P \equiv Q$ .

Soit  $F$  un fermé et  $\varphi$  une fonction continue définie par :

$$\varphi(t) = \begin{cases} 1 & \text{si } t \leq 0 \\ 1 - t & \text{si } 0 \leq t \leq 1 \\ 0 & \text{si } 1 \leq t \end{cases}$$

et  $f$  définie par :

$$f(x) = \varphi\left(\frac{1}{\varepsilon}d(x, F)\right)$$

Notons, pour tout entier  $u$ ,

$$\varphi_u(t) = \varphi(ut)$$

et

$$f_u(x) = \varphi_u(d(x, F))$$

On a  $(f_u)$  est une suite décroissante d'éléments de  $C_b(S)$  qui converge ponctuellement vers  $\mathbb{1}_F$ . Par le théorème de la convergence dominée :

$$P(F) = \lim_u \int f_u dP = \int f dP$$

et

$$Q(F) = \lim_u \int f_u dQ = \int Q dP$$

Donc,

$$P(F) = Q(F) \text{ pour tout fermé } F$$

La propriété 1.2.1.2 implique  $P \equiv Q$  •

**Définition 1.2.1.4** Une famille  $\Pi$  est tendue si

$$\forall \varepsilon > 0, \exists K_\varepsilon \text{ un compact tel que } P(K) > 1 - \varepsilon, \forall P \in \Pi$$

**Définition 1.2.1.5** Une famille  $\Pi$  est relativement compacte si de toute suite  $(P_n)$  de  $\Pi$  on peut extraire une sous-suite  $(P'_n)$  et une mesure de probabilité  $Q$  définie sur  $(S, \Gamma)$  (pas forcément un élément de  $\Pi$ ) tel que  $P'_n \Rightarrow Q$ .

**Définition 1.2.1.6** Un espace  $E$  est précompact si pour tout  $\varepsilon > 0$ , on peut recouvrir  $E$  par un nombre fini de boules ouvertes de rayon  $\varepsilon$ .

On a le théorème suivant

**Théorème 1.2.1.7** Si  $S$  est séparable et complet alors toute mesure de probabilité  $P$  dans  $(S, d)$  est tendue.

**Preuve**

Soit  $S$  séparable et complet. Montrons que toute mesure de probabilité dans  $(S, d)$  est tendue i.e  $\forall \varepsilon > 0, \exists K_\varepsilon$  un compact tel que  $P(K) > 1 - \varepsilon$ .

$S$  étant séparable alors  $S$  est couvert par une famille dénombrable de boules ouvertes de rayon  $\varepsilon > 0$ .

Soit  $n \geq 1$  et  $\varepsilon = \frac{1}{n}$ . Il existe  $\{B_k^{(n)}\}_{k \geq 1}$  une famille de boules ouvertes de rayon  $\frac{1}{n}$  vérifiant :

$$S = \cup_{k \geq 1} B_k^{(n)}$$

D'autre part,

$$\begin{aligned} 1 = P(S) &= P(\cup_{k \geq 1} B_k^{(n)}) \\ &= P(\cup_{m \geq 1} \cup_{k=1}^m B_k^{(n)}) \\ &= \lim_{m \rightarrow \infty} P(\cup_{k=1}^m B_k^{(n)}) \quad , \forall n \geq 1 \end{aligned}$$

Ainsi  $\forall \varepsilon > 0, \forall n \geq 1, \exists m_n \geq 1$  tel que

$$\forall m \geq m_n, P(\cup_{k=1}^m B_k^{(n)}) > 1 - \frac{\varepsilon}{2^n}$$

D'où

$$\begin{aligned} P(\cap_{n \geq 1} \cup_{k=1}^m B_k^{(n)}) &= 1 - P(\cup_{n \geq 1} \cap_{k=1}^m \overline{B_k^{(n)}}) \\ &\geq 1 - \sum_{n \geq 1} P(\cap_{k=1}^m \overline{B_k^{(n)}}) \\ &\geq 1 - \sum_{n \geq 1} \frac{\varepsilon}{2^n} \\ &> 1 - \varepsilon \end{aligned}$$



Posons

$$K = \overline{\bigcap_{n \geq 1} \bigcup_{k=1}^m B_k^{(n)}}$$

Par conséquent

$$P(K) \geq 1 - \varepsilon$$

Comme  $\bigcap_{n \geq 1} \bigcup_{k=1}^m B_k^{(n)}$  est précompact dans  $(S, d)$  qui est complet. Donc il est relativement compact. D'où sa fermeture qui est  $K$  est un compact. •

Le résultat suivant fournit une autre caractérisation.

**Théorème 1.2.1.8** *On a  $P_n \Rightarrow P$  ssi toute sous-suite  $(P_{n_k})$  contient une sous-suite  $(P_{n_{k_j}})$  telle que  $P_{n_{k_j}} \Rightarrow P$ .*

## 1.2.2 Théorème de Portmanteau

**Théorème 1.2.2.1** *Soit  $(P_n)$  et  $P$  des mesures de probabilités sur  $(S, d)$ . Les propriétés suivantes sont équivalentes :*

1.  $P_n \Rightarrow P$
2.  $\forall f \in C_b(S) ; \int f dP_n \rightarrow \int f dP$
3.  $\forall F$  un fermé de  $S$ ;  $\limsup_{n \rightarrow \infty} P_n(F) \leq P(F)$
4.  $\forall O$  un ouvert de  $S$ ;  $\liminf_{n \rightarrow \infty} P_n(O) \geq P(O)$
5.  $\forall A \in \Gamma$  tel que  $P(\partial A) = 0$ ;  $P_n(A) \rightarrow P(A)$

## 1.2.3 Théorème de l'application continue

**Théorème 1.2.3.1** *Soit  $P_n$  et  $P$  des mesures de probabilités sur  $(S, \Gamma)$ . Soit  $h : (S, \Gamma) \rightarrow (S', \Gamma')$  une fonction mesurable et  $D_h$  l'ensemble des points de discontinuité de  $h$ . Si  $P_n \Rightarrow P$  et  $P(D_h) = 0$  alors :*

$$P_n h^{-1} \Rightarrow P h^{-1} \text{ dans } S'$$

### Preuve

Soit  $P_n \Rightarrow P$  et  $P(D_h) = 0$ . Montrons que  $P_n h^{-1} \Rightarrow P h^{-1}$  dans  $S'$  en utilisant le point (3) du théorème de Portmanteau.

Soit  $F$  un fermé de  $\Gamma'$ .

$$\begin{aligned} \limsup_{n \rightarrow \infty} (P_n h^{-1})(F) &:= \limsup_{n \rightarrow \infty} P_n(h^{-1}(F)) \\ &\leq \limsup_{n \rightarrow \infty} P_n(\overline{h^{-1}(F)}) \\ &\leq P(\overline{h^{-1}(F)}) \end{aligned}$$

D'autre part,

$$\begin{aligned}
x \in \overline{h^{-1}(F)} &\Rightarrow \exists x_n \in h^{-1}(F) \text{ tel que } x_n \rightarrow x \\
&\Rightarrow \exists x_n \in S, h(x_n) \in F \text{ et } x_n \rightarrow x \\
&\Rightarrow (\exists x_n \in S, h(x_n) \in F \text{ et } x_n \rightarrow x \wedge x \in D_h) \vee (\exists x_n \in S, h(x_n) \in F \text{ et } x_n \rightarrow x \wedge x \notin D_h) \\
&\Rightarrow x \in D_h \vee (\exists x_n, h(x_n) \in F, h(x_n) \rightarrow h(x), x \notin D_h) \\
&\Rightarrow x \in D_h \vee (\exists x_n \in S, h(x_n) \rightarrow h(x), h(x_n) \in F, h(x) \in F) \\
&\Rightarrow x \in D_h \vee h(x) \in F \\
&\Rightarrow x \in D_h \vee x \in h^{-1}(F) \\
&\Rightarrow x \in (D_h \cup h^{-1}(F))
\end{aligned}$$

D'où

$$\overline{h^{-1}(F)} \subset (D_h \cup h^{-1}(F))$$

Donc

$$\begin{aligned}
P(\overline{h^{-1}(F)}) &\leq P(D_h \cup h^{-1}(F)) \\
&\leq P(D_h) + P(h^{-1}(F)) \\
&\leq P(h^{-1}(F))
\end{aligned}$$

Ainsi, par le point (3) du théorème de Portmanteau on a

$$P_n h^{-1} \Rightarrow P h^{-1} \quad \bullet$$

**Corollaire 1.2.3.2** Soit  $(S, \Gamma)$  et  $(S', \Gamma')$  deux espaces métriques et  $h$  une fonction continue de  $(S, \Gamma)$  dans  $(S', \Gamma')$ . Si  $P_n \Rightarrow P$  alors

$$P_n h^{-1} \Rightarrow P h^{-1} \text{ dans } S'$$

**Proposition 1.2.3.3** Si  $X_n \xrightarrow{\mathcal{L}} X$  et  $P(X \in D_h) = 0$  alors

$$h(X_n) \xrightarrow{\mathcal{L}} h(X)$$

**Corollaire 1.2.3.4** Si  $X_n \xrightarrow{\mathcal{P}} a$  et si  $h$  est continue en  $a$  alors

$$h(X_n) \xrightarrow{\mathcal{P}} h(a)$$

**Théorème 1.2.3.5** .

(i) Si pour toute fonction continue bornée  $h$ ,  $P_n h^{-1} \Rightarrow P h^{-1}$  alors  $P_n \Rightarrow P$

(ii) Si  $h$  est une fonction réelle mesurable et bornée avec  $P(D_h) = 0$  et  $P_n \Rightarrow P$  alors  $\int h dP_n \rightarrow \int h dP$

**Preuve**

(i) Supposons  $P_n h^{-1} \Rightarrow P h^{-1}$ , pour tout  $h$  continue et bornée. Montrons que  $P_n \Rightarrow P$ .

$$\begin{aligned}
[P_n h^{-1} \Rightarrow P h^{-1}] &\Rightarrow \forall f \text{ continue bornée, } \int_{\mathbb{R}} f(x) d(P_n h^{-1})(x) \rightarrow \int_{\mathbb{R}} f(x) d(P h^{-1})(x) \\
&\Rightarrow \forall f \text{ continue bornée, } \int_{\mathbb{R}} f(h(x)) dP_n(x) \rightarrow \int_{\mathbb{R}} f(h(x)) dP(x)
\end{aligned}$$

$\exists M > 0$  tel que  $|h(x)| \leq M$ .

Posons  $f_0(x) = \begin{cases} -M & \text{si } x \leq -M \\ x & \text{si } -M \leq x \leq M \\ M & \text{si } x \geq M \end{cases}$

On a  $f_0$  continue et bornée. Par suite,

$$\int_{\mathbb{R}} f_0(h(x)) dP_n(x) \rightarrow \int_{\mathbb{R}} f_0(h(x)) dP(x)$$

Comme  $-M < h(x) < M$

$$\int_{\mathbb{R}} h(x) dP_n(x) \rightarrow \int_{\mathbb{R}} h(x) dP(x) \quad \forall h \in C_b$$

Donc

$$P_n \Rightarrow P$$

(ii) Supposons que  $P_n \Rightarrow P$ . Soit  $h$  fonction réelle mesurable et bornée et  $P(D_h) = 0$ . Montrons que  $\int h dP_n \rightarrow \int h dP$ .

Comme  $P_n \Rightarrow P$  et  $P(D_h) = 0$  par hypothèse, le théorème 1.2.3.1 implique  $P_n h^{-1} \Rightarrow P h^{-1}$ . En passant par les mêmes étapes que pour (i) on obtient

$$\int h dP_n \rightarrow \int h dP \quad \bullet$$

#### 1.2.4 Théorème de Prohorov

Soit  $(S, d)$  un espace métrique, la tribu de Borel sur  $S$  et  $\Pi = \{P_i, i \in I\}$  une famille de probabilités définies sur  $(S, \varphi)$ . On cherche des conditions suffisantes pour que la famille  $\Pi$  soit relativement compacte.

On a besoin du théorème suivant.

##### **Théorème 1.2.4.1 (Théorème de Helly)**

Soit  $(F_n)$  une suite de fonctions de répartition sur  $\mathbb{R}^k$  alors il existe une suite  $(n_k) \nearrow \infty$  et une fonction  $F$  définie sur  $\mathbb{R}^k$  croissante, continue à droite sur  $\mathbb{R}^k$ ,  $0 \leq F \leq 1$  et telle qu'on ait  $F_{n_k}(x) \rightarrow F(x)$  en tout point de continuité de  $F$ .

##### **Théorème 1.2.4.2 (Théorème de Prohorov)**

(i) Si  $\Pi$  est tendue alors elle est relativement compacte

(ii) Si  $S$  est séparable et complet et  $\Pi$  relativement compacte alors  $\Pi$  est tendue.

##### **Preuve**

i) Supposons  $\Pi$  tendue et montrons qu'elle est relativement compacte.

1) Cas  $S = \mathbb{R}^k$ .

Si  $(P_n)$  est une suite de  $\Pi$  alors, par le théorème de Helly,  $F_n$  la f.d.r associée à  $P_n$  admet une sous suite  $F_{n_k}$  tel que

$$F_{n_k}(x) \rightarrow F(x) \text{ aux points de continuité de } F \quad (*)$$

avec  $F$  vérifiant :

i)  $0 \leq F \leq 1$

ii)  $F$  croissante

iii)  $F$  continue à droite

Alors il existe  $\mu$  une mesure de  $(\mathbb{R}^k, \mathcal{R}^k)$  tel que  $\mu(a, b]$  définie par les valeurs de  $F$  sur les sommets du pavé  $k$ -dimensionnel  $(a, b]$ . Donc on aura  $P_{n_k} \Rightarrow \mu$ . Ainsi,  $\Pi$  sera relativement compacte si on arrive à montrer que  $\mu(\mathbb{R}^k) = 1$ . Pour cela il suffit de montrer que

$$\forall \varepsilon > 0, \exists a, b \text{ tel que } \mu(a, b] \geq 1 - \varepsilon$$

$\Pi$  étant tendue, pour  $\varepsilon > 0$ , il existe un compact  $K$  tel que  $P_{n_k}(K) > 1 - \varepsilon \quad \forall k \geq 1$ .

Comme  $F$  est une f.d.r,  $D_F$  l'ensemble des points de discontinuité de  $F$  est au plus dénombrable. Soit  $a, b$  deux points de continuité de  $F$ ,  $a \leq b$  tels que  $K \subset (a, b]$  et les  $2^k$  sommets de  $(a, b]$  sont des points de continuité de  $F$ . Puisque  $P_{n_k}(a, b]$  est définie par les valeurs de  $F_{n_k}$  sur les sommets du pavé  $(a, b]$ , alors si (\*) implique

$$P_{n_k}(a, b] \rightarrow \mu(a, b]$$

De

$$P_{n_k}(a, b] \geq P_{n_k}(K) > 1 - \varepsilon$$

on déduit

$$\mu(a, b] = \lim_{n \rightarrow \infty} P_{n_k}(a, b] \geq 1 - \varepsilon$$

Alors  $\mu(\mathbb{R}^k) = 1$ .

Donc  $\Pi$  est relativement compacte.

2) Cas  $S = \mathbb{R}^\infty$ .

On a besoin du lemme suivant :

**Lemme1** Si  $\Pi$  est tendue de  $(S, \Gamma)$  et  $h$  est une application continue de  $S$  vers  $S'$  alors :

$$\{Ph^{-1} : P \in \Pi\} \text{ est une famille tendue dans } (S', \Gamma')$$

**preuve :** Pour un  $\varepsilon$  donné on choisie dans  $S$  un compact  $K$  tel que  $P(K) > 1 - \varepsilon$  pour tout  $P$  dans  $\Pi$ . Si  $K' = hK$  alors  $K'$  est compact et  $h^{-1}K' \supset K$  et donc  $Ph^{-1}(K') > 1 - \varepsilon$  pour tout  $P$  dans  $\Pi$ . •

Si  $\Pi$  est tendue dans  $(\mathbb{R}^\infty, \mathcal{R}^\infty)$ . Par le lemme1 on a  $\{P\pi_k^{-1} : P \in \Pi\}$  est pour tout  $k$  tendue dans  $(\mathbb{R}^k, \mathcal{R}^k)$  où  $\pi_k$  est une projection sur  $\mathbb{R}^k$ . De toute suite  $(P_n)$  on peut extraire une sous-suite  $(P_{n_j})$  tel que  $P_{n_j}\pi_k^{-1}$  converge faiblement vers une mesure de probabilité  $\mu_k$  dans  $(\mathbb{R}^k, \mathcal{R}^k)$  : on peut choisir  $P_{n_j}$  tel que

$$P_{n_j}\pi_k^{-1} \Rightarrow \mu_k$$

simultanément pour tout les  $k$ . Comme  $\mu_k$  satisfait les conditions du théorème d'existence de Kolmogorov ([1] p228) alors il existe une probabilité  $Q$  dans  $(\mathbb{R}^\infty, \mathcal{R}^\infty)$  tel que

$$Q\pi_k^{-1} = \mu_k \text{ pour tout } k$$

Donc

$$P_{n_j}\pi_k^{-1} \Rightarrow Q\pi_k^{-1} \text{ pour tout } k$$

d'où les lois de dimension finie de  $P_{n_j}$  convergent vers celles de  $Q$  ce qui implique que

$$P_{n_j} \Rightarrow Q$$

On a donc tendue qui implique relativement compact dans  $\mathbb{R}^\infty$ .

3) Pour le cas  $\sigma$ -compact et le cas général on a besoin d'un autre lemme.

Soit  $S_0$  un Borélien de  $S$  :  $S_0 \in \Gamma$ . Notons  $\Gamma_0$  la classe des Boréliens de  $S_0$  :

$$\Gamma_0 = \{A : A \subset S_0, A \in \mathcal{A}\}$$

En particulier

$$\Gamma_0 \subset \Gamma$$

Si  $P$  est une mesure de probabilité dans  $(S, \Gamma)$  avec  $P(S_0) = 1$ , notons  $P^r$  la probabilité dans  $(S_0, \Gamma_0)$  obtenue par la restriction de  $\Gamma$  à  $\Gamma_0$ .

Si  $P$  est une mesure de probabilité dans  $(S_0, \Gamma_0)$  notons  $P^e$  la mesure de probabilité dans  $(S, \Gamma)$  avec  $P^e(A) = P(A \cap S_0)$  pour  $A \in \Gamma$ . On remarque que  $P^e(S_0) = 1$ .

Si  $P$  est une mesure de probabilité sur  $(S, \Gamma)$  avec  $P(S_0) = 1$  alors

$$(P^r)^e = P \tag{1.1}$$

Si  $P$  est une mesure de probabilité dans  $(S_0, \Gamma_0)$  alors

$$(P^e)^r = P \tag{1.2}$$

On admet le lemme suivant

**Lemme2 :**

- Si  $\Pi$  est une famille tendue dans  $(S_0, \Gamma_0)$  alors  $\Pi^e = \{P^e : P \in \Pi\}$  est une famille tendue dans  $(S, \Gamma)$ .
- Si  $P_n \Rightarrow P$  dans  $(S_0, \Gamma_0)$  alors  $P_n^e \Rightarrow P^e$  dans  $(S, \Gamma)$ .
- Si  $P_n \Rightarrow P$  dans  $(S, \Gamma)$  et  $P_n(S_0) = P(S_0) = 1$  alors  $P_n^r \Rightarrow P^r$  dans  $(S_0, \Gamma_0)$ .

### 3.1) Cas $S$ $\sigma$ -compact

Si  $S$  est  $\sigma$ -compact alors il est séparable donc on peut lui associer un homéomorphisme dans  $\mathbb{R}^\infty$  et son image par cet homéomorphisme est un  $\sigma$ -compact en particulier cette image est un Borélien de  $\mathbb{R}^\infty$ . Par le théorème 1.2.3.1 la convergence faible et la compacité relative sont vérifiées sous l'homéomorphisme. Comme la compacité est vraie aussi sous l'homéomorphisme alors la tension reste aussi. Donc  $S$  peut être remplacé par son image par cet homéomorphisme.

$S$  est un Borélien de  $\mathbb{R}^\infty$ . Si  $\Pi$  est tendue dans  $(S, \Gamma)$  alors, en appliquant le lemme 2 à  $\mathbb{R}^\infty$  et son sous-ensemble  $S$ ,  $\Pi^e$  est tendue dans  $(\mathbb{R}^\infty, \mathcal{R}^\infty)$ . On a vu que le théorème est vérifié dans le cas  $\mathbb{R}^\infty$  donc  $\Pi^e$  est relativement compacte. Donc, de toute suite  $(P_n)$  de  $\Pi$  qui lui correspond la suite  $(P_n^e)$  on peut extraire une sous-suite qui lui correspond  $(P_{n'}^e)$  telle que  $(P_{n'}^e) \Rightarrow Q$ . Par la tension de  $\Pi$ , pour tout  $\varepsilon > 0$  et tout  $n'$  il existe un compact  $K \subset S$  tel que

$$P_{n'}^e(K) := P_{n'}(K) > 1 - \varepsilon$$

D'où

$$Q(S) \geq Q(K) \geq \limsup_{n'} P_{n'}^e(K) \geq 1 - \varepsilon$$

Ainsi par le lemme2 et (1.2),  $P_{n'}$  converge faiblement vers  $Q^r$ . On conclut que la tension implique la compacité relative dans le cas  $\sigma$ -compact.

### 3.2) Cas général

Pour tout espace  $S$ , si  $K_i$  un compact de  $S$ ,  $S_0 = \cup_i K_i$  tel que  $P(K_i) > 1 - \frac{1}{i}$  pour tout  $P$  dans  $\Pi$  alors  $S_0$  contient le support de tout éléments de  $\Pi$  et  $\Pi^r = \{P^r : P \in \Pi\}$  est une famille tendue dans  $(S_0, \Gamma_0)$ . Comme  $\Pi^r$  est relativement compact, pour toute suite  $(P_n)$  de  $\Pi$  qui lui correspond  $(P_n^r)$  contient une sous-suite  $(P_{n'})$  qui converge faiblement vers  $Q$ . Par le lemme2 et (1.1)  $P_{n'}$  converge faiblement vers  $Q^e$ . Donc  $\Pi$  est relativement compact.

ii) Pour l'implication inverse (cf [1] p204)      •

## 1.3 Convergence faible dans $C[0,1]$

### 1.3.1 Définitions

Cette partie concerne la convergence faible dans l'espace  $C = C[0, 1]$  l'espace des fonctions réelles continues sur l'intervalle  $[0, 1]$  muni de la distance : pour  $x$  et  $y$  dans  $C$  :

$$\rho(x, y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)|$$

Le module de continuité d'un élément  $x$  dans  $C$  est défini par :

$$w_x(\delta) = w(x, \delta) = \sup_{|s-t| < \delta} |x(s) - x(t)|, 0 < \delta \leq 1$$

Le théorème suivant caractérise les compacts de  $C$ .

**Théorème 1.3.1.1 (Théorème d'Arzelà-Ascoli)** *Un sous-ensemble  $A$  est compact dans  $C$  si et seulement si :*

- (i)  $\sup_{x \in A} |x(0)| < \infty$
- (ii)  $\lim_{\delta \rightarrow 0} \sup_{x \in A} w_x(\delta) = 0$

Nous aurons besoin des résultats suivants.

**Théorème 1.3.1.2** *Soit  $(X_n)$  et  $(X'_n)$  des suites de v.a indépendantes tel que  $X_n \Rightarrow Y$  et  $X'_n \Rightarrow Y'$ . Alors*

$$(X_n, X'_n) \xrightarrow{\mathcal{L}} (Y, Y')$$

**Théorème 1.3.1.3** *Soit  $(X_n)$  et  $(Y_n)$  des suites de v.a. et  $X$  une v.a. à valeurs dans  $C$ . Si  $X_n \Rightarrow X$  et  $\rho(X_n, Y_n) \xrightarrow{\mathbf{P}} 0$  alors*

$$Y_n \Rightarrow X$$

**Théorème 1.3.1.4** *Soit  $\pi_{t_1, \dots, t_k}$  les projections définies par :*

$$\begin{aligned} \pi_{t_1, \dots, t_k} : (C, \mathcal{C}, P) &\rightarrow (\mathbb{R}^k, \mathcal{B}^k, P\pi_{t_1, \dots, t_k}^{-1}) \\ X(\omega) &\equiv x \mapsto \{x(t_1), \dots, x(t_k)\} \end{aligned}$$

*Soit  $P$  et  $Q$  deux probabilités sur  $(C, \mathcal{C})$ .*

*Supposons que les lois de dimensions finies vérifient : pour tout  $k$  et  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq 1$ ,*

$$P\pi_{t_1, \dots, t_k}^{-1} = Q\pi_{t_1, \dots, t_k}^{-1}$$

*alors*

$$P = Q$$

Nous avons le théorème fondamental suivant.

**Théorème 1.3.1.5** *Soit  $(P_n)$  et  $P$  des probabilités sur  $(C, \mathcal{C})$ . Alors  $P_n \Rightarrow P$  si et seulement si :*

- (i)  $(P_n)$  est tendue

(ii) Les lois de dimensions finies de  $P_n$  converge faiblement vers celles de  $P$  : pour tout  $k \geq 1$

$$P_n \pi_{t_1, \dots, t_k}^{-1} \Rightarrow P \pi_{t_1, \dots, t_k}^{-1}$$

**Preuve**

Soit  $(P_n)$  tendue et  $P_n \pi_{t_1, \dots, t_k}^{-1} \Rightarrow P \pi_{t_1, \dots, t_k}^{-1}$ . Montrons que  $P_n \Rightarrow P$ .

On sait que de toute suite  $P_{n_k}$  on peut extraire une sous-suite  $P_{n_{k_j}}$  qui converge faiblement vers une limite  $Q$ . Alors par le théorème de l'application continue et comme les projections sont continues on a : pour tout  $(t_1, t_2, \dots, t_k)$  :

$$P_{n_{k_j}} \pi_{t_1, \dots, t_k}^{-1} \Rightarrow Q \pi_{t_1, \dots, t_k}^{-1}$$

par l'hypothèse (ii) on a :  $P_{n_{k_j}} \pi_{t_1, \dots, t_k}^{-1} \Rightarrow P \pi_{t_1, \dots, t_k}^{-1}$  D'où

$$Q \pi_{t_1, \dots, t_k}^{-1} = P \pi_{t_1, \dots, t_k}^{-1}$$

Donc, le théorème 1.3.1.4 donne

$$Q = P$$

Ainsi toute suite  $(P_{n_k})$  contient une sous-suite  $P_{n_{k_j}}$  qui converge faiblement vers  $P$ . Par le théorème 1.2.1.8 on a

$$P_n \Rightarrow P \quad \bullet$$

Le résultat suivant fournit des critères pour qu'une suite de probabilités soit tendue dans  $C$ .

**Théorème 1.3.1.6** Soit  $(P_n)$  une suite de mesures de probabilités dans  $(C, \mathcal{C})$ .

On dit que  $(P_n)$  est tendue si et seulement si :

(i)  $\forall \eta \geq 1, \exists a$  tel que :

$$P_n \{x : |x(0)| > a\} \leq \eta,$$

(ii)  $\forall \varepsilon$  et  $\eta > 0 \exists \delta, 0 < \delta < 1$  et  $n_0$  tels que :  $\forall n \geq n_0$

$$P_n \{x : w_x(\delta) \geq \varepsilon\} \leq \eta,$$

**Preuve**

1. Soit  $(P_n)$  tendue et montrons les deux conditions (i) et (ii).

Soit  $\varepsilon > 0, \eta > 0$  et  $K$  un compact de  $C$  tel que  $\forall n P_n(K) > 1 - \eta$ . Par le théorème d'Arzelà-Ascoli  $K \subset \{x : |x(0)| \leq a\}$  pour  $a$  assez grand et aussi  $K \subset \{x : w_x(\delta) < \varepsilon\}$  pour  $\delta$  assez petit.

D'où

$$P_n(K) \leq P_n \{x : |x(0)| \leq a\}$$

donc

$$1 - P_n(K) > P_n \{x : |x(0)| > a\}$$

et alors

$$P_n \{x : |x(0)| > a\} \leq \eta$$

De même, on trouve

$$P_n \{x : w_x(\delta) \geq \varepsilon\} \leq \eta$$



2. Soit  $(P_n)$  vérifiant les deux conditions (i) et (ii). Montrons que  $(P_n)$  est tendue. Soit  $\eta > 0$  et choisissons  $a$  tel que si

$$A = \{x : |x(0)| \leq a\}$$

alors, pour tout  $n$

$$P_n(A) \geq 1 - \frac{1}{2}\eta$$

De même on choisit  $\delta_k$  tel que si

$$A_k = \{x : w_x(\delta_k) \leq \frac{1}{k}\}$$

alors, pour tout  $n$

$$P_n(A_k) \geq 1 - \frac{\eta}{2^{k+1}}$$

Posons

$$K = \overline{A \cap (\bigcap_{k \geq 1} A_k)}$$

Par définition des ensembles  $A$  et  $A_k$  on a  $A \cap (\bigcap_{k \geq 1} A_k)$  qui vérifie les conditions du théorème d'Arzela-Ascoli. Donc  $K$  est compact dans  $C$ . Il reste à montrer que  $P_n(K) \geq 1 - \eta$ .

On a

$$\begin{aligned} P_n(K^C) &= P_n[\overline{A \cap (\bigcap_{k \geq 1} A_k)}^C] \leq P_n[(\overline{A})^C \cup (\bigcup_{k \geq 1} \overline{A_k})^C] \\ &\leq P_n(\overline{A})^C + \sum_{k \geq 1} P(\overline{A_k})^C \\ &\leq \frac{1}{2}\eta + \sum_{k \geq 1} \frac{\eta}{2^{k+1}} \\ &\leq \eta \end{aligned}$$

Par suite

$$P_n(K) \geq 1 - \eta$$

Donc  $(P_n)$  est tendue. •

Un autre critère de tension dans  $C$ .

**Corollaire 1.3.1.7** *la suite  $(P_n)$  est tendue si et seulement si :*

(i)  $\forall \eta, \exists a$  tel que :

$$P_n\{x : |x(0)| > a\} \leq \eta, \eta \geq 1$$

(ii) pour tout  $\varepsilon$  et  $\eta, \exists \delta, 0 < \delta < 1$  et  $n_0$  tel que :

$$\frac{1}{\delta} P_n\{x : \sup_{t \leq s \leq t+\delta} |x(s) - x(t)| \geq \varepsilon\} \leq \eta, n \geq n_0$$

### 1.3.2 Théorème de Donsker

On appelle la mesure de Wiener la loi de probabilité du mouvement brownien standard  $W$  sur  $C$  et elle est notée  $P_W$ .

Soit  $(e_n)$  une suite de v.a. indépendantes définie sur  $(\Omega, \mathcal{B}, P)$  et à valeurs dans  $C[0, 1]$ . Soit  $S_n = e_1 + \dots + e_n$  les sommes partielles et  $(X_n)$  une suite de v.a. à valeurs dans  $C[0, 1]$  définie par :

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} e_{[nt]+1}(\omega)$$

Le résultat suivant donne la tension de  $(P_{X_n})$  ou de  $(X_n)$ .

**Proposition 1.3.2.1** *La suite  $(X_n)$  est tendue si  $\forall \varepsilon \geq 0, \exists \lambda > 1$  et un entier  $n_0$  tel que pour tout  $n \geq n_0$  on a  $\forall k$*

$$P\{\max_{i \leq n} |S_{k+i} - S_k| \geq \lambda \sigma \sqrt{n}\} \leq \frac{\varepsilon}{\lambda^2}$$

**Théorème 1.3.2.2 (Théorème de Donsker)** *Soit  $(e_n)$  des v.a. i.i.d. sur  $(\Omega, \mathcal{B}, P)$  et*

$$E(e_n) = 0, \quad E(e_n^2) = \sigma^2$$

*Soit  $S_n = e_1 + \dots + e_n$  et  $(X_n)$  une suite de v.a. à valeurs dans  $C[0, 1]$  définie par :*

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} e_{[nt]+1}(\omega)$$

Alors

$$P_{X_n} \Rightarrow P_W$$

#### Preuve

1. Montrons d'abord que les lois de dimensions finies de  $X_n$  convergent vers celles de  $W$ .

i) Considérons un point  $s$ , et montrons que

$$X_n(s) \xrightarrow{\mathcal{L}} W_s$$

Par définition de  $X_n$  on a

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}}| \leq \left| \frac{1}{\sigma\sqrt{n}} e_{[ns]+1} \right|$$

Par l'inégalité de Chebyshev,

$$\begin{aligned} P\left(\left| \frac{1}{\sigma\sqrt{n}} e_{[ns]+1} \right| > \varepsilon\right) &= P(|e_{[ns]+1}| > \varepsilon \sigma \sqrt{n}) \\ &\leq \frac{1}{n\varepsilon^2} \rightarrow 0 \end{aligned}$$

D'où

$$\frac{1}{\sigma\sqrt{n}} e_{[ns]+1} \xrightarrow{\mathcal{P}} 0$$

Donc

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]}| \xrightarrow{\mathcal{P}} 0$$

Par le théorème central limite, on a aussi

$$\frac{1}{\sigma\sqrt{n}}S_{[ns]} \xrightarrow{\mathcal{L}} W_s \hookrightarrow N(o, s)$$

D'où, par le théorème 1.3.1.3

$$X_n(s) \xrightarrow{\mathcal{L}} W_s$$

ii) Pour deux points  $s$  et  $t$  avec  $s < t$ , montrons que

$$(X_n(s), X_n(t)) \xrightarrow{\mathcal{L}} (W_s, W_t)$$

En effet du

$$\frac{1}{\sigma\sqrt{n}}S_{[ns]} \xrightarrow{\mathcal{L}} W_s, \quad \frac{1}{\sigma\sqrt{n}}S_{[nt]} \xrightarrow{\mathcal{L}} W_t$$

et comme les  $e_n$  sont indépendants, et  $S_{[ns]}$  et  $S_{[nt]} - S_{[ns]}$  sont aussi indépendants on en déduit du théorème 1.3.1.2

$$\left(\frac{1}{\sigma\sqrt{n}}S_{[ns]}, \frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})\right) \xrightarrow{\mathcal{L}} (W_s, W_t - W_s)$$

Par suite, puisque

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}}S_{[ns]}| \xrightarrow{\mathcal{P}} 0$$

on obtient

$$\left(\frac{1}{\sigma\sqrt{n}}S_{[ns]}, \frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})\right) - (X_n(s), X_n(t) - X_n(s)) \xrightarrow{\mathcal{P}} 0$$

et donc

$$(X_n(s), X_n(t) - X_n(s)) \xrightarrow{\mathcal{L}} (W_s, W_t - W_s)$$

D'où

$$(X_n(s), X_n(t)) \xrightarrow{\mathcal{L}} (W_s, W_t)$$

iii) Le même raisonnement se fait pour trois points ou plus. D'où, les lois de dimensions finies de  $X_n$  convergent vers celles de  $W$ .

2. Montrons que  $(X_n)$  est tendue.

Nous avons besoin du lemme suivant :

**Lemme**

Soit  $e_1, \dots, e_n$  des v.a. indépendantes de moyenne 0 et de variance  $\sigma_i^2$ . Posons  $S_i = e_1 + \dots + e_i$  et  $s_i^2 = \sigma_1^2 + \dots + \sigma_i^2$ . Alors

$$P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} \leq 2P\{|S_m| \geq (\lambda - \sqrt{2})s_m\}$$

**Preuve :**

L'inégalité est triviale pour  $\lambda \leq \sqrt{2}$ .

Pour  $\lambda \geq \sqrt{2}$ , considérons l'ensemble

$$E_i = \{\max_{j < i} |S_j| < \lambda s_m \leq |S_i|\}$$

On a

$$\begin{aligned} P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} &= P\{\bigcup_{i=1}^m E_i\} \\ &= \sum_{i=1}^{m-1} P(E_i) + P\{E_m\} \\ &= \sum_{i=1}^{m-1} P(E_i \cap \Omega) + P(\{|S_m| \geq \lambda s_m\} \cap \Omega) \end{aligned}$$

Ce qui donne

$$\begin{aligned} P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} &= \sum_{i=1}^{m-1} P(\{E_i\} \cap \{|S_m| \geq (\lambda - \sqrt{2})s_m\}) + \sum_{i=1}^{m-1} P(\{E_i\} \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}) \\ &+ P(\{|S_m| \geq \lambda s_m\} \cap \{|S_m| \geq (\lambda - \sqrt{2})s_m\}) + P(\{|S_m| \geq \lambda s_m\} \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}) \end{aligned}$$

D'où

$$P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} \leq \sum_{i=1}^{m-1} P(\{E_i\} \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}) + P(\{|S_m| \geq (\lambda - \sqrt{2})s_m\})$$

On a que

$$\begin{aligned} \sum_{i=1}^{m-1} P(\{E_i\} \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}) &= \sum_{i=1}^{m-1} P(\{E_i\} \cup \{|S_i| \geq \lambda s_m\} \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}) \\ &\leq \sum_{i=1}^{m-1} P(\{E_i\})P(\{|S_m| - |S_i| < \sqrt{2}s_m\}) \\ &\leq \sum_{i=1}^{m-1} P(\{E_i\}) \frac{1}{2s_m^2} \sum_{k=i+1}^m \sigma_k^2 \\ &\leq \frac{1}{2} \sum_{i=1}^{m-1} P(E_i) \\ &\leq \frac{1}{2} P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} \end{aligned}$$

Donc

$$P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} \leq \frac{1}{2} P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} + P(\{|S_m| \geq (\lambda - \sqrt{2})s_m\})$$

D'où

$$P\{\max_{i \leq m} |S_i| \geq \lambda s_m\} \leq 2P\{|S_m| \geq (\lambda - \sqrt{2})s_m\} \quad \circ$$

Revenons à la preuve de notre théorème. Appliquons le lemme aux v.a.  $e_n$  de l'énoncé. Pour  $\lambda > 2\sqrt{2}$ , on a

$$|S_n| \geq (\lambda - \sqrt{2}) \Rightarrow |S_n| \geq \frac{1}{2}\lambda$$

D'où

$$P\{\max_{i \leq n} |S_i| \geq \lambda\sigma\sqrt{n}\} \leq 2P\{|S_n| \geq \frac{1}{2}\lambda\sigma\sqrt{n}\}$$

Par le théorème central limite et l'inégalité de Chebyshev on a

$$P\{|S_n| \geq \frac{1}{2}\lambda\sigma\sqrt{n}\} \rightarrow P\{|N| \geq \frac{1}{2}\lambda\} < \frac{8}{\lambda^3}E\{|N|^3\}$$

où  $N$  est un v.a.  $N(0,1)$ . Par suite, pour  $\varepsilon > 0$  il existe  $\lambda > 1$  tel  $\frac{8}{\lambda}E\{|N|^3\} < \varepsilon$  et alors

$$\limsup_{n \rightarrow \infty} P\{\max_{i \leq n} |S_i| \geq \lambda\sigma\sqrt{n}\} < \frac{\varepsilon}{\lambda^2}$$

Par la proposition 1.3.2.1, on a la tension de  $(P_{X_n})$ .

De 1 et 2, par le théorème 1.3.1.5,

$$X_n \xrightarrow{\mathcal{L}} W \quad \bullet$$

## 1.4 Convergence faible dans $D[0,1]$

### 1.4.1 Définitions

Dans cette partie nous étudions la convergence faible dans  $D[0,1]$ . Soit  $D = D[0,1]$  l'espace des fonctions  $x$  définies sur  $[0,1]$  qui sont continues à droite avec limite à gauche :

– (i) pour  $0 \leq t < 1$ ,  $x(t^+) = \lim_{s \downarrow t} x(s)$  existe et  $x(t^+) = x(t)$

– (ii) pour  $0 < t \leq 1$ ,  $x(t^-) = \lim_{s \uparrow t} x(s)$  existe

Une fonction  $x$  a une discontinuité du premier type si  $x(t^+)$  et  $x(t^-)$  existent et sont différents de  $x(t)$ . Les discontinuités dans  $D$  sont du premier type.

Pour  $x \in D$  et  $T_0 \subset [0,1]$  posons :

$$w_x(T_0) = \sup\{|x(s) - x(t)| : s, t \in T_0\}$$

Le module de continuité d'un élément  $x$  de  $D$  est défini par :

$$w_x(\delta) = \sup_{0 \leq t \leq 1-\delta} w_x[t, t + \delta]$$

Nous aurons besoin des lemmes suivants.

**Lemme 1.4.1.1** *Pour  $x$  dans  $D$  et pour  $\varepsilon$  positif il existe des points  $t_0, t_1, \dots, t_r$  tels que :*

$$0 = t_0 < t_1 < \dots < t_r = 1$$

et

$$w_x[t_{i-1}, t_i] < \varepsilon, \quad i = 1, 2, \dots, r$$

**Lemme 1.4.1.2** *On définit le module de continuité suivant (jouant dans  $D$  le même rôle que celui dans  $C$ ).*

*Pour  $0 < \delta < 1$*

$$w'_x(\delta) = \inf_{t_i} \max_{0 < i \leq r} w_x[t_i, t_i]$$

avec  $t_i$  vérifiant  $\begin{cases} 0=t_0 < t_1 < \dots < t_r = 1 \\ t_i - t_{i-1} > \delta, \quad i = 1, 2, \dots, r \end{cases}$

**Conséquences** On a

- $\sup_t |x(t)| < \infty$
- Le lemme 1.4.1.1 est équivalent à

$$\lim_{\delta \rightarrow 0} w'_x(\delta) = 0$$

- pour  $\delta < \frac{1}{2}$ , l'intervalle  $[0,1]$  peut être décomposé en sous-intervalles  $[t_{i-1}, t_i]$  avec  $\delta < t_i - t_{i-1} \leq 2\delta$  et on a

$$w'_x(\delta) \leq w_x(2\delta), \quad \text{si } \delta < \frac{1}{2}$$

- Si  $x \in C$

$$w_x(\delta) \leq 2w'_x(\delta)$$

## 1.4.2 Topologie de Skorohod

### Distance de $D$

Soit  $\Lambda$  la classe des applications continues et strictement croissantes de  $[0, 1]$  à  $[0, 1]$  vérifiant  $\lambda(0) = 0$  et  $\lambda(1) = 1$ . Pour tout  $x$  et  $y$  dans  $D$  on définit  $d(x, y)$  l'infimum des  $\varepsilon > 0$  pour lesquels il existe  $\lambda \in \Lambda$  tel que

$$\sup_t |\lambda t - t| \leq \varepsilon$$

et

$$\sup_t |x(t) - y(\lambda t)| \leq \varepsilon$$

Montrons que  $d$  est une distance.

- Si  $d(x, y) = 0$  alors  $\forall t$  on a  $\lambda(t) = t$  et  $x(t) = y(\lambda(t))$ .  
D'où  $\forall t, x(t) = y(t)$ . Donc  $x = y$ .
- Soit  $\lambda \in \Lambda$  et  $\lambda^{-1} \in \Lambda$  l'application inverse de  $\lambda$  et qui sont strictement croissantes sur  $[0, 1]$ . Donc

$$\sup_t |\lambda^{-1}t - t| = \sup_t |\lambda t - t|$$

et

$$\sup_t |x(\lambda^{-1}t) - y(t)| = \sup_t |x(t) - y(\lambda t)|$$

D'où,  $d(x, y) = d(y, x)$

- Comme  $\lambda_1$  et  $\lambda_2$  sont à valeurs dans  $D$  alors  $\lambda_2 \lambda_1$  l'est aussi. D'où

$$\sup_t |\lambda_2 \lambda_1 t - t| \leq \sup_t |\lambda_1 t - t| + \sup_t |\lambda_2 t - t|$$

et

$$\sup_t |x(t) - z(\lambda_2 \lambda_1 t)| \leq \sup_t |x(t) - y(\lambda_1 t)| + \sup_t |y(t) - z(\lambda_2 t)|$$

Ainsi  $d$  est une distance et elle définit la topologie de Skorohod.

### Convergence au sens de Skorohod

Une suite  $(x_n)$  de  $D$  converge vers une limite  $x$  au sens de la topologie de Skorohod (on notera  $x_n \xrightarrow{S} x$ ) si  $d(x_n, x) \rightarrow 0$ . Par définition de  $d$  ceci est équivalent à dire qu'il existe une fonction  $\lambda_n$  dans  $\Lambda$  tel que

$$\lim_n x_n(\lambda_n t) = x(t) \text{ uniformément en } t$$

et

$$\lim_n \lambda_n t = t \text{ uniformément en } t$$

On a le corollaire suivant.

#### **Corollaire 1.4.2.1** .

- Si  $x_n$  converge uniformément vers  $x$  alors  $x_n \xrightarrow{S} x$ .
- $x_n \xrightarrow{S} x$  n'implique pas forcément que  $x_n$  converge uniformément vers  $x$ .

*Exemple : Soit  $x_n = \mathbb{1}_{[0, \frac{1}{2} + \frac{1}{n}]}$  et  $x = \mathbb{1}_{[0, \frac{1}{2}]}$*

*On a  $x_n \xrightarrow{S} x$  mais  $x_n$  ne converge pas vers  $x$  en  $t = \frac{1}{2}$*

- $x_n \xrightarrow{S} x$  implique la convergence uniforme pour les points de continuité de  $x$ .  
Si  $x$  est continue sur  $[0, 1]$  alors la convergence uniforme est équivalente à celle au sens de Skorohod car

$$|x_n(t) - x(t)| \leq |x_n(t) - x(\lambda_n^{-1}t)| + |x(\lambda_n^{-1}t) - x(t)|$$

**Remarque 1.4.2.2** L'espace  $D$  muni de la distance  $d$  n'est pas complet.

En effet, si

$$x_n = \mathbb{1}_{[\frac{1}{2}, \frac{1}{2} + \frac{1}{n})}$$

alors

$$d(x_n, x_m) = \left| \frac{1}{n} - \frac{1}{m} \right|$$

Ainsi  $(x_n)$  est de Cauchy mais elle ne converge pas dans  $D$ . Donc l'espace  $(D, d)$  n'est pas complet

On doit donc chercher une distance équivalente à  $d$  pour laquelle  $D$  serait complet.

### Distance pour laquelle $D$ est complet

Soit  $\lambda$  une fonction croissante dans  $[0, 1]$  avec  $\lambda(0) = 0$  et  $\lambda(1) = 1$ . Posons

$$\|\lambda\| = \sup_{s \neq t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|$$

On définit  $d_0(x, y)$  l'infimum des  $\varepsilon$  pour lesquels il existe  $\lambda \in \Lambda$  tel que

$$\|\lambda\| \leq \varepsilon$$

et

$$\sup_t |x(t) - y(\lambda t)| \leq \varepsilon$$

On a :

### **Corollaire 1.4.2.3** .

- $d_0$  est une distance.
- $(D, d_0)$  est un espace complet.
- les distances  $d$  et  $d_0$  sont équivalentes.

### Convergence faible dans $D$

Comme l'espace  $(D, d_0)$  est séparable et complet, une famille de probabilités sur  $D$  est relativement compacte si et seulement si elle est tendue (Th. Prokorov). Mais le fait que les projections ne Soit pas continues complique un peu les choses.

Soit  $T_p$  l'ensemble des  $t \in [0, 1]$  où la projection  $\pi_t$  est continue. On a le résultat suivant sur l'ensemble  $T_p$

### **Corollaire 1.4.2.4** .

- $0$  et  $1$  appartiennent toujours à  $T_p$ .
- Si  $0 < t < 1$  alors  $t \in T_p$  si et seulement si  $P(J_t) = 0$  avec

$$J_t = \{x : x(t) \neq x(t-)\}$$

et  $P$  est une probabilité sur  $D$ .

- Pour  $0 < t < 1$ ,  $\pi_t$  est continue en  $x$  si et seulement si  $x$  est continue en  $t$ .



-  $P(J_t) > 0$  est possible au plus pour un nombre dénombrable de points  $t$ .

Le théorème suivant donne la convergence faible dans  $D$ .

**Théorème 1.4.2.5** Si  $(P_n)$  est tendue et si  $P_n \pi_{t_1 \dots t_k}^{-1} \Rightarrow P \pi_{t_1 \dots t_k}^{-1}$  pour tout  $t_1, \dots, t_k$  dans  $T_p$  alors

$$P_n \Rightarrow P$$

Le résultat suivant donne un critère de tension dans  $D$ .

**Théorème 1.4.2.6**  $(P_n)$  est tendue si et seulement si

i  $\forall \eta > 0, \exists a > 0$  tel que

$$P_n \{x : |x(0)| > a\} \leq \eta, \quad n \geq 1$$

ii  $\forall \varepsilon > 0$  et  $\eta > 0, \exists n_0$  tel que

$$P_n \{x : w_x(\delta) \geq \varepsilon\} \leq \eta, \quad n \geq n_0$$

### 1.4.3 Théorème de Donsker

Dans ce résultat nous avons un exemple de suite de v.a. qui converge faiblement dans  $D$ .

**Théorème 1.4.3.1** Soit  $(e_n)_n$  une suite de variables aléatoires identiquement distribuées tel que :

$$E(e_n) = 0 \quad E(e_n^2) = \sigma^2$$

Soit  $S_n = e_1 + e_2 + \dots + e_n$  et  $X_n$  définie par :

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega)$$

Alors

$$P_{X_n} \Rightarrow P_W$$

où  $P_W$  est la loi de probabilité du mouvement brownien standard  $W$  sur  $C$ .

#### **Preuve :**

La preuve repose sur les étapes suivantes :

1. Les lois de dimensions finies de  $(X_n)$  convergent vers celles de  $W$  (qui se fait comme dans la preuve du théorème de Donsker dans  $C[0, 1]$ )
2. Pour montrer la tension de la suite  $(P_{X_n})$  on doit utiliser le théorème 2.5.2.6 •

# Chapitre 2

## Modèles de survie

Dans ce chapitre nous présentons des résultats qui sont dans [2],[3],[4] et [5],

### 2.1 Généralités

On note  $T$  une variable aléatoire positive qui est la durée de survie jusqu'à l'arrivée d'un évènement  $E$  (l'unique évènement d'intérêt).

#### 2.1.1 Fonctions associées aux distributions de survie

La loi de probabilité de la durée de survie  $T$  peut être définie par l'une des fonctions équivalentes suivantes :

1. **La fonction de répartition :**

$$F(t) = P(T \leq t)$$

Pour  $t$  fixé, c'est la probabilité que l'évènement survienne avant l'instant  $t$

2. **La fonction de survie :**

$$S(t) = P(T > t); t \geq 0$$

Pour  $t$  fixé  $S(t)$  désigne la probabilité que l'individu survive jusqu'à l'instant  $t$

On a :

- $S(t) = 1 - F(t)$
- $S$  est continue à droite, i.e  $S(t^+) = S(t)$ , pour  $t > 0$
- $S$  est décroissante.  $S(0) = 1$  et  $\lim_{t \rightarrow \infty} S(t) = 0$

3. **La densité de probabilité :**

Soit  $T$  de densité  $f(t)$  et pour tout  $t \geq 0$

$$F(t) = \int_0^t f(s) ds$$

Si la fonction de répartition a une dérivée au point  $t$  alors

$$f(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt} = F'(t) = -S'(t)$$

Pour  $t$  fixé, la densité de probabilité caractérise la probabilité que l'évènement  $E$  survienne dans un petit intervalle de temps après l'instant  $t$

#### 4. La fonction de risque (ou taux de hasard) :

Le taux de hasard est défini par :

$$\begin{aligned} h(t) &= \lim_{dt \rightarrow 0} \frac{P(t < T \leq t + dt | T > t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{1}{dt} \frac{P(t < T \leq t + dt)}{P(T > t)} \\ &= \frac{f(t)}{S(t)} \end{aligned}$$

#### 5. Le taux de hasard cumulé :

Il est défini par :

– Si  $F$  est continue :

$$H(t) = \int_0^t h(u) du$$

– Si  $F$  est quelconque :

$$H(t) = \int_0^t \frac{dF(s)}{1 - F(s^-)}$$

#### L'équivalence entre ces fonctions :

$$\begin{aligned} S(t) &= \exp[-H(t)] = \exp\left[-\int_0^t h(u) du\right] \\ H(t) &= -\ln(S(t)) \\ f(t) &= h(t)S(t) = h(t)\exp[-H(t)] \\ F(t) &= 1 - S(t) = 1 - \exp[-H(t)] \end{aligned}$$

### 2.1.2 Censures

Dans plusieurs situations les observations sont "censurées" : Une donnée est dite censurée si elle n'est pas observée. Il existe plusieurs mécanismes de censures. Particulièrement la censure aléatoire à droite : une variable aléatoire de durée  $T$  est censurée à droite par une variable de censure  $C$  si on n'arrive pas à observer  $T$  quand  $T > C$  (ou on observe la variable  $T$  que lorsque  $T \leq C$ ). Cela arrive particulièrement dans des données de survie où on n'a pas accès à toutes les observations. Par exemple dans le suivi de traitement de  $n$  patients, nous associons à chaque patient  $i$ , non seulement son temps de guérison  $T_i$  mais aussi la censure  $C_i$  (soit "perdu de vue", soit on arrête l'étude et le patient n'est toujours pas guéri ect...). Dans ce cas on observe  $T_i \wedge C_i = X_i$

En général les v.a.  $T_i$  sont i.i.d de f.d.r  $F$  les v.a  $C_i$  sont i.i.d. de f.d.r  $G$  et  $(T_i)_i$  est indépendante de  $(C_i)_i$ . Pour le  $i$ ème individu, on associe : la durée  $T_i$  et la v.a.  $D_i = \mathbf{1}_{\{T_i \leq C_i\}}$  qui vaut 1 si  $T_i$  est observée et 0 si  $T_i$  est censurée. Lorsqu'on ordonne les données  $(X_i, D_i)$  :

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}, \quad D_{(1)} \leq D_{(2)} \leq \dots \leq D_{(n)}$$

la v.a.  $D_{(i)}$  est l'indicateur de censure (à droite) associé à  $X_{(i)}$ .

### 2.1.3 Processus ponctuels associés

Au milieu des années 1970, Aalen présente une théorie des martingales pour le processus de comptage qui offre un cadre unifié pour les méthodes statistiques de l'analyse de survie. Les processus de comptage sont utilisés dans les représentations intégrales des statistiques des données censurées. La théorie fournit des formes simples et unifiées des estimateurs, des statistiques de test et des méthodes de régression. Ces méthodes permettent aussi d'obtenir des expressions simples de statistiques compliquées et des distributions asymptotiques des test et des estimateurs.

Dans l'analyse des durées de survie, chaque "individu" a une durée de vie  $T$ , de densité  $f$ , de fonction de répartition  $F$  et de fonction de survie  $S = 1 - F$  et les v.a.  $T$  et la censure  $C$  sont supposées indépendantes. A chaque "individu" on associe un couple de processus ponctuel  $((Y(t), N(t)), t \geq 0)$  donné par : (où  $X = T \wedge C$ )

$$\begin{aligned} Y(t) &= \mathbf{1}_{\{X \geq t\}} = \mathbf{1}_{\{T \wedge C \geq t\}} \\ &= \begin{cases} 1 & \text{si l'individu est en observation et à risque à l'instant } t \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

Le processus  $(Y(t), t \geq 0)$  est dit processus risque : il indique la présence du risque juste avant l'instant  $t$  (le sujet n'a pas encore subi l'évènement à  $t^-$ ). Le processus  $N(t)$  est défini par :

$$N(t) = \text{l'indicateur de l'"évènements d'intérêt" dans l'intervalle } ]0, t]$$

et nous avons avec l'indicatrice de censure  $D = \mathbf{1}_{\{T \leq C\}}$

$$N(t) = \mathbf{1}_{\{X \leq t, D=1\}} = \mathbf{1}_{\{T \leq t, D=1\}} = \mathbf{1}_{\{T \leq t\}} \mathbf{1}_{\{D=1\}} = \mathbf{1}_{\{T \leq t\}} \mathbf{1}_{\{T \leq C\}} = D \mathbf{1}_{\{T \leq t\}}$$

Rappelons que la fonction de risque

$$h(t) = \lim_{dt \rightarrow 0} \frac{\mathbf{P}(t < T \leq t + dt / T > t)}{dt}$$

et donc pour  $dt$  "petit" on peut écrire :

$$\begin{aligned} h(t)dt &\approx \mathbf{P}(t < T \leq t + dt / T > t) \\ &\approx \mathbf{P}(t < T \leq t + dt / T > t, C > t) \end{aligned}$$

Si on note  $dN(t)$  l'accroissement du processus  $N$  sur  $]t, t + dt]$ , avec  $dt$  "petit"  $dN(t)$  ne prend que 0 (pas d'évènement) ou 1 (survenue de l'évènement) et donc

$$\begin{aligned} \mathbf{P}(dN(t) = 1 / T > t, C > t) &= \mathbf{P}(t < T \leq t + dt / T > t, C > t) \\ &\approx h(t)dt \end{aligned}$$

$dN(t)$  est une v.a. de Bernoulli et sa moyenne est

$$\mathbf{E}(dN(t) / T > t, C > t) = h(t)dt$$

Si on note la tribu  $\mathcal{F}_t = \sigma(N(s), \mathbf{1}_{\{X \leq s, D=0\}}, s \leq t)$  et la tribu  $\mathcal{F}_{t-} = \sigma(N(s), \mathbf{1}_{\{X \leq s, D=0\}}, s < t)$  (évènements avant  $t$ ). Alors conditionnellement à  $\mathcal{F}_{t-}$  la v.a.  $dN(t)$  admet une loi de Bernoulli de paramètre  $Y(t)h(t)dt$

$$\mathbf{P}(dN(t) = 1 / \mathcal{F}_{t-}) = Y(t)h(t)dt$$

$$\mathbf{P}(dN(t) = 0 / \mathcal{F}_{t-}) = 1 - Y(t)h(t)dt$$

et

$$\mathbf{E}(dN(t) / \mathcal{F}_{t-}) = Y(t)h(t)dt$$

Le processus stochastique défini par

$$M(t) = N(t) - \int_0^t Y(u)h(u)du \quad (2.1)$$

est une martingale par rapport à  $\mathcal{F}_{t-}$  et  $\int_0^t h(u)Y(u)du$  est le processus croissant de  $N(t)$ . En effet pour la propriété de martingale on a :

$$\begin{aligned} \mathbf{E}(dM(t) / \mathcal{F}_{t-}) &= \mathbf{E}(dN(t) - Y(t)h(t)dt / \mathcal{F}_{t-}) \\ &= \mathbf{E}(dN(t) / \mathcal{F}_{t-}) - \mathbf{E}(Y(t)h(t)dt / \mathcal{F}_{t-}) \\ &= \mathbf{P}(dN(t) = 1 / \mathcal{F}_{t-}) - Y(t)h(t)dt \\ &= 0 \end{aligned}$$

Les processus

$$(\lambda(t) = Y(t)h(t), t \geq 0)$$

et

$$\Lambda(t) = \int_0^t \lambda(u)du = \int_0^t Y(u)h(u)du = \int_0^{t \wedge X} h(u)du = H(t \wedge X)$$

sont appelés respectivement processus intensité et processus intensité cumulée du processus  $(N(t), t \geq 0)$ . Le processus  $(\Lambda(t), t \geq 0)$  est appelée "compensateur prévisible" de  $N(t)$  (il est déterminé par  $\mathcal{F}_{t-}$ ). Enfin on a

$$\mathbf{E}(dN(t) / \mathcal{F}_{t-}) = \lambda(t)dt$$

$$\mathbf{V}(dN(t) / \mathcal{F}_{t-}) = (\lambda(t)dt)(1 - \lambda(t)dt) \quad (2.2)$$

**Notations :** Si dans l'étude nous avons  $n$  "individus" (machine, patients ...ect) indexés par  $i = 1, \dots, n$ , on note  $T_1, T_2, \dots, T_n$  les instants de survenue de "l'évènement d'intérêt". On observe  $X_i = T_i \wedge C_i$  et  $D_i = \mathbf{1}_{\{T_i \leq C_i\}}$ . Donc on dispose des observations  $(X_i, D_i)$ ,  $i = 1, \dots, n$ . A chaque "individu  $i$ " correspond le processus risque  $Y_i(t)$  défini par

$$Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}$$

et le processus ponctuel  $N_i(t)$  indicateur de "survenue de l'évènement" dans  $]0, t]$

$$N_i(t) = \mathbf{1}_{\{X_i \leq t, D_i=1\}}$$

On définit le processus risque  $\bar{Y}_n(t)$  par

$$\bar{Y}_n(t) = \sum_{i=1}^n Y_i(t)$$

qui est le nombre d'"individus" à risque à l'instant  $t$  et en observation. De même on définit le processus  $\bar{N}_n(t)$  par

$$\bar{N}_n(t) = \sum_{i=1}^n N_i(t)$$

qui est le nombre d'"événements" observés dans l'intervalle  $]0, t]$ . Le processus à sauts  $\bar{Y}_n(t)$  est continu à gauche, les autres processus (intensité...ect...) sont définis de manière similaire.

Le processus  $(N(t) , t \geq 0)$  est continu à droite avec limite à gauche et  $N(0) = 0$ .

Le processus  $(Y(t) , t \geq 0)$  est continu à gauche avec limite à droite.

## 2.2 Théorème de Rebolledo

Dans cette partie nous présentons les résultats de [3] p204-225

Pour montrer le théorème de Rebolledo on a besoin des résultats suivants.

**Théorème 2.2.1** Soit  $W_1, \dots, W_r$  des mouvements browniens indépendants et  $f_1, \dots, f_r$  des fonctions mesurables positives telles que :

$$\alpha_l(t) = \int_0^t f_l^2(s) ds < \infty \text{ pour tout } t > 0, l = 1, \dots, r$$

. Soit  $\{X_{l,n,j}\}$  une matrice de martingales par rapport à  $\{\mathcal{F}_{n,j}\}$ ,  $W_{l,n}(t) = \sum_{j=1}^{r_n(t)} X_{l,n,j}$  et  $W_{l,n} = \{W_{l,n}(t) : 0 \leq t \leq \tau\}$  pour tout  $l = 1, \dots, r$ .

Pour toute martingale  $\{X_{l,n,j}\}_{l=1, \dots, r, t \in [0, \tau]}$  vérifiant :

1.

$$\sum_{j=1}^{r_n(t)} E_{j-1} X_{l,n,j}^2 \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \alpha_l(t)$$

2. pour tout  $\varepsilon > 0$

$$\sum_{j=1}^{r_n(t)} E_{j-1} (X_{l,n,j}^2 \mathbf{1}_{\{|X_{l,n,j}| > \varepsilon\}}) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

3. si  $l \neq l'$

$$\sum_{j=1}^{r_n(t)} E_{j-1} (X_{l,n,j} X_{l',n,j}) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

avec  $E_j(Z) := E(Z/F_{n,j})$ , on a

$$(W_{1,n}, \dots, W_{r,n}) \Rightarrow_{n \rightarrow \infty} \left( \int f_1 dW_1, \dots, \int f_r dW_r \right)$$

dans  $(D[0, \tau])^r$

**Théorème 2.2.2 (Théorème de Doléans-Dade)** Soit  $(X(s), 0 \leq s \leq t)$  une surmartingale adaptée à la filtration  $\mathcal{F}_s$  et  $(A(s), 0 \leq s \leq t)$  un processus croissant, continu et intégrable tel que  $X - A$  est une martingale.

Alors pour toute subdivision  $S = (t_0, t_1, \dots, t_m)$  de  $[0, t]$

$$\begin{aligned} A^S(t) &= \sum_{j=0}^{m-1} E\{A(t_{j+1}) - A(t_j) / \mathcal{F}_{t_j}\} \\ &= \sum_{j=0}^{m-1} E\{X(t_{j+1}) - X(t_j) / \mathcal{F}_{t_j}\} \end{aligned}$$

on a

$$E|A^S(t) - A(t)| \xrightarrow{n \rightarrow \infty} 0$$

**Corollaire 2.2.3** Soit  $M$  une martingale de carré intégrable tel que  $\langle M, M \rangle$  est continue. Pour toute subdivision  $S = (t_0, t_1, \dots, t_m)$  de  $[0, t]$  on a

$$\sum_{j=0}^{m-1} E[M(t_{j+1}) - M(t_j)]^2 / F_{t_j} \xrightarrow{\mathcal{L}} \langle M, M \rangle_t$$

**Lemme 2.2.4 .**

1. Soit  $Y$  une variable aléatoire,  $(Y_n : n \geq 1)$  et  $(Y_{n,k} : n, k \geq 1)$  des suites de variables aléatoires. Supposons que :

a) pour tout  $n$ ,  $Y_{n,k} \xrightarrow[k \rightarrow \infty]{\mathbf{P}} Y_n$

b)  $Y_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}} Y$

Alors pour toute suite croissante  $k_n$  tel que  $k_n \rightarrow \infty$  on a

$$\lim_{n \rightarrow \infty} Y_{n, k_n} = Y \quad \text{en probabilité}$$

2. Soit  $Y(\cdot)$  un processus stochastique croissant et continue,  $\{Y_n(\cdot) : n \geq 1\}$  et  $\{Y_{n,k}(\cdot) : n, k \geq 1\}$  des suites de processus croissants. Supposons, pour tout  $t > 0$ , les conditions a) et b) sont vérifiées pour  $Y_{n,k}(t), Y_n(t)$  et  $Y(t)$ . Alors pour tout  $c > 0$ , et  $k_n \rightarrow \infty$ , on a

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq c} |Y_{n, k_n}(t) - Y(t)| = 0 \quad \text{en probabilité}$$

**Lemme 2.2.5** Soit  $\{A_{n,k} : k \geq 1, n \geq 1\}$  une famille d'évènements dans l'espace de probabilité  $(\Omega, \mathcal{F}, P)$  et  $\{\mathcal{F}_{n,k} : k \geq 1, n \geq 1\}$  une suite de filtrations en  $n$ . Soit  $A_{n,k} \in \mathcal{F}_{n,k}$  et  $K_n$  une suite d'entiers positifs aléatoires tel que ,pour tout  $n$ ,  $K_n$  est un temps d'arrêt par rapport à  $\{\mathcal{F}_{n,k} : k \geq 1\}$ .

On a

$$\lim_{n \rightarrow \infty} P \left\{ \bigcup_{k=1}^{K_n} A_{n,k} \right\} = 0$$

si et seulement si

$$\sum_{k=1}^{k_n} E_{k-1}(1_{\{A_{n,k}\}}) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

**Lemme 2.2.6** Soit  $x \in D[0, \infty)$  et  $\{Q(k) : k \geq 1\} = \{\{t_{j,k} : j \geq 0\} : k \geq 1\}$  une R-séquence de  $\mathbb{R}$  (cf. déf 5.2.1 de [3] p. 205). Pour tout  $t > 0$  on définit  $k(t) = \max\{j : t_{j,k} \leq t\}$  et

$$a_k = \max_{j \leq k(t)} \sup_{u,v \in [t_{j,k}, t_{j+1,k}]} |x(u) - x(v)|$$

alors

$$\lim_{k \rightarrow \infty} a_k = \max_{s \leq t} |\Delta x(s)| \text{ où } \Delta x(s) = x(s) - x(s^-) \text{ (saut en } s)$$

**Théorème 2.2.7 Théorème de Rebolledo.**

Soit  $(M_i^{(n)})$  une suite de martingales telle que  $M_i^{(n)} = N_i^{(n)} - A_i^{(n)}$  où  $N_i^{(n)}$  désigne un processus de comptage et  $A_i^{(n)}$  son compensateur. Considérons la suite des processus

$$U^{(n)}(t) = \sum_{i=1}^n \int_0^t f_i^{(n)}(s) dM_i^{(n)}(s) = \sum_{i=1}^n U_i^{(n)}(t)$$

et pour  $\varepsilon > 0$ ,

$$U_\varepsilon^{(n)}(t) = \int_0^t f^n(s) \mathbf{1}_{\{|f^n(s)| > \varepsilon\}} dM^{(n)}$$

où  $f_n$  et  $f$  sont des processus  $\mathcal{F}_{s^-}$  prévisibles et localement bornés tels que  $\sup_s |f_n(s) - f(s)| \rightarrow_{n \rightarrow \infty} 0$ . Supposons également que les processus  $N_i^{(n)}, A_i^{(n)}$  et  $f_i^{(n)}$  sont bornés par  $C^n$

Posons pour tout  $t > 0$ ,  $\alpha(t) = \int_0^t f^2(s) ds < \infty$ .

Si

$$\langle U^{(n)} \rangle_t \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \alpha(t)$$

et pour tout  $\varepsilon > 0$

$$\langle U_\varepsilon^{(n)} \rangle_t \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

Alors

$$(U^{(n)}(t), t > 0) \xrightarrow[n \rightarrow \infty]{} \left( \int_0^t f(s) dW(s), t > 0 \right)$$

dans  $D[0, \infty)$  muni de la topologie de Skohorod et  $W$  est un mouvement brownien.

**Preuve :**

Montrons d'abord la convergence faible dans  $D[0, \tau]$ ,  $\forall \tau > 0$  ce qui va impliquer la convergence faible dans  $D[0, \infty)$ . Pour une suite de partition  $\{Q(k) : k \geq 1\} = \{\{t_{j,k} : j \geq 0\} : k \geq 1\}$ , on définit  $W_k^{(n)}(t) = \sum_{j=0}^{k(t)} X_{j,k}^{(n)}$  avec  $X_{j,k}^{(n)} = U^{(n)}(t_{j+1,k}) - U^{(n)}(t_{j,k})$  et  $k(t) = \max\{j : t_{j,k} \leq t\}$ . Montrons que il existe une suite  $k_n$  croissante tel que pour  $W^{(n)}(t) = W_{(k_n)}^{(n)}(t)$  on ait

$$\sup_{t \leq \tau} |U^{(n)}(t) - W^{(n)}(t)| \xrightarrow{\mathbf{P}} 0$$

et

$$W^{(n)} \Rightarrow U^\infty \text{ dans } D[0, \tau]$$

et on conclut en appliquant le théorème 1.3.1.3 (chap.1).

Premierement nous allons montrer que

$$\sup_{t \leq \tau} |U^{(n)}(t) - W^{(n)}(t)| \xrightarrow{\mathbf{P}} 0$$



Par construction on a :

$$\sup_{t \leq \tau} |U^{(n)}(t) - W_k^{(n)}(t)| \leq \max_{j \leq k(\tau)} \sup_{u, v \in [t_{j,k}, t_{j+1,k} \wedge \tau]} |U^{(n)}(u) - U^{(n)}(v)|$$

Par le lemme 2.2.6 on a

$$\lim_{k \rightarrow \infty} \max_{j \leq k(\tau)} \sup_{u, v \in [t_{j,k}, t_{j+1,k} \wedge \tau]} |U^{(n)}(u) - U^{(n)}(v)| = \max_{t \leq \tau} |\Delta U^{(n)}(t)|$$

Il reste à montrer que

$$\max_{t \leq \tau} |\Delta U^{(n)}(t)| \xrightarrow{\mathbf{P}} 0 \text{ quand } n \rightarrow \infty$$

ce que l'on fera par le lemme 2.2.4. Remarquons d'abord que pour tout  $s$

$$\Delta U^{(n)}(s) \mathbf{1}_{\{|\Delta U^{(n)}(s)| > \varepsilon\}} = \sum_{i=1}^n \Delta U_i^{(n)}(s) \mathbf{1}_{\{|\Delta U_i^{(n)}(s)| > \varepsilon\}} \text{ presque surement}$$

Comme  $A_i^{(n)}$  est continue, les conditions du théorème 2.2.2 sont vérifiées pour

$$X(t) = \int_0^t H_i^{(n)}(s)^2 \mathbf{1}_{\{|H_i^{(n)}(s)| > \varepsilon\}} dN_i^{(n)}(s)$$

et

$$A(t) = \int_0^t H_i^{(n)}(s)^2 \mathbf{1}_{\{|H_i^{(n)}(s)| > \varepsilon\}} dA_i^{(n)}(s)$$

Donc, quand  $k \rightarrow \infty$ , on utilise le corollaire du théorème pour avoir

$$\begin{aligned} Q_k^{(n)} &:= \sum_{j=0}^{k(\tau)} E \left\{ \sum_{t_{j,k} < s \leq (t_{j+1,k} \wedge \tau)} \Delta U^{(n)}(s)^2 \mathbf{1}_{\{|\Delta U^{(n)}(s)| > \varepsilon\}} / F_{t_{j,k}}^{(n)} \right\} \\ &= \sum_{i=1}^n \sum_{j=0}^{k(\tau)} E \left\{ \sum_{t_{j,k} < s \leq (t_{j+1,k} \wedge \tau)} \Delta U_i^{(n)}(s)^2 \mathbf{1}_{\{|\Delta U^{(n)}(s)| > \varepsilon\}} / F_{t_{j,k}}^{(n)} \right\} \\ &\rightarrow \sum_{i=1}^n \langle U_{i,\varepsilon}^{(n)} \rangle_\tau = \langle U_\varepsilon^{(n)} \rangle_\tau \text{ dans } L^1 \end{aligned}$$

La deuxième hypothèse du théorème :  $\langle U_\varepsilon^{(n)} \rangle_\tau \xrightarrow{\mathbf{P}} 0$  quand  $n \rightarrow \infty$ . Le lemme 2.2.4 donne qu'il existe  $k_n$  tel que

$$Q_{k_n}^{(n)} \xrightarrow{\mathbf{P}} 0 \text{ quand } n \rightarrow \infty$$

D'autre part, on a

$$\sum_{t_{j,k} < s \leq (t_{j+1,k} \wedge \tau)} \Delta U^{(n)}(s)^2 \mathbf{1}_{\{|\Delta U^{(n)}(s)| > \varepsilon\}} \geq \varepsilon^2 \mathbf{1}_{\{\max_{t_{j,k} < s \leq (t_{j+1,k} \wedge \tau)} |\Delta U^{(n)}(s)| > \varepsilon\}}$$

De ce qui précède

$$\sum_{j=0}^{k_n(\tau)} E \left\{ \mathbf{1}_{\{\max_{t_{j,k} < s \leq (t_{j+1,k} \wedge \tau)} |\Delta U^{(n)}(s)| > \varepsilon\}} / F_{t_{j,k_n}}^{(n)} \right\} \xrightarrow{\mathbf{P}} 0 \text{ quand } n \rightarrow \infty$$

ce qui donne par le lemme 2.2.6

$$P\{\max_{s \leq \tau} |\Delta U^{(n)}(s)| > \varepsilon\} \rightarrow 0$$

D'où

$$\sup_{t \leq \tau} |U^{(n)}(t) - W^{(n)}(t)| \xrightarrow{\mathbf{P}} 0$$

Montrons le deuxième point :

$$W^{(n)} \Rightarrow U^\infty \text{ dans } D[0, \tau]$$

Pour cela, nous allons montrer qu'il existe une suite  $(k_n)$  pour laquelle les 3 conditions du théorème 2.2.1 sont vérifiées.

la première condition du théorème 2.2.1 :

$$\sum_{j=1}^{r_n(t)} E_{j-1} X_{n,j}^2 \xrightarrow{\mathbf{P}} \int_0^t f^2(s) ds = \alpha(t)$$

peut s'écrire

$$A_{k_n}^{(n)}(t) \xrightarrow{\mathbf{P}} \alpha(t) \text{ pour tout } t \in [0, \tau]$$

quand

$$A_k^{(n)}(t) = \sum_{j=0}^{k(t)} E \left[ \{U^{(n)}(t_{j+1,k}) - U^{(n)}(t_{j,k})\}^2 / F_{t_{j,k}}^{(n)} \right]$$

Comme  $\langle U^{(n)} \rangle$  est continue, le corollaire 2.2.3 implique

$$\lim_{k \rightarrow \infty} \sum_{j=0}^{k(t)} E \left[ \{U^{(n)}(t_{j+1,k} \wedge t) - U^{(n)}(t_{j,k})\}^2 / F_{t_{j,k}}^{(n)} \right] = \langle U^{(n)} \rangle_t \text{ en probabilité}$$

Aussi, par la continuité à droite de  $U^{(n)}$

$$\lim_{k \rightarrow \infty} U^{(n)}(t_{k(t)+1,k}) - U^{(n)}(t) = 0 \text{ presque sûrement}$$

(car  $t_{k(t)+1,k} \rightarrow t$ ) Donc

$$A_k^{(n)}(t) \xrightarrow{\mathbf{P}} \langle U^{(n)} \rangle_t \text{ quand } k \rightarrow \infty$$

D'où

$$\langle U^{(n)} \rangle_t \xrightarrow{\mathbf{P}} \alpha(t) \text{ quand } n \rightarrow \infty \text{ pour tout } t \in [0, \tau]$$

Comme  $A_k^{(n)}, \langle U^{(n)} \rangle$  et  $\alpha(\cdot)$  sont des processus croissants et  $\alpha(\cdot)$  est continue, on applique le lemme 2.2.4 partie (2) ce qui donne l'existence de la suite  $k_n$  tel que

$$A_{k_n}^{(n)}(t) \xrightarrow{\mathbf{P}} \alpha(t) \text{ pour tout } t \in [0, \tau]$$

Par suite la première condition du théorème 2.2.1 est vérifiée.

Pour la deuxième condition du théorème 2.2.1

$$\sum_{j=1}^{r_n(t)} E_{j-1} (X_{n,j})^2 \mathbb{1}_{\{|X_{l,n,j}| > \varepsilon\}} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0 \text{ pour tout } \varepsilon > 0$$

elle peut s'écrire, à  $t = \tau$ ,

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{k_n(\tau)} E \left\{ [U^{(n)}(t_{j+1, k_n}) - U^{(n)}(t_{j, k_n})]^2 \mathbf{1}_{\{|U^{(n)}(t_{j+1, k_n}) - U^{(n)}(t_{j, k_n})| > \varepsilon\}} / F_{t_j, k_n}^{(n)} \right\} = 0 \quad (2.3)$$

en probabilité pour tout  $\varepsilon > 0$

Fixons  $\delta > 0$  et notons  $U^{(n)} = \overline{U}_\delta^{(n)}(t) + \underline{U}_\delta^{(n)}(t)$  avec

$$\overline{U}_\delta^{(n)}(t) = \sum_{i=1}^n \overline{U}_{i, \delta}^{(n)}(t) = \sum_{i=1}^n \int_0^t H_i^{(n)}(s) \mathbf{1}_{\{|H_i^{(n)}(s)| > \delta\}} dM_i^{(n)}(s)$$

et

$$\underline{U}_\delta^{(n)}(t) = \sum_{i=1}^n \underline{U}_{i, \delta}^{(n)}(t) = \sum_{i=1}^n \int_0^t H_i^{(n)}(s) \mathbf{1}_{\{|H_i^{(n)}(s)| \leq \delta\}} dM_i^{(n)}(s)$$

Or pour  $a_i = b_i + c_i$  avec  $i = 1, 2$  on a

$$(a_1 - a_2)^2 \mathbf{1}_{\{|a_1 - a_2| > \varepsilon\}} \leq 4(b_1 - b_2)^2 \mathbf{1}_{\{|b_1 - b_2| > \frac{\varepsilon}{2}\}} + (c_1 - c_2)^2 \mathbf{1}_{\{|c_1 - c_2| > \frac{\varepsilon}{2}\}}$$

D'où

$$\sum_{j=0}^{k_n(\tau)} E \left[ \{U^{(n)}(t_{j+1, k_n}) - U^{(n)}(t_{j, k_n})\}^2 \mathbf{1}_{\{|U^{(n)}(t_{j+1, k_n}) - U^{(n)}(t_{j, k_n})| > \varepsilon\}} / F_{t_j, k_n}^{(n)} \right] \leq 4Q_{1, k_n}^{(n)} + 4Q_{2, k_n}^{(n)}$$

avec

$$Q_{1, k_n}^{(n)} = \sum_{j=0}^{k_n(\tau)} E \left[ \{\underline{U}_\delta^{(n)}(t_{j+1, k_n}) - \underline{U}_\delta^{(n)}(t_{j, k_n})\}^2 \mathbf{1}_{\{|\underline{U}_\delta^{(n)}(t_{j+1, k_n}) - \underline{U}_\delta^{(n)}(t_{j, k_n})| > \frac{\varepsilon}{2}\}} / F_{t_j, k_n}^{(n)} \right]$$

et

$$Q_{2, k_n}^{(n)} = \sum_{j=0}^{k_n(\tau)} E \left[ \{\overline{U}_\delta^{(n)}(t_{j+1, k_n}) - \overline{U}_\delta^{(n)}(t_{j, k_n})\}^2 \mathbf{1}_{\{|\overline{U}_\delta^{(n)}(t_{j+1, k_n}) - \overline{U}_\delta^{(n)}(t_{j, k_n})| > \frac{\varepsilon}{2}\}} / F_{t_j, k_n}^{(n)} \right]$$

Pour montrer (2.3) il suffit de montrer que  $Q_{1, k_n}^{(n)} \xrightarrow{\mathbf{P}} 0$  et  $Q_{2, k_n}^{(n)} \xrightarrow{\mathbf{P}} 0$  pour une suite  $k_n$  telle que  $k_n \rightarrow \infty$ .

Par hypothèse les processus  $A_i^{(n)}, N_i^{(n)}, f_i^{(n)}$  sont bornés par  $C^n$ . Donc pour tout  $n$  il existe une constante  $b_n$  tel que

$$\sup_s |\underline{U}_\delta^{(n)}(s)| \leq b_n$$

avec  $b_n \rightarrow \infty$  quand  $n \rightarrow \infty$

D'où

$$Q_{1, k_n}^{(n)} \leq (2b_n)^2 \sum_{j=0}^{k_n(\tau)} E \left[ \mathbf{1}_{\{|\underline{U}_\delta^{(n)}(t_{j+1, k_n}) - \underline{U}_\delta^{(n)}(t_{j, k_n})| > \frac{\varepsilon}{2}\}} / F_{t_j, k_n}^{(n)} \right]$$

$\underline{U}_\delta^{(n)}$  saute au plus  $\delta$  puisque les deux composantes du processus ne saute pas au même temps et que leurs sauts ne dépasse pas  $\delta$ . Ainsi en posant  $\delta < \frac{\varepsilon}{2}$ , on a pour un  $n$  fixé (sauts ne dépassant pas  $\delta$ )

$$\lim_{n \rightarrow \infty} P \left\{ \max_{j \leq k(\tau)} |\underline{U}_\delta^{(n)}(t_{j+1, k}) - \underline{U}_\delta^{(n)}(t_{j, k})| > \frac{\varepsilon}{2} \right\} = 0$$

Par suite du lemme 2.3.4 on en déduit :

$$(2b_n)^2 \sum_{j=0}^{k_n(\tau)} E \left[ \mathbb{1}_{\{|\underline{U}_\delta^{(n)}(t_{j+1,k_n}) - \underline{U}_\delta^{(n)}(t_{j,k_n})| > \frac{\varepsilon}{2}\}} / F_{t_{j,k_n}}^{(n)} \right] \xrightarrow[k \rightarrow \infty]{\mathbf{P}} 0$$

On conclut de nouveau par le lemme 2.2.4, qu'il existe une suite  $(k_n)$  tel que

$$Q_{1,k_n}^{(n)} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

Pour le second terme  $Q_{2,k_n}^{(n)}$ , on a

$$Q_{2,k_n}^{(n)} \leq \sum_{j=0}^{k_n(\tau)} E \left[ \{\bar{U}_\delta^{(n)}(\tau \wedge t_{j+1,k}) - \bar{U}_\delta^{(n)}(t_{j,k})\}^2 / F_{t_{j,k}}^{(n)} \right] + E \left[ \{\bar{U}_\delta^{(n)}(t_{k(\tau)+1,k}) - \bar{U}_\delta^{(n)}(\tau)\}^2 / F_{t_{j,k}}^{(n)} \right]$$

Le corollaire 2.2.3 donne

$$\sum_{j=0}^{k_n(\tau)} E \left[ \{\bar{U}_\delta^{(n)}(\tau \wedge t_{j+1,k}) - \bar{U}_\delta^{(n)}(t_{j,k})\}^2 / F_{t_{j,k}}^{(n)} \right] \xrightarrow[k \rightarrow \infty]{L^1} \langle \bar{U}_\delta^{(n)} \rangle_\tau$$

On a aussi

$$|\bar{U}_\delta^{(n)}(t_{k(\tau)+1,k}) - \bar{U}_\delta^{(n)}(\tau)|^2 \xrightarrow[k \rightarrow \infty]{} 0 \text{ dans } L^1$$

car les processus sont continus à droite et bornés par une constante dépendant que de  $n$   
D'où

$$E \left[ \{\bar{U}_\delta^{(n)}(t_{k(\tau)+1,k}) - \bar{U}_\delta^{(n)}(\tau)\}^2 / F_{t_{j,k}}^{(n)} \right] \rightarrow 0 \text{ dans } L^1$$

Comme la convergence dans  $L^1$  implique la convergence en probabilité,

$$\sum_{j=0}^{k_n(\tau)} E \left[ \{\bar{U}_\delta^{(n)}(\tau \wedge t_{j+1,k}) - \bar{U}_\delta^{(n)}(t_{j,k})\}^2 / F_{t_{j,k}}^{(n)} \right] + E \left[ \{\bar{U}_\delta^{(n)}(t_{k(\tau)+1,k}) - \bar{U}_\delta^{(n)}(\tau)\}^2 / F_{t_{j,k}}^{(n)} \right] \xrightarrow[k \rightarrow \infty]{\mathbf{P}} \langle \bar{U}_\delta^{(n)} \rangle_\tau$$

Par la deuxième condition du théorème on a

$$\langle \bar{U}_\delta^{(n)} \rangle_\tau \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

Ainsi, le lemme 2.2.4 donne

$$Q_{2,k_n}^{(n)} \xrightarrow{\mathbf{P}} 0$$

Ce qui termine la démonstration du théorème. •

## 2.3 Estimation du taux de risque cumulé

### 2.3.1 Introduction

Pour l'individu  $i$ , on associe le couple  $(N_i(t), Y_i(t), t \geq 0)$  et  $(\bar{Y}_n(t), \bar{N}_n(t), t \geq 0)$  définis précédemment. On définit le processus  $J_n(t) = \mathbf{1}_{\{\bar{Y}_n(t) > 0\}}$  qui est l'indicateur de présence d'au moins un "individu" à risque. Dans cette partie on prend un taux de hasard  $h_i(u) = h(u), \forall i = 1, \dots, n$ . Le risque cumulé avec existence d'au moins un "individu" à risque sur  $[0, t]$  est :

$$\tilde{H}_n(t) = \int_0^t J_n(u) h(u) du$$

On pose

$$\hat{H}_n(t) = \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{N}_n(u)$$

qui sera un estimateur de  $\tilde{H}_n(t)$ . Calculons leur différence :

$$\begin{aligned} \hat{H}_n(t) - \tilde{H}_n(t) &= \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} (d\bar{N}_n(u) - \bar{Y}_n(u)h(u)du) \\ &= \sum_{i=1}^n \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} dM_i(u) \end{aligned}$$

où la martingale

$$M_i(u) = N_i(u) - \int_0^u Y_i(s)h(s)ds$$

est d'espérance nulle et elle représente la différence entre le nombre d'évènements observés sur  $[0, u]$  et son processus croissant (qui est le nombre d'évènements "prédit" par le modèle pour l'individu  $i$ , il est noté :  $\Lambda_i(t)$  et est appelé "compensateur prévisible" de  $N_i(t)$  puisqu'il est déterminé par  $\mathcal{F}_{t^-}$ .

On pose  $\bar{M}_n(u) = \sum_{i=1}^n M_i(u)$

### 2.3.2 Estimateur de Nelson-Aalen

#### 1) Construction

L'estimateur de Nelson-Aalen  $\hat{H}_{NA}(t)$  du taux de risque cumulé  $H(t)$  a été proposé par Nelson en 1972 et ensuite par Aalen en 1978. Cet estimateur est fondé sur la remarque suivante pour  $\Delta$  petit :

$$H(s + \Delta) - H(s) \approx h(s)\Delta \approx \mathbf{P}(s < T \leq s + \Delta \mid T > s) = \frac{\mathbf{P}(s < T \leq s + \Delta)}{\mathbf{P}(T > s)}$$

C'est la probabilité que l'évènement survienne dans un petit intervalle  $]s, s + \Delta]$  sachant qu'il y a un risque à l'instant  $s$ . Cette probabilité est approchée par la fréquence suivante :

$$\frac{[\bar{N}_n(s + \Delta) - \bar{N}_n(s)]}{\bar{Y}_n(s)} = \frac{d\bar{N}_n(s)}{\bar{Y}_n(s)}$$

où  $d\bar{N}_n(s) = \bar{N}_n(s + \Delta) - \bar{N}_n(s)$  est le nombre "d'évènements" observés dans l'intervalle  $]s, s + \Delta]$  et  $\bar{Y}_n(s)$  est le nombre d'"individus" à risque à l'instant  $s$ .

En sommant les quantités sur tous les sous intervalles formant  $]0, t]$  et en faisant tendre  $\Delta$  vers 0, de telle sorte que chaque intervalle contienne au plus un "évènement". Ainsi l'estimateur de Nelson-Aalen est défini par

$$\hat{H}_{NA}(t) = \int_0^t \frac{d\bar{N}_n(s)}{\bar{Y}_n(s)}$$

En effet : cet estimateur peut aussi s'écrire si on note  $t_1, t_2, \dots, t_n$  les instants ordonnés de survenue de l'évènement" pour les  $n$  "individus" (processus à sauts) :

$$\begin{aligned} \int_0^t \frac{d\bar{N}_n(s)}{\bar{Y}_n(s)} &= \sum_{\{i, t_i \leq t\}} \int_{t_{i-1}}^{t_i} \frac{d\bar{N}_n(s)}{\bar{Y}_n(s)} = \sum_{\{i, t_i \leq t\}} \frac{\bar{N}_n(t_i) - \bar{N}_n(t_{i-1})}{\bar{Y}_n(t_i)} \\ &= \sum_{\{i, t_i \leq t\}} \frac{\Delta \bar{N}_n(t_i)}{\bar{Y}_n(t_i)} = \hat{H}_{NA}(t) \end{aligned}$$

où sur  $]t_{i-1}, t_i]$  le processus  $\bar{Y}_n(s)$  est constant égal à  $\bar{Y}_n(t_i)$  et

$$\Delta \bar{N}_n(t_i) = \bar{N}_n(t_i) - \bar{N}_n(t_{i-1})$$

est le nombre "d'évènements" observés dans l'intervalle  $]t_{i-1}, t_i]$  i.e. en  $t_i$ . On peut directement construire un estimateur de la fonction de risque  $h(t_i)$  à l'instant  $t_i$  en utilisant la remarque précédente. En effet on pose

$$\hat{h}(t_i) = \frac{m(t_i)}{r(t_i)}$$

où  $m(t_i)$  désigne le nombre "d'évènements" observés jusqu'à l'instant  $t_i$  et  $r(t_i)$  est le nombre "d'individus" à risque à l'instant  $t_i$ .

Ainsi on estime le taux de risque cumulé  $H(t)$  en sommant les risques aux instants  $t_i \leq t$  par

$$\hat{H}_{NA}(t) = \sum_{\{i, t_i \leq t\}} \hat{h}(t_i) = \sum_{\{i, t_i \leq t\}} \frac{m(t_i)}{r(t_i)}$$

## 2) Exemple :

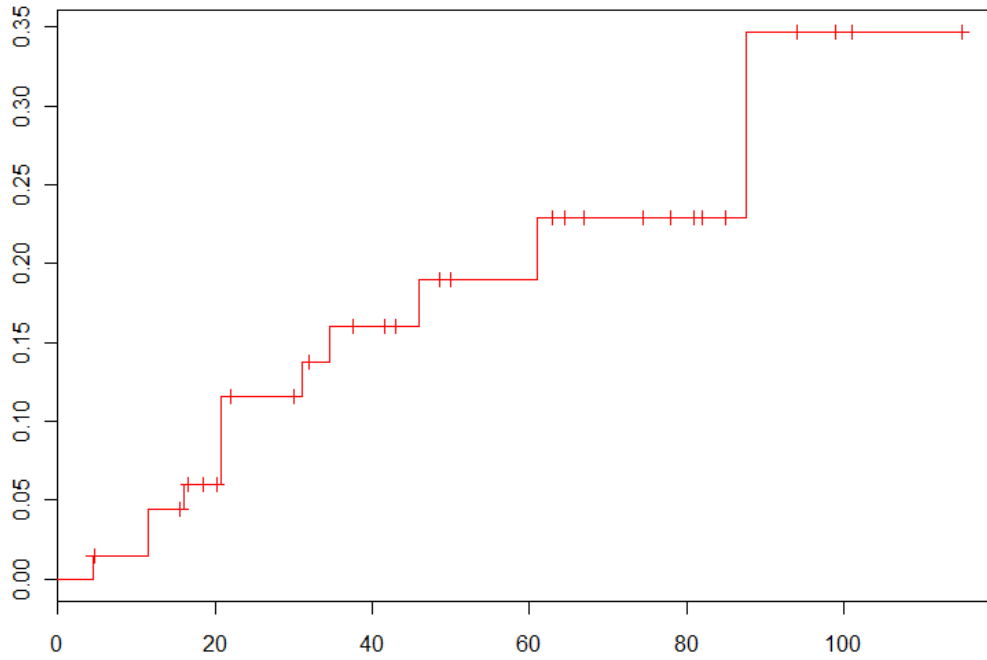
### Données des ventilateurs (Données de Aalen)

On étudie la durée de vie de 70 ventilateurs, en milliers d'heures de fonctionnement. La question qui se pose est de savoir si la fonction de risque  $h$  est décroissante dans le temps (qui signifie que le risque de panne est décroissant avec le temps et donc une bonne durée de fonctionnement des ventilateurs). On a le tableau des données suivant :

4.5	4.6 <sup>+</sup>	11.5	11.5	15.6 <sup>+</sup>	16.0	16.6 <sup>+</sup>	18.5 <sup>+</sup>	18.5 <sup>+</sup>	18.5 <sup>+</sup>
18.5 <sup>+</sup>	18.5 <sup>+</sup>	20.3 <sup>+</sup>	20.3 <sup>+</sup>	20.3 <sup>+</sup>	20.7	20.7	20.8	22.0 <sup>+</sup>	30.0 <sup>+</sup>
30.0 <sup>+</sup>	30.0 <sup>+</sup>	30.0 <sup>+</sup>	31.0	32.0 <sup>+</sup>	34.5	37.5 <sup>+</sup>	37.5 <sup>+</sup>	41.5 <sup>+</sup>	41.5 <sup>+</sup>
41.5 <sup>+</sup>	41.5 <sup>+</sup>	43.0 <sup>+</sup>	43.0 <sup>+</sup>	43.0 <sup>+</sup>	43.0 <sup>+</sup>	46.0	48.5 <sup>+</sup>	48.5 <sup>+</sup>	48.5 <sup>+</sup>
48.5 <sup>+</sup>	50.0 <sup>+</sup>	50.0 <sup>+</sup>	50.0 <sup>+</sup>	61.0 <sup>+</sup>	61.0	61.0 <sup>+</sup>	61.0 <sup>+</sup>	63.0 <sup>+</sup>	64.5 <sup>+</sup>
64.5 <sup>+</sup>	67.0 <sup>+</sup>	74.5 <sup>+</sup>	78.0 <sup>+</sup>	78.0 <sup>+</sup>	81.0 <sup>+</sup>	81.0 <sup>+</sup>	82.0 <sup>+</sup>	85.0 <sup>+</sup>	85.0 <sup>+</sup>
85.0 <sup>+</sup>	87.5 <sup>+</sup>	87.5	87.5 <sup>+</sup>	94.0 <sup>+</sup>	99.0 <sup>+</sup>	101.0 <sup>+</sup>	101.0 <sup>+</sup>	101.0 <sup>+</sup>	115.0 <sup>+</sup>

(+ indique une donnée censurée). Si on note  $t_1, t_2, \dots, t_n$  avec  $n = 70$ , les instants de panne observés.

1. *Taux de risque cumulé* : On a le graphe suivant :



Par exemple pour tout  $t \geq t_n = 16000$  heures :

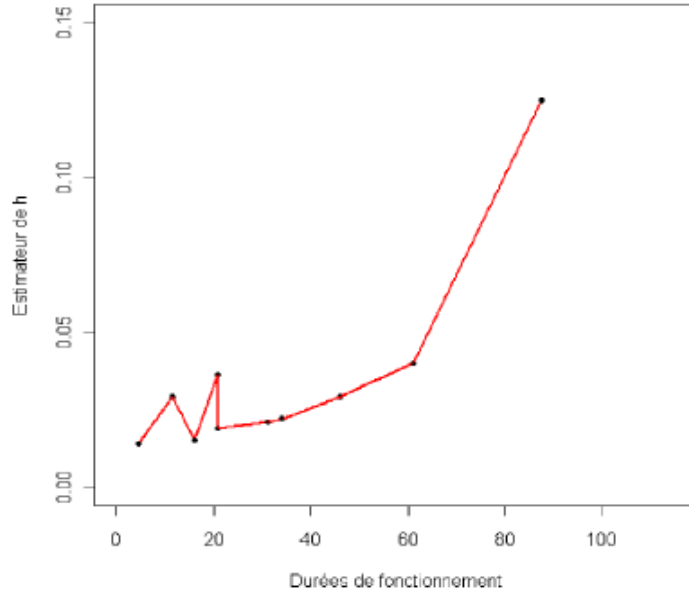
$$\begin{aligned}\hat{H}_{NA}(t) &= \frac{1}{70} + \frac{2}{68} + \frac{1}{65} \\ &= 0.0591\end{aligned}$$

Interpretation : on a 5.91% de risque de tomber en panne pour ce lot de ventilateurs durant la période de fonctionnement (16000 heures).

2. *Fonction de risque h* : Des estimations  $\hat{h}_n(t_i) = \frac{\Delta \bar{N}(t_i)}{Y(t_i)} = \frac{m_i}{r_i}$  de  $h(t_i)$ , on peut construire un graphe de  $h$  par interpolation linéaire des points  $(t_i, \hat{h}_n(t_i))$ . On a le tableau suivant :

Durées $t_i$	$m_i$	$r_i$	$\hat{h}_n(t_i)$
4.5	1	70	$1/70 = 0.014$
11.5	2	68	$2/68 = 0.029$
16.0	1	65	$1/65 = 0.015$
20.7	2	55	$2/55 = 0.036$
20.8	1	53	$1/53 = 0.019$
31.0	1	47	$1/47 = 0.021$
34.5	1	45	$1/45 = 0.022$
46.0	1	34	$1/34 = 0.029$
61.0	1	25	$1/25 = 0.04$
87.5	1	8	$1/8 = 0.125$

On obtient donc le graphe suivant de  $h$  :



Le graphe donne une fonction de risque  $h$  croissante dans le temps.

### 3) Consistance et loi asymptotique

Le théorème suivant donne la convergence en probabilité de  $\hat{H}_{NA}$ .

**Théorème 2.3.1** On a pour  $t \geq 0$

$$|\hat{H}_{NA}(t) - H(t)| \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

Le résultat suivant donne la convergence uniforme en probabilité de l'estimateur de  $\hat{H}_{NA}$ .

**Théorème 2.3.2** Si pour  $t > 0$ ,  $F(t) < 1$  et  $\bar{Y}_n(t) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \infty$  alors

$$\sup_{0 \leq s \leq t} |\hat{H}_{NA}(s) - H(s)| \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

Le théorème suivant donne la loi limite de  $\hat{H}_{NA}(t)$ .

**Théorème 2.3.3** .

1. L'estimateur  $\hat{H}_{NA}(t)$  est un estimateur avec biais de  $H(t)$ .
2. Sous l'hypothèse  $F(t) < 1$  nous avons

$$\sqrt{n} \left( \hat{H}_{NA}(t) - H(t) \right) \Longrightarrow_{n \rightarrow \infty} U(t)$$

où  $U(t)$  est une martingale gaussienne centrée de variance

$$\mathbf{V}(U(t)) = \int_0^t \frac{h(s)}{[1 - F(s)][1 - G(s^-)]} ds$$

et  $G$  la f.d.r de  $C$ .



**Preuve :**

1. Le biais de  $\hat{H}_{NA}(t)$  est donné grace à (2.1) et à la martingale

$$\bar{M}_n(u) = \sum_{i=1}^n M_i(u)$$

On a

$$\begin{aligned} \mathbf{E}(\hat{H}_{NA}(t)) &= \mathbf{E}\left(\int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{N}_n(u)\right) \\ &= \mathbf{E}\left(\int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} [d\bar{M}_n(u) + \bar{Y}_n(u)h(u)du]\right) \\ &= \mathbf{E}\left(\int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{M}_n(u)\right) + \mathbf{E}\left(\int_0^t J_n(u)h(u)du\right) \\ &:= E(M^*(t)) + \mathbf{E}\left(\int_0^t J_n(u)h(u)du\right) \\ &= E(M^*(0)) + \mathbf{E}\left(\int_0^t J_n(u)h(u)du\right) \quad (M^*(t) \text{ est une martingale}) \\ &= \int_0^t \mathbf{E}[J_n(u)]h(u)du \\ &= \int_0^t \mathbf{P}(\bar{Y}_n(u) > 0) h(u)du \\ &= \int_0^t h(u)du - \int_0^t \mathbf{P}(\bar{Y}_n(u) = 0) h(u)du \\ &= H(t) - \int_0^t \mathbf{P}(\bar{Y}_n(u) = 0) h(u)du \end{aligned}$$

Soit  $F$  la f.d.r des durées de survie  $T_i$  et  $G$  la f.d.r des  $C_i$ . On a vu au th 2.3.1 que

$$\int_0^t \mathbf{P}(\bar{Y}_n(u) = 0) h(u)du \xrightarrow{n \rightarrow \infty} 0$$

D'où le résultat.

2. La fonction de survie  $S_X$  des  $X_i$  est :

$$S_X(t) = (1 - F(t))(1 - G(t^-))$$

Par le théorème de Glivenko-Cantelli pour  $\bar{Y}_n(t) = \sum_{i=1}^n \mathbf{1}_{\{X_i \geq t\}}$  on a

$$\sup_{s \in [0, t]} \left| \frac{\bar{Y}_n(s)}{n} - (1 - F(s))(1 - G(s^-)) \right| \xrightarrow{n \rightarrow \infty} 0$$

Par ailleurs  $J_n(t) = \mathbf{1}_{\{\bar{Y}_n(t) > 0\}}$ , on en déduit que

$$1 - J_n(t) = \mathbf{1}_{\{\bar{Y}_n(t) = 0\}} \xrightarrow{n \rightarrow \infty} 0$$

En effet

$$\mathbf{P}(\bar{Y}_n(u) = 0) = \mathbf{P}\left(\sum_{i=1}^n Y_i(t) = 0\right)$$

qui est la probabilité d'avoir 0 individus à risque sur  $(0, t]$  parmi les  $n$  individus.  
Or pour un individu

$$\begin{aligned}\mathbf{P}(Y_i(t) = 1) &= \mathbf{P}(T_i > t, C_i \geq t) \\ &= \mathbf{P}(T_i > t)P(C_i \geq t) \\ &= (1 - F(t))(1 - G(t^-)) \text{ car } F \text{ est continue}\end{aligned}$$

Comme

$$\bar{Y}_n(t) \hookrightarrow B(n, (1 - F(t))(1 - G(t^-)))$$

$$P(\bar{Y}_n(t) = 0) = C_n^0 [1 - (1 - F(t))(1 - G(t^-))]^n \xrightarrow{n \rightarrow \infty} 0$$

Par suite

$$J_n(t) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 1$$

Or par (2.1) on a

$$dM(t) = dN(t) - Y(t)h(t)dt$$

et de

$$d \langle M \rangle_t = \text{var}(dM(t)/\mathcal{F}_{t^-})$$

on obtient par (2.2)

$$d \langle M \rangle_t = \text{var}(dN(t) - Y(t)h(t)dt/\mathcal{F}_{t^-}) = \text{var}(dN(t)/\mathcal{F}_{t^-}) = Y(t)h(t)dt$$

Par conséquent pour

$$L_t = \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} dM_i(u)$$

son processus croissant est

$$\langle L \rangle_t = \int_0^t \frac{J_n^2(u)}{\bar{Y}_n^2(u)} d \langle M \rangle_u$$

D'autre part

$$\begin{aligned}\langle \sqrt{n} \sum_{i=1}^n \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} dM_i(u) \rangle_t &= \sum_{i=1}^n n \int_0^t \frac{J_n^2(u)}{\bar{Y}_n^2(u)} Y_i(u) h(u) du \\ &= \int_0^t n \frac{J_n^2(u)}{\bar{Y}_n^2(u)} \sum_{i=1}^n Y_i(u) h(u) du \\ &= \int_0^t n \frac{J_n^2(u)}{\bar{Y}_n^2(u)} \bar{Y}_n(u) h(u) du \\ &= \int_0^t n \frac{J_n(u)}{\bar{Y}_n(u)} h(u) du\end{aligned}$$

D'autre part, le processus

$$\sqrt{n} (\hat{H}_{NA}(t) - H(t)) = \sqrt{n} \sum_{i=1}^n \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} dM_i(u)$$

est une martingale que l'on notera  $Z_n(t)$ . On applique le théorème central limite pour les martingales (Théorème de Rebolledo)

Vérifions les conditions du théorème pour la martingale  $Z_n(t) = \sqrt{n} \left( \hat{H}_{NA}(t) - H(t) \right)$ .

1. On a

$$\langle Z_n \rangle_t = \int_0^t n \frac{J(u)}{\bar{Y}(u)} h(u) du$$

et

$$J_n(u) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 1$$

et aussi, en appliquant le théorème de Glivenko-Cantelli pour  $\bar{Y}_n(t) = \sum_{i=1}^n \mathbf{1}_{\{X_i \geq t\}}$ , on a

$$\frac{\bar{Y}_n(u)}{n} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} (1 - F(u))(1 - G(u^-))$$

Donc

$$\langle Z_n \rangle_t = \int_0^t \frac{J_n(u)}{\frac{\bar{Y}_n(u)}{n}} h(u) du \xrightarrow[n \rightarrow \infty]{} \int_0^t \frac{h(u) du}{(1 - F(u))(1 - G(u^-))} = \alpha(t)$$

qui est déterministe. D'où la 1<sup>ère</sup> condition.

2. Pour la 2<sup>ème</sup> condition posons

$$Z_{n,\varepsilon}(t) = \int_0^t \sqrt{n} \frac{J(u)}{\bar{Y}(u)} \mathbf{1}_{\left\{ \left| \sqrt{n} \frac{J(u)}{\bar{Y}(u)} \right| > \varepsilon \right\}} d\bar{M}_n(u)$$

Montrons que  $\langle Z_{n,\varepsilon} \rangle_t \xrightarrow[n \rightarrow \infty]{} 0$  en probabilité :

$$\begin{aligned} \langle Z_{n,\varepsilon} \rangle_t &= \int_0^t n \frac{J(u)}{\bar{Y}(u)^2} \mathbf{1}_{\left\{ \left| \sqrt{n} \frac{J(u)}{\bar{Y}(u)} \right| > \varepsilon \right\}} d \langle \bar{M}_n \rangle_u \\ &= \int_0^t n \frac{J_n(u)}{\bar{Y}_n(u)^2} \mathbf{1}_{\left\{ \left| \sqrt{n} \frac{J(u)}{\bar{Y}(u)} \right| > \varepsilon \right\}} \bar{Y}(u) h(u) du \\ &= \int_0^t n \frac{J_n(u)}{\bar{Y}_n(u)} \mathbf{1}_{\left\{ \left| \sqrt{n} \frac{J(u)}{\bar{Y}(u)} \right| > \varepsilon \right\}} h(u) du \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0 \end{aligned}$$

car

$$n \frac{J_n(u)}{\bar{Y}_n(u)} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \frac{1}{(1 - F(u))(1 - G(u^-))}$$

et donc

$$\sqrt{n} \frac{J_n(u)}{\bar{Y}_n(u)} = \frac{1}{\sqrt{n}} n \frac{J_n(u)}{\bar{Y}_n(u)} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

D'où la 2<sup>ème</sup> condition.

Par conséquent :

$$(Z_n(t), t > 0) \xrightarrow[n \rightarrow \infty]{} \left( \int_0^t f(s) dW(s), t > 0 \right)$$

avec  $\alpha(t) = \int_0^t f^2(s) ds$ .

D'où pour tout  $t > 0$

$$Z_n(t) \xrightarrow[n \rightarrow \infty]{} Z(t) = \int_0^t f(s) dW(s) \hookrightarrow N(0, \alpha(t))$$

D'où le résultat du théorème. •

## 2.4 Estimation de la fonction de survie

### 2.4.1 Introduction

La fonction de survie  $S$  est *inconnue* et elle caractérise la loi de  $T$ . On va construire un estimateur en utilisant les observations

$$(X_i, D_i) = (T_i \wedge C_i, \mathbf{1}_{\{T_i \leq C_i\}}) \quad i = 1, \dots, n$$

Si les données ne sont pas censurées,  $S(t)$  peut être estimée par la f.d.r. empirique

$$S_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{T_i > t\}} = 1 - F_n(t)$$

où  $F_n(t)$  est la f.d.r. empirique des  $T_i$ . Si on dispose d'observations  $(T_i)$  avec présence de censures, estimer  $S$  seulement par les données  $(T_i)_{i=1, \dots, n}$  non censurées ( $D_i = 1$ ) fournit une estimation avec biais. En effet si on définit un estimateur  $\tilde{S}_n$

$$\tilde{S}_n(t) = \frac{\sum_{i=1}^n \mathbf{1}_{\{T_i > t, D_i = 1\}}}{\sum_{i=1}^n \mathbf{1}_{\{D_i = 1\}}}$$

alors par la loi forte des grands nombres on a

$$\tilde{S}_n(t) = \frac{\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{T_i > t, D_i = 1\}}}{\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{D_i = 1\}}} \xrightarrow{n \rightarrow \infty} \frac{\mathbf{P}(T_i > t, D_i = 1)}{\mathbf{P}(D_i = 1)} \neq S(t)$$

car

$$\begin{aligned} \frac{\mathbf{P}(T_i > t, D_i = 1)}{\mathbf{P}(D_i = 1)} &= \frac{\int_t^\infty (1 - G(u)) f(u) du}{\mathbf{P}(T_i \leq C_i)} = \frac{\int_t^\infty (1 - G(u)) f(u) du}{\int_0^\infty (1 - G(u)) f(u) du} \\ &= \frac{\int_t^\infty f(u) du - \int_t^\infty G(u) f(u) du}{1 - \int_0^\infty G(u) f(u) du} \neq \int_t^\infty f(u) du = S(t) \end{aligned}$$

Ainsi dans le cas de données censurées, l'estimation de la fonction de survie  $S$  nécessite des outils spécifiques. Il existe plusieurs méthodes non paramétriques parmi elles la plus connue est celle de *Kaplan-Meier*. D'abord on définit l'estimateur de *Fleming-Harington* qui se déduit immédiatement de l'estimation du taux cumulé.

### 2.4.2 Estimateur de Fleming-Harington

#### 1) Construction

L'estimateur de Fleming-Harington est l'estimateur qui découle de l'estimateur de Nelson-Aalen du risque cumulé  $H$  en utilisant la relation  $S = \exp(-H)$

$$\widehat{S}_{FH}(t) = \exp\left(-\sum_{i, t_i \leq t} \frac{m(t_i)}{r(t_i)}\right)$$

Il peut aussi s'écrire :

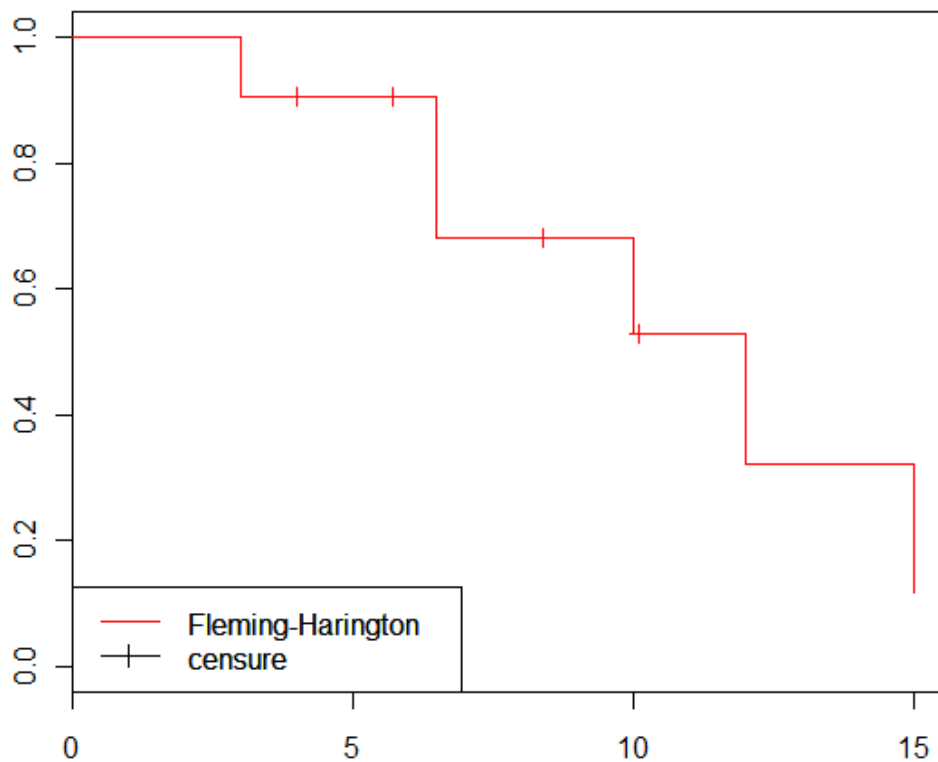
$$\widehat{S}_{FH}(t) = \exp\left(-\sum_{i, t_i \leq t} \frac{m(t_i)}{r(t_i)}\right)$$

## 2) Exemple

Sur 10 patients atteints de tumeurs, on a observé les durées de survie suivantes, exprimées en mois : 3, 4+, 5.7+, 6.5, 6.5, 8.4+, 10, 10.1+, 12, 15

L'estimateur de Fleming-Harington de la fonction de survie  $S(t)$  se calcule de la manière suivante :

temps	$R_i$	$m_i$	Survie	Intervalle
0	10	0	1	$[0, 3[$
3	10	1	0.905	$[3, 6.5[$
6.5	7	2	0.680	$[6.5, 10[$
10	4	1	0.530	$[10, 12[$
12	2	1	0.321	$[12, 15[$
15	1	1	0.118	



## 3) Consistence et loi asymptotique

**Théorème 2.4.1 .**

1) L'estimateur  $\hat{S}_{FH}(t)$  possède un biais.

2) Si on suppose que :

i) pour tout  $s \in [0, t]$

$$n \int_0^s \frac{J_n(u)}{\bar{Y}_n(u)} h(u) du \xrightarrow{\mathcal{P}}_{n \rightarrow \infty} \sigma^2(s)$$

ii) pour tout  $\varepsilon > 0$

$$n \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} h(u) \mathbb{1}_{\{\sqrt{n} |\frac{J_n(u)}{\bar{Y}_n(u)}| > \varepsilon\}} du \xrightarrow{\mathcal{P}}_{n \rightarrow \infty} 0$$

iii)

$$\sqrt{n} \int_0^t (1 - J_n(u)) h(u) du \xrightarrow{\mathcal{P}}_{n \rightarrow \infty} 0$$

Alors pour tout  $t > 0$

$$\sqrt{n}(\hat{S}_{FH}(t) - S(t)) \Rightarrow_{n \rightarrow \infty} -U(t)S(t)$$

où  $U$  est une martingale gaussienne centrée.

**Preuve :** Rappelons que  $\hat{S}_{FH}(t) = \exp(-\hat{H}_{NA}(t))$  où  $\hat{H}_{NA}(t) = \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{N}_n(u)$

**A).** Calcul du biais : posons  $\tilde{S}_n(t) = e(-\tilde{H}_n(t))$  où  $\tilde{H}_n(t) = \int_0^t J_n(u)h(u)du$

Nous avons pour  $t \in [0, \tau[$  et  $\tau > 0$  (en appliquant le th. 3.3.3 Annexe)

$$\begin{aligned} \frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} - 1 &= - \int_0^t \frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} d(\hat{H}_{NA} - \tilde{H}_n)(u) \\ &= - \int_0^t \frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} d(\hat{H}_{NA} - \tilde{H}_n)(u) \\ &= - \int_0^t \frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} d\left(\sum_{i=1}^n \int_0^u \frac{J_n(s)}{\bar{Y}_n(s)} dM_i(s)\right) \\ &= - \int_0^t \frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} \frac{J_n(u)}{\bar{Y}_n(u)} d\left(\sum_{i=1}^n M_i(u)\right) \\ &= - \int_0^t \frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{M}_n(u) \end{aligned}$$

où  $d\bar{M}_n(t) = d\bar{N}_n(t) - \bar{Y}_n(t)h(u)du$ . Puisque  $\frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} - 1$  est une martingale locale de carré intégrable sur  $[0, \tau[$  nulle en  $t = 0$ , On en déduit que

$$\mathbf{E} \left( \frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} \right) = 1 \text{ pour tout } t \in [0, \tau[$$

On a

$$S(t) \leq \tilde{S}_n(t)$$

Donc

$$\frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} \leq \frac{\hat{S}_{FH}(t)}{S(t)}$$

et alors

$$\mathbf{E} \left( \frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} \right) \leq \mathbf{E} \left( \frac{\hat{S}_{FH}(t)}{S(t)} \right)$$

D'où :  $1 \leq \frac{\mathbf{E}(\hat{S}_{FH}(t))}{S(t)}$  ou encore  $\mathbf{E}(\hat{S}_{FH}(t)) \geq S(t)$ .

Par suite le biais est :  $\mathbf{E}(\hat{S}_{FH}(t)) - S(t) \geq 0$

**B).** Pour la convergence en loi vérifions les conditions du Th. de Robolledo :  
par la condition **1** on a

$$\langle \sqrt{n} \left( \frac{\hat{S}_{FH}}{\tilde{S}_n} - 1 \right) \rangle_{t=n} = n \int_0^t \frac{\hat{S}_{FH}(u^-)^2 J_n(u)}{\tilde{S}_n(u)^2 \bar{Y}_n(u)} h(u) du \xrightarrow[n \rightarrow \infty]{P} \sigma^2(s)$$

Par la preuve du Th.2.3.1 on en déduit  $\frac{\hat{S}_{FH}(u^-)}{\tilde{S}_n(u)} \xrightarrow[n \rightarrow \infty]{P} 1$ . D'où la 1<sup>ère</sup> condition du Th. de Robolledo.

Pour la 2<sup>ème</sup> condition du Th. de Robolledo, la condition **2** et de façon analogue à la preuve du Th 2.3.3 (2<sup>ème</sup> condition ) on trouve que

$$n \int_0^t \frac{\hat{S}_{FH}(u^-)^2 J(u)}{\tilde{S}_n(u)^2 \bar{Y}_n(u)} \mathbf{1}_{\{\sqrt{n} \left| \frac{J(u)}{\bar{Y}_n(u)} \right| > \varepsilon\}} h(u) du \xrightarrow[n \rightarrow \infty]{} 0$$

Donc pour chaque  $t > 0$

$$\sqrt{n} \int_0^t \frac{\hat{S}_{FH}(u^-) J(u)}{\tilde{S}_n(u) \bar{Y}_n(u)} d\bar{M}_n(u) \implies U(t)$$

où  $U(t) \hookrightarrow N(0, \sigma^2(t))$ .

Enfin

$$\sqrt{n} \left( \frac{\hat{S}_{FH}(t)}{\tilde{S}_n(t)} - 1 \right) \implies -U(t)$$

Le fait que  $S(u) \leq \tilde{S}_n(u)$ , pour  $u \in [0, s[$  et la condition **3** implique

$$\begin{aligned} \sqrt{n} \left| \frac{S(t)}{\tilde{S}_n(t)} - 1 \right| &\leq \sqrt{n} \int_0^t \frac{S(u)}{\tilde{S}_n(u)} d(H - \tilde{H}_n)(u) \\ &\leq \sqrt{n} \int_0^t (1 - J(u)) h(u) du \xrightarrow[n \rightarrow \infty]{P} 0 \end{aligned}$$

comme  $\tilde{S}_n(t) \xrightarrow[n \rightarrow \infty]{} S(t)$  on en déduit que

$$\sqrt{n} (\tilde{S}_n(t) - S(t)) \xrightarrow[n \rightarrow \infty]{P} 0$$

Il s'ensuit que

$$\begin{aligned} \sqrt{n} (\hat{S}_{FH}(t) - S(t)) &= \sqrt{n} (\hat{S}_{FH}(t) - \tilde{S}_n(t)) + \sqrt{n} (\tilde{S}_n(t) - S(t)) \\ &= \frac{\sqrt{n} (\hat{S}_{FH}(t) - \tilde{S}_n(t))}{\tilde{S}_n(t)} \tilde{S}_n(t) + \sqrt{n} (\tilde{S}_n(t) - S(t)) \\ &\implies_{n \rightarrow \infty} -U(t)S(t) \end{aligned}$$

D'où le résultat du théorème. •

### 2.4.3 Estimateur de Kaplan Meier

#### 1) Construction

*Kaplan* et *Meier* ont proposé un estimateur de  $S$  nommé aussi estimateur produit-limite. Il repose sur l'idée suivante : un "individu" est en vie après l'instant  $t$ , c'est être en vie juste avant l'instant  $t$  et ne pas "mourir" en  $t$ . Cette idée se traduit par les relations suivantes :

$$\begin{aligned} S(t) &= \mathbf{P}(T > t) \\ &= \mathbf{P}(T > t/T > t - 1)\mathbf{P}(T > t - 1) \\ &= \dots \\ &= \mathbf{P}(T > t/T > t - 1) \dots \mathbf{P}(T > 1/T > 0)\mathbf{P}(T > 0) \end{aligned}$$

Si l'on choisit les instants de conditionnement où il se produit un évènement  $t_i$  (mort, panne ou censure ...) on aura à estimer des quantités de la forme :

$$\mathbf{P}(T > t_{(i)} / T > t_{(i-1)}) = p_i$$

où les  $t_{(i)} < t_{(i-1)}$  et  $p_i$  est la probabilité de survivre pendant l'intervalle de temps  $I_i = ]t_{(i-1)}, t_{(i)}]$  sachant qu'on était "vivant" au début de cet intervalle.

Notons comme précédemment,  $r_i$  le nombre des sujets qui sont "vivants" (donc à risque) juste avant l'instant  $t_{(i)}$  et  $m_i$  le nombre d'"évènements" à l'instant  $t_{(i)}$ . Or  $q_i = 1 - p_i$  est la probabilité de la survenue de l'évènement" durant  $I_i$ . Un estimateur naturel de  $q_i$  est la fréquence

$$\hat{q}_i = \frac{m_i}{r_i}$$

Si on suppose qu'il n'y ait pas d'ex-aequo (plusieurs pannes en  $t_{(i)}$ ) : si  $D_i = 1$ , c'est qu'il y a eu un "évènement" en  $t_{(i)}$  et donc  $m_i = 1$ , si  $D_i = 0$ , c'est qu'il y a eu une censure en  $t_{(i)}$  et donc  $m_i = 0$ .

Par suite

$$\begin{aligned} \hat{p}_i &= 1 - \hat{q}_i = 1 - \frac{1}{r_i} \text{ si } D_i = 1 \\ &= 1 \text{ sinon} \end{aligned}$$

On a  $r_i = n - (i - 1)$  si  $D_i = 1$  (car il y'a eu  $i - 1$  "évènements" avant  $t_{(i)}$  et il y'a  $n$  individus dans l'étude). L'estimateur de *Kaplan-Meier* dans ce cas est

$$\hat{S}_{KM}(t) = \prod_{T_{(i)} \leq t} \left( 1 - \frac{1}{n - i + 1} \right)^{D_{(i)}}$$

**Remarque :** Pour  $\frac{m_i}{r_i}$  petit  $\hat{S}_{KM}(t) \approx \hat{S}_{FH}(t)$  car  $1 - x \approx e^{-x}$  au voisinage de zéro

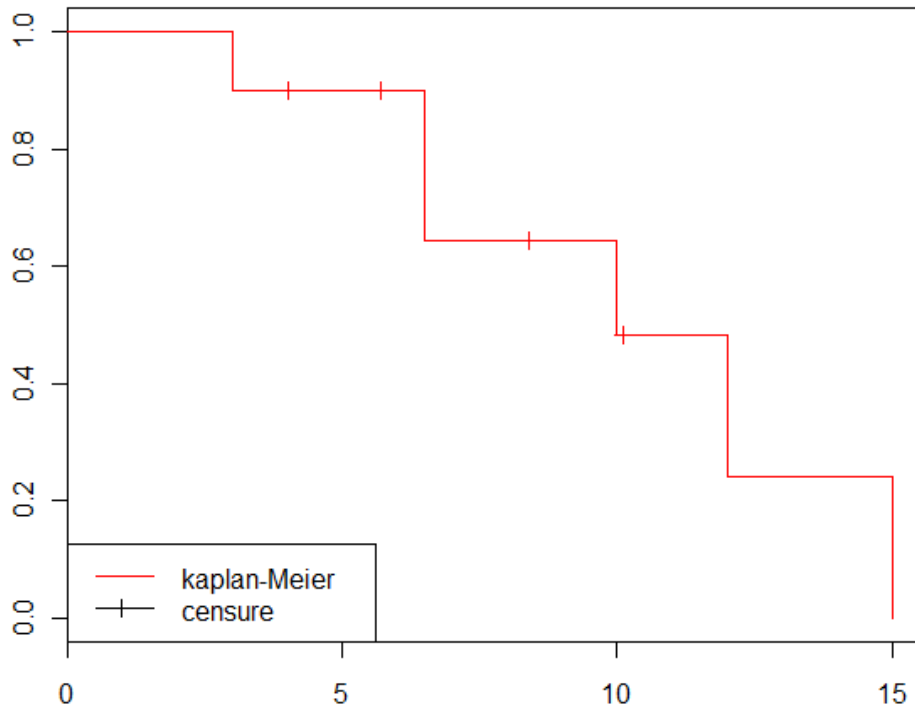
#### 2) Exemple

Sur 10 patients atteints de tumeurs, on a observé les durées de survie suivantes, exprimées en mois : 3, 4+, 5.7+, 6.5, 6.5, 8.4+, 10, 10.1+, 12, 15

L'estimateur de Kaplan-Meier de la fonction de survie  $S(t)$  se calcule de la manière suivante :



temps	$R_i$	$m_i$	Survie	Intervalle
0	10	0	1	$[0, 3[$
3	10	1	$9/10=0,9$	$[3, 6.5[$
6.5	7	2	$9/10*5/7=0,643$	$[6.5, 10[$
10	4	1	$9/10*5/7*3/4=0,482$	$[10, 12[$
12	2	1	$9/10*5/7*3/4*1/2=0,241$	$[12, 15[$
15	1	1	0	



### 3) Consistence

Le résultat suivant donne la consistance uniforme de l'estimateur  $\hat{S}_{KM}$  de *Kaplan-Meier*.

**Théorème 2.4.2** Si pour  $t > 0$ ,  $F(t) < 1$  et  $\bar{Y}_n(t) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \infty$  alors

$$\sup_{0 \leq s \leq t} \left| \hat{S}_{KM}(s) - S(s) \right| \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0$$

## 2.5 Estimation de la densité des durées de survie

### 2.5.1 Données non censurées

#### 1) Introduction

Dans cette partie nous rappelons quelques propriétés de l'estimateur à noyau de la densité sans démonstrations. Soit  $T_1, T_2, \dots, T_n$  des durées de survie de f.d.r.  $F$  et soit  $V_{x_0} = ]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}]$  un voisinage de  $t_0$ , où  $(b_n)_{n \geq 0}$  est suite de paramètres positifs appelée fenêtre.

Un estimateur de la densité  $f$  au point  $t_0$  est donné par

$$\begin{aligned} f_n(t_0) &= \frac{F_n(t_0 + \frac{b_n}{2}) - F_n(t_0 - \frac{b_n}{2})}{b_n} \\ &= \frac{\text{nombre d'évènements dans } ]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}]}{b_n} \end{aligned}$$

où  $F_n$  est la fonction de répartition empirique. On peut aussi écrire

$$\begin{aligned} f_n(t_0) &= \frac{1}{nb_n} \sum_{i=1}^n \mathbf{1}_{]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}]}(T_i) \\ &= \frac{1}{nb_n} \sum_{i=1}^n \mathbf{1}_{]-\frac{1}{2}, \frac{1}{2}]} \left( \frac{t_0 - T_i}{b_n} \right) \end{aligned}$$

Cette écriture fait apparaître un noyau de la forme

$$K(\cdot) = \mathbf{1}_{]-\frac{1}{2}, \frac{1}{2}]}(\cdot)$$

D'où la définition d'un estimateur à noyau de  $f$ , pour  $t \in ]-\frac{1}{2}, \frac{1}{2}]$

$$f_n(t) = \frac{1}{nb_n} \sum_{i=1}^n K\left(\frac{t - T_i}{b_n}\right)$$

où le noyau  $K$  est une application  $:\mathbb{R} \rightarrow \mathbb{R}^+$ , borné, d'intégrale égale à 1 et symétrique et  $b_n \rightarrow 0$ .

**Exemples de noyaux :** noyau gaussien, noyau uniforme, noyau d'Epanechnikov, noyau triangulaire, noyau quadratique, ...etc

Un noyau  $K$  est dit de Parzen-Roseblatt si  $K$  est symétrique et si  $\lim_{|u| \rightarrow \infty} |u| |K(u)| = 0$ .

Si on pose  $K_{b_n}(y) = \frac{1}{b_n} K\left(\frac{y}{b_n}\right)$ , l'estimateur à noyau  $f_n$  s'écrit

$$f_n(t) = \int_{-\infty}^{\infty} K_{b_n}(t-s) d\mu_n(s) = (K_{b_n} * \mu_n)(t)$$

où  $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{T_i}$ . L'étude de  $f_n$  repose en particulier sur le lemme de Bochner

#### 2) Comportement asymptotique de $f_n$

Nous avons les résultats suivants sur  $f_n$  [2].

On pose  $C_t$  l'ensemble des fonctions continues en un point  $t \in \mathbb{R}$ .

**Théorème 2.5.1.1** Si  $f_n$  est un estimateur à noyau et si le noyau  $K$  est de Parzen-Rosemblatt, alors

$$(b_n \rightarrow 0 \text{ et } nb_n \rightarrow \infty) \iff \left[ \forall t \in \mathbb{R}, \forall f \in C_t, \mathbf{E} [f_n(t) - f(t)]^2 \xrightarrow[n \rightarrow \infty]{} 0 \right]$$

**Théorème 2.5.1.2** Soit  $K$  un noyau positif et à variations bornées et  $f$  uniformément continue, on a

$$\left( b_n \rightarrow 0 \text{ et } \forall \gamma > 0, \sum_{n \geq 0} e^{-\gamma nb_n^2} < \infty \right) \iff \sup_t |f_n(t) - f(t)| \xrightarrow[n \rightarrow \infty]{} 0 \text{ p.co.}$$

**Théorème 2.5.1.3** On a la loi limite gaussienne :

$$\frac{f_n(t) - \mathbf{E}f_n(t)}{\sigma(f_n(t))} \xrightarrow[n \rightarrow \infty]{} N(0, 1)$$

### 3) Choix optimaux du noyau $K$ et de la fenêtre $b_n$

Les résultats de convergence de l'estimateur à noyau  $f_n$  nécessitent des conditions sur  $K$  et  $b_n$ , d'où se pose le problème de choix optimaux de  $(K, b_n)$ .

On pose

$$\begin{aligned} \Delta_n(t) &= \mathbf{E} [f_n(t) - f(t)]^2 \\ &= \mathbf{E} [f_n(t) - \mathbf{E}f_n(t) + \mathbf{E}f_n(t) - f(t)]^2 \\ &= [\mathbf{E}f_n(t) - f(t)]^2 + \mathbf{E} [f_n(t) - \mathbf{E}f_n(t)]^2 \end{aligned}$$

On suppose que le support du noyau  $K$  est  $[-1, 1]$

Le théorème suivant donne le comportement asymptotique du biais et de la variance de  $f_n$

**Théorème 2.5.1.4** Si  $f$  est de classe  $C^2$  et  $K$  est un noyau de Parzen-Rosemblatt. Alors si  $b_n \rightarrow 0$  et  $nb_n \rightarrow \infty$  on a

1.

$$\mathbf{E}f_n(t) - f(t) = \frac{b_n^2}{2} f''(t) \int_{-1}^1 u^2 K(u) du + O(b_n^2)$$

2.

$$\mathbf{E} [f_n(t) - \mathbf{E}f_n(t)]^2 = \frac{1}{nb_n} f(t) \int_{-1}^1 K^2(u) du + O\left(\frac{1}{nb_n}\right)$$

3.

$$\Delta_n(t) = \frac{b_n^4}{4} f''^2(t) \left[ \int_{-1}^1 u^2 K(u) du \right]^2 + \frac{1}{nb_n} f(t) \int_{-1}^1 K^2(u) du + O\left(b_n^4 + \frac{1}{nb_n}\right)$$

Le choix de la fenêtre optimale  $b_n^*$  se fait ici si  $K$  est donnée.

**Théorème 2.5.1.5** Sous les conditions du théorème précédent, la fenêtre optimale est

$$b_n^* = \left( \frac{f(t)}{f''^2(t) \left[ \int_{-1}^1 u^2 K(u) du \right]^2 \int_{-1}^1 K^2(u) du} \right)^{1/5} n^{-1/5}$$

Pour les démonstrations des théorèmes précédents

## 2.5.2 Données censurées

### 1) Introduction

Le but de cette partie est l'estimation de la densité des durées de survie  $T$  dans le cadre de données censurées voir 5.

Soit  $T_1, T_2, \dots, T_n$  une suite de v.a. positives i.i.d. de fonction de répartition  $F$  représentant les durées de survie et  $C_1, C_2, \dots, C_n$  est la suite de v.a i.i.d. représentant les censures de f.d.r.  $G$ . On suppose que la suite  $(T_i)$  est indépendante de la suite  $(C_i)$  et on observe

$$X_i = T_i \wedge C_i, D_i = \mathbf{1}_{\{T_i \leq C_i\}} i = 1, \dots, n$$

L'estimateur de Kaplan-Meier de la fonction de survie  $S$  est donné par

$$\hat{S}_{KM}(t) = \prod_{t^{(i)} \leq t} \left( 1 - \frac{1}{n - i + 1} \right)^{D_{(i)}}$$

qui s'écrit aussi

$$\hat{S}_{KM}(t) = \begin{cases} \prod_{X_{(i)} \leq t} \left( 1 - \frac{D_{(i)}}{n - i + 1} \right) & \text{si } t \leq X_{(n)} \\ 1 - \hat{F}_n(X_{(n)}) & \text{si } t > X_{(n)} \end{cases}$$

où  $X_{(1)} < X_{(2)} < \dots < X_{(n)}$  est la statistique d'ordre de l'échantillon  $X_1, X_2, \dots, X_n$  et  $D_{(i)}$  est la valeur de  $D_i$  (l'indicateur de censure à droite) associée à  $X_{(i)}$ . (Pour  $t > X_{(n)}$  on a posé  $1 - \hat{F}_n(t) = 1 - \hat{F}_n(X_{(n)})$  puisque il n'y a plus d'observations après  $t$ )

Nous reprenons les mêmes notations. On suppose que  $F$  admet une densité  $f$  par rapport à la mesure de Lebesgue qu'on se propose d'estimer en utilisant les observations  $(X_i, D_i)$ ,  $i = 1, \dots, n$ . En se basant sur l'estimateur de Kaplan-Meier, Blum et Susarla (1980) ont proposé un estimateur de la densité  $f$  par la méthode du noyau donné par

$$f_n(t) = \frac{1}{b_n} \int_0^\infty K\left(\frac{t-s}{b_n}\right) d\hat{F}_n(s) \quad (2.4)$$

où  $\hat{F}_n$  est une fonction empirique,  $(b_n)_{n \geq 1}$  est la fenêtre avec  $b_n \xrightarrow[n \rightarrow \infty]{} 0$  et  $K$  un noyau de support  $[-1, 1]$ . Cet estimateur  $f_n$  a été étudié en particulier par Blum et Susarla (1980), Földes et al. (1981) et Mielniczuk (1986). Dans cette partie on quelques propriétés asymptotiques de  $f_n$  en suivant [5].

Nous utilisons les notations suivantes :

Soit

$$\begin{aligned} S(t) &= 1 - F(t), S_C(t) = 1 - G(t) \\ \pi(t) &= \mathbf{P}(X_1 \geq t) = S(t)S_C(t^-) \\ h(t) &= \frac{f(t)}{S(t)}, H(t) = \int_0^t h(s)ds \end{aligned}$$

où  $h$  et  $H$  sont respectivement le taux de hasard et le taux de hasard cumulé.

On considère les processus ponctuels suivants définis sur  $[0, \infty[$  :

- Le processus  $N_i(t)$  indicateur de "survenue de l'évènement"  $N_i(t) = \mathbf{1}_{\{X_i \leq t, D_i=1\}}$
- Le processus  $\bar{N}_n(t) = \sum_{i=1}^n N_i(t)$  : le nombre d'"évènements" observés dans l'intervalle  $]0, t]$
- Le processus risque :  $Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}$

– Le processus  $\bar{Y}_n(t) = \sum_{i=1}^n Y_i(t)$  : le nombre d'"individus" à risque à l'instant  $t$   
On note aussi les processus suivants :

$$\begin{aligned}\Lambda(t) &= \int_0^t \bar{Y}_n(s) dH(s) \text{ le processus intensité cumulée} \\ \bar{M}_n(t) &= \bar{N}_n(t) - \Lambda(t) = \bar{N}_n(t) - \int_0^t \bar{Y}_n(s) h(s) ds \\ \hat{H}_{NA} &= \int_0^t \frac{d\bar{N}_n(u)}{\bar{Y}_n(u)} \\ \tilde{H}_n(t) &= \int_0^t J_n(u) h(u) du \\ J_n(t) &= \mathbf{1}_{\{\bar{Y}_n(t) > 0\}} \\ \tilde{f}_n(t) &= \frac{1}{b_n} \int_0^\infty K\left(\frac{t-s}{b_n}\right) \hat{S}_{KM}(s^-) d\tilde{H}_n(s)\end{aligned}\tag{2.5}$$

Le processus  $(M(t), t \geq 0)$  est une  $\mathcal{F}_t$ -martingale où

$$\mathcal{F}_t = \sigma(\mathbf{1}_{\{X_i \leq s, D_i=1\}}, \mathbf{1}_{\{X_i \leq s, D_i=0\}}; 0 \leq s \leq t, i = 1, \dots, n)$$

$(M(t), t \geq 0)$  est aussi une martingale locale de carré intégrable (cf.3.3.4 Annexe).

On suppose que :

– **C1** : Le noyau  $K$  est borné sur  $[-1, 1]$  et satisfait les conditions suivantes

$$\int_{-1}^1 K(t) dt = 1, \quad \int_{-1}^1 tK(t) dt = 0, \quad \int_{-1}^1 t^2 K(t) dt = k_2 \neq 0$$

## 2) Etude du biais de l'estimateur $\hat{f}_n$

On note dorénavant l'estimateur  $f_n$  dans (2.4) par  $\hat{f}_n$

Dans le cas général où  $F$  est une f.d.r. quelconque, la fonction de survie  $S(t) = e\{-H(t)\}$  est l'unique solution de l'équation intégrale suivante :

$$S(t) = 1 - F(t) = 1 - \int_0^t S(s^-) dH(s)$$

et l'estimateur de Kaplan-Meier Vérifie

$$\hat{S}_{KM}(t) = 1 - \int_0^t \hat{S}_{KM}(s^-) d\hat{H}_{NA}(s) =: 1 - \hat{F}_n(t)$$

où encore

$$d\hat{S}_{KM}(t) = -\hat{S}_{KM}(t^-) d\hat{H}_{NA}(t)$$

En effet

$$\begin{aligned}\hat{S}_{KM}(t) - \hat{S}_{KM}(t^-) &= \Delta \hat{S}_{KM}(t) \\ &= -\hat{S}_{KM}(t^-) \left[ \hat{H}_{NA}(t) - \hat{H}_{NA}(t^-) \right] \\ &= : -\hat{S}_{KM}(t^-) \Delta \hat{H}_{NA}(t) \\ &= -\hat{S}_{KM}(t^-) \frac{\Delta \bar{N}_n(t)}{\bar{Y}_n(t)} \text{ par construction de } \hat{H}_{NA} \text{ (cf chap.3)}\end{aligned}$$

D'où

$$\hat{S}_{KM}(t) = \hat{S}_{KM}(t^-) \left( 1 - \frac{\Delta \bar{N}_n(t)}{\bar{Y}_n(t)} \right)$$

où  $\Delta \bar{N}(t) = \bar{N}(t) - \bar{N}(t^-)$ . Comme  $\hat{H}_{NA}$  présente des sauts seulement aux points  $t_i$ , on obtient

$$\hat{S}_{KM}(t) = \prod_{t_i \leq t} \left( 1 - \frac{\Delta \bar{N}(t_i)}{\bar{Y}(t_i)} \right) = \prod_{t_i \leq t} \left( 1 - \Delta \hat{H}_{NA}(t_i) \right)$$

Ainsi de l'équation  $\hat{S}_{KM}(t) = 1 - \hat{F}_n(t)$ , l'estimateur à noyau (2.4) peut s'écrire -en changeant de notation- comme suit :

$$f_n(t) = \hat{f}_n(t) = \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \hat{S}_{KM}(s^-) d\hat{H}_{NA}(s) \quad (2.6)$$

car  $d\hat{F}_n(t) = -d\hat{S}_{KM}(t)$ .

Nous avons montré que  $\hat{H}_{NA}(t) - \tilde{H}_n(t) = \int_0^t \frac{J_n(u)}{\bar{Y}_n(u)} d\bar{M}_n(u)$ . Par suite de (2.4) et (2.6) on a

$$\begin{aligned} \hat{f}_n(t) - \tilde{f}_n(t) &= \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \hat{S}_{KM}(s^-) d(\hat{H}_{NA} - \tilde{H}_n)(s) \\ &= \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \hat{S}_{KM}(s^-) \frac{J_n(s)}{\bar{Y}_n(s)} d\bar{M}_n(s) \\ &= \sum_{i=1}^n \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \hat{S}_{KM}(s^-) \frac{J_n(s)}{\bar{Y}_n(s)} dM_i(s) \\ &= \sum_{i=1}^n \int_0^\infty L_i(s) dM_i(s) \end{aligned} \quad (2.7)$$

Comme  $(L_i(s))$  sont des processus localement bornés (sur  $[0, t]$ ) et  $\mathcal{F}_s$ -prévisibles, et  $M_i(s)$  est une martingale locale de carré intégrable (cf. corol.3.3.4 Annexe), alors par le Th. 2.4.5 de [3]  $(\hat{f}_n(t) - \tilde{f}_n(t), t \geq 0)$  est une martingale locale de carré intégrable.

On note

$$f_n^*(t) = \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) f(s) ds = \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) S(s) dH(s)$$

On a le résultat suivant

**Théorème 2.5.2.1** *Supposons que  $f$  est continue.*

(i) Si  $b_n \xrightarrow{n \rightarrow \infty} 0$  alors

$$f_n^*(t) - f(t) = \int_{-1}^1 K(u) [f(t - b_n u) - f(t)] du \xrightarrow{n \rightarrow \infty} 0$$

(ii) Pour tout  $t \geq 0$  on a

$$\mathbf{E} \hat{f}_n(t) = \mathbf{E} \tilde{f}_n(t) = \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \mathbf{E} \left( \hat{S}_{KM}(s^-) J_n(s) \right) dH(s)$$

(iii) Si la fonction  $\pi(\cdot) = \mathbf{P}(X_1 \geq \cdot)$  est strictement positive dans un voisinage d'un point  $t$ , alors on a

$$\begin{aligned} \mathbf{E}(\hat{f}_n(t) - f_n^*(t)) &= \mathbf{E}(\tilde{f}_n(t) - f_n^*(t)) \\ &= \frac{1}{b_n} \int_0^\infty K \left( \frac{t-s}{b_n} \right) \mathbf{E} \left( \mathbf{1}_{\{\tau_{\bar{Y}} < s\}} \frac{\hat{S}_{KM}(\tau_{\bar{Y}}) S(s)}{S(\tau_{\bar{Y}})} \right) dH(s) \end{aligned}$$

où  $\tau_{\bar{Y}} = \inf \{s : \bar{Y}(s) = 0\}$

(v) Sous les mêmes conditions de (iii), on a, pour  $n$  assez grand

$$\left| \mathbf{E}(\hat{f}_n(t) - f_n^*(t)) \right| \leq e^{-n\pi(t+b_n)} \int_{-1}^1 |K(u)| h(t - b_n u) du$$

ie  $\left| \mathbf{E}(\hat{f}_n(t) - f_n^*(t)) \right| \xrightarrow[n \rightarrow \infty]{} 0$  avec une vitesse exponentielle.

(v) Si les conditions de (iii) sont vérifiées,  $f$  de classe  $C^2$  et  $b_n \rightarrow 0$  alors

$$\mathbf{E}(\hat{f}_n(t) - f(t)) = \frac{1}{2} b_n^2 f''(t) k_2 + o(b_n^2) + o\left(\frac{1}{n}\right)$$

où  $k_2$  est donné par **C1**.

**Remarque** le point (v) montre que l'estimateur  $\hat{f}_n$  est asymptotiquement sans biais.

### 3) Consistence et normalité asymptotique de $\hat{f}_n$

Le théorème suivant donne la convergence en probabilité de l'estimateur  $\hat{f}_n$ .

**Théorème 2.5.2.2** Si  $f$  est continue en  $t$  et  $\pi$  est positive au voisinage de  $t$ ,  $b_n \rightarrow 0$  et  $nb_n \rightarrow \infty$  alors

$$\hat{f}_n(t) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} f(t)$$

Le théorème suivant donne la normalité asymptotique de  $\hat{f}_n$ .

**Théorème 2.5.2.3** Supposons que  $f$  est continue en un point  $t$  et  $\pi$  est continue et strictement positive.

Si  $b_n \rightarrow 0$  et  $nb_n \rightarrow \infty$ , alors

$$\sqrt{nb_n} \left( \hat{f}_n(t) - f_n^*(t) \right) \xRightarrow[n \rightarrow \infty]{} N(0, \sigma^2(t))$$

$$\sigma^2(t) = \frac{f(t)}{S_C(t^-)} \int_{-1}^1 K^2(u) du$$

**Preuve** On décompose :

$$\sqrt{nb_n} (\hat{f}_n(t) - f_n^*(t)) = \sqrt{nb_n} (\hat{f}_n(t) - \tilde{f}_n(t)) + \sqrt{nb_n} (\tilde{f}_n(t) - f_n^*(t))$$

Montrons que le premier terme à une loi limite et le deuxième terme tend vers 0 en proba.

Par (2.7) on peut écrire en posant :

$$\begin{aligned} U_n(t) &::= \sqrt{nb_n} (\hat{f}_n(t) - \tilde{f}_n(t)) \\ &= \sqrt{nb_n^{-1}} \int_0^\infty K\left(\frac{t-s}{b_n}\right) \hat{S}_{KM}(s^-) \frac{J_n(s)}{\bar{Y}_n(s)} d\bar{M}_n(s) \\ &:= \int_0^\infty L_n(t, s) d\bar{M}_n(s) \end{aligned} \tag{2.8}$$

Le processus croissant de  $(U_n(t), t \geq 0)$  est donné par (cf Th. 2.4.3 de [3] p. 70)

$$\begin{aligned} \langle U_n \rangle_t &= \int_0^\infty L_n^2(t, s) d\langle \bar{M} \rangle_s \\ &= \int_0^\infty L_n^2(t, s) \bar{Y}_n(s) dH(s) \\ &= \int_{-1}^1 nK^2(u) \hat{S}_{KM}^2((t - b_n u)^-) \frac{J_n(t - b_n u)}{\bar{Y}_n(t - b_n u)} h(t - b_n u) du \end{aligned}$$

Posons

$$\begin{aligned}
A_1(t) & : = \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) \frac{J_n(t-b_n u)}{\frac{1}{n} \bar{Y}_n(t-b_n u)} h(t) du \\
A_2(t) & : = \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) \frac{J_n(t-b_n u)}{\pi(t-b_n u)} h(t) du \\
A_3(t) & : = \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) \frac{J_n(t-b_n u)}{\pi(t)} h(t) du \\
A_4(t) & : = \int_{-1}^1 K^2(u) S^2((t-b_n u)^-) \frac{J_n(t-b_n u)}{\pi(t)} h(t) du \\
A_5(t) & : = \int_{-1}^1 K^2(u) S^2(t) \frac{J_n(t-b_n u)}{\pi(t)} h(t) du
\end{aligned}$$

En utilisant  $\hat{S}_{KM}(t) \leq 1$  et  $J_n(t) \leq 1$ , on a

$$\begin{aligned}
| \langle U_n \rangle_t - A_1(t) | & = \left| \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) \frac{J_n(t-b_n u)}{\frac{1}{n} \bar{Y}_n(t-b_n u)} [h(t-b_n u) - h(t)] du \right| \\
& \leq \frac{1}{\frac{1}{n} \bar{Y}_n(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} |h(s) - h(t)| \int_{-1}^1 K^2(u) du \quad (2.9)
\end{aligned}$$

De même on majore la différence :

$$\begin{aligned}
|A_1(t) - A_2(t)| & = \left| \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) J_n(t-b_n u) h(t) \left( \frac{1}{\frac{1}{n} \bar{Y}_n(t-b_n u)} - \frac{1}{\pi(t-b_n u)} \right) du \right| \\
& \leq \frac{h(t)}{\pi(t+b_n) n^{-1} \bar{Y}_n(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} \left| \frac{1}{n} \bar{Y}_n(s) - \pi(s) \right| \int_{-1}^1 K^2(u) du
\end{aligned}$$

De même on majore la différence

$$\begin{aligned}
|A_2(t) - A_3(t)| & = \left| \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t-b_n u)^-) J_n(t-b_n u) h(t) \left[ \frac{1}{\pi(t-b_n u)} - \frac{1}{\pi(t)} \right] du \right| \\
& \leq \frac{h(t)}{\pi(t)\pi(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} |\pi(t) - \pi(s)| \int_{-1}^1 K^2(u) du
\end{aligned}$$

De même on a

$$\begin{aligned}
|A_3(t) - A_4(t)| & = \left| \int_{-1}^1 K^2(u) \left[ \hat{S}_{KM}^2((t-b_n u)^-) - S^2((t-b_n u)^-) \right] \frac{J_n(t-b_n u)}{\pi(t)} h(t) du \right| \\
& \leq 2 \frac{h(t)}{\pi(t)} \sup_{s \in [t-b_n, t+b_n]} \left| \hat{S}_{KM}(s) - S(t) \right| \int_{-1}^1 K^2(u) du
\end{aligned}$$

Aussi on a

$$\begin{aligned}
|A_4(t) - A_5(t)| & = \left| \int_{-1}^1 K^2(u) \left[ S^2((t-b_n u)^-) - S^2(t) \right] \frac{J_n(t-b_n u)}{\pi(t)} h(t) du \right| \\
& \leq 2 \frac{h(t)}{\pi(t)} \sup_{s \in [t-b_n, t+b_n]} |S(t) - S(s)| \int_{-1}^1 K^2(u) du
\end{aligned}$$



Enfin, en remarquant que  $\pi(t) = S_C(t)S(t)$  et que  $S(t) = f(t)/h(t)$  on a

$$\begin{aligned}
|A_5(t) - \sigma^2(t)| &= \left| \frac{h(t)S^2(t)}{\pi(t)} \int_{-1}^1 K^2(u) J_n(t - b_n u) du - \frac{f(t)}{S_C(t^-)} \int_{-1}^1 K^2(u) du \right| \\
&= \frac{f(t)}{S_C(t^-)} \int_{-1}^1 K^2(u) [1 - J_n(t - b_n u)] du \\
&\leq \frac{f(t)}{S_C(t^-)} \mathbf{1}_{\{\bar{Y}_n(t+b_n)=0\}} \int_{-1}^1 K^2(u) du
\end{aligned} \tag{2.10}$$

En utilisant les majorations (2.9) à (2.10) on obtient

$$\begin{aligned}
|< U_n >_t - \sigma^2(t)| &\leq |< U_n, U_n >_t - A_1(t)| + |A_1(t) - A_2(t)| + |A_2(t) - A_3(t)| \\
&\quad |A_3(t) - A_4(t)| + |A_4(t) - A_5(t)| + |A_5(t) - \sigma^2(t)| \\
&\leq \frac{1}{\frac{1}{n}\bar{Y}_n(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} |h(s) - h(t)| \int_{-1}^1 K^2(u) du \\
&\quad + \frac{h(t)}{\pi(t+b_n)n^{-1}\bar{Y}_n(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} \left| \frac{1}{n}\bar{Y}(s) - \pi(s) \right| \int_{-1}^1 K^2(u) du \\
&\quad + \frac{h(t)}{\pi(t)\pi(t+b_n)} \sup_{s \in [t-b_n, t+b_n]} |\pi(s) - \pi(t)| \int_{-1}^1 K^2(u) du \\
&\quad + 2 \frac{h(t)}{\pi(t)} \sup_{s \in [t-b_n, t+b_n]} \left| \hat{S}_{KM}(s) - S(t) \right| \int_{-1}^1 K^2(u) du \\
&\quad + 2 \frac{h(t)}{\pi(t)} \sup_{s \in [t-b_n, t+b_n]} |S(s) - S(t)| \int_{-1}^1 K^2(u) du \\
&\quad + \frac{f(t)}{S_C(t^-)} \mathbf{1}_{\{\bar{Y}_n(t+b_n)=0\}} \int_{-1}^1 K^2(u) du
\end{aligned}$$

Chacun des majorants tend vers 0 en probabilité. En effet :

- Pour le premier, on utilise :  $\frac{1}{n}\bar{Y}(t+b_n) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} (1 - F(t))(1 - G(t^-))$  uniformément, la continuité de  $h$  et la condition **C1**.
- Pour le deuxième, on utilise  $\frac{1}{n}\bar{Y}(t+b_n) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} (1 - F(t))(1 - G(t^-))$  uniformément, la continuité de  $\pi$  et **C1**.
- Pour le troisième, on utilise la continuité de  $\pi$  et **C1**.
- Pour le quatrième, on utilise le Th. 3.4.2 [3] p.115 et **C1**.
- Pour le cinquième, on utilise la continuité de  $S$  et **C1**.
- Pour le sixième, on utilise  $\bar{Y}_n(t+b_n) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \infty$  (car  $\frac{1}{n}\bar{Y}_n(t+b_n) \rightarrow^{\mathbf{P}} (1 - F(t))(1 - G(t^-))$ ) et **C1**.

Par suite

$$< U_n >_t \xrightarrow[n \rightarrow \infty]{\mathbf{P}} \sigma^2(t) \tag{2.11}$$

Pour établir la loi limite on utilise le th. de Robolledo.

On pose

$$U_{n,\varepsilon}(t) := \int_0^\infty L_n(t, s) \mathbf{1}_{\{|L_n(t,s)| \geq \varepsilon\}} d\bar{M}_n(s)$$

où  $L_n$  est donnée par (2.8). Par le théorème 2.4.3 de [3], p. 70 on a

$$\begin{aligned} &< U_{n,\varepsilon} >_t = \int_0^\infty L_n^2(t, s) \mathbf{1}_{\{|L_n(t,s)| \geq \varepsilon\}} \bar{Y}_n(s) dH(s) \\ &= \int_{-1}^1 K^2(u) \hat{S}_{KM}^2((t - b_n u)^-) \frac{J_n(t - b_n u)}{n^{-1} \bar{Y}_n(t - b_n u)} h(t - b_n u) \mathbf{1}_{A(n,u)} du \\ &\leq \frac{1}{n^{-1} \bar{Y}_n(t + b_n)} \left[ \sup_{s \in [t - b_n, t + b_n]} h(s) \right] \mathbf{1}_{B(n)} \int_{-1}^1 K^2(u) du \end{aligned}$$

où

$$A(n, u) = \left\{ \left| K(u) \hat{S}_{KM}((t - b_n u)^-) \frac{J_n(t - b_n u)}{n^{-1} \bar{Y}_n(t - b_n u)} \right| \geq \varepsilon \sqrt{nb_n} \right\}$$

et

$$B(n) = \left\{ \frac{1}{n} \bar{Y}_n(t + b_n) \sup_{u \in [-1, 1]} |K(u)| \geq \varepsilon \sqrt{nb_n} \right\}$$

Le théorème de Glivenko-Cantelli implique que

$$\frac{1}{n} \bar{Y}_n(t + b_n) \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} S(t) (1 - G(t))$$

et comme  $nb_n \rightarrow \infty$  on en déduit que

$$\mathbf{1}_{B(n)} \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} 0$$

Par suite

$$< U_{n,\varepsilon} >_t \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} 0 \quad (2.12)$$

Par conséquent, de (2.11) et (2.12) et par le théorème de Rebolledo on a

$$U_n(t) \equiv \sqrt{nb_n} \left( \hat{f}_n(t) - \tilde{f}_n(t) \right) \Longrightarrow_{n \rightarrow \infty} N(0, \sigma^2(t)) \quad (2.13)$$

Montrons maintenant que

$$\sqrt{nb_n} \left( \tilde{f}_n(t) - f_n^*(t) \right) \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} 0 \quad (2.14)$$

Par la démonstration du théorème 2.5.2.2(cf [4] p67 (2.9)) on a

$$\begin{aligned} \sqrt{nb_n} \left| \tilde{f}_n(t) - f_n^*(t) \right| &\leq \sqrt{nb_n} \left[ \sup_{s \in [t - b_n, t + b_n]} h(s) \sup_{s \in [t - b_n, t + b_n]} \left| \hat{S}_{KM}(s) - S(s) \right| \right] \int_{-1}^1 |K(u)| du \\ &\quad + \sqrt{nb_n} \mathbf{1}_{\{\bar{Y}(t + ub_n) = 0\}} \left[ \sup_{s \in [t - b_n, t + b_n]} f(s) \right] \int_{-1}^1 |K(u)| du \end{aligned} \quad (2.15)$$

Pour le deuxième terme, nous avons du fait que  $1 - x \leq e^{-x}$  :

$$\mathbf{E} \left( \sqrt{nb_n} \mathbf{1}_{\{\bar{Y}(t + ub_n) = 0\}} \right) = \sqrt{nb_n} (1 - \pi(t + b_n))^n \leq \sqrt{nb_n} e^{-n\pi(t + b_n)}$$

puisque  $\pi$  est positive au voisinage de  $t$ . Donc

$$\sqrt{nb_n} \mathbf{1}_{\{\bar{Y}(t + ub_n) = 0\}} \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} 0$$

Par conséquent

$$\sqrt{nb_n} \mathbf{1}_{\{\bar{Y}(t+ub_n)=0\}} \left[ \sup_{s \in [t-b_n, t+b_n]} f(s) \right] \int_{-1}^1 |K(u)| \xrightarrow{\mathbf{P}}_{n \rightarrow \infty} 0$$

Pour le premier terme, par le théorème 6.3.1 de [3] p. 235, on a : pour  $\varepsilon > 0$

$$\sqrt{n} \left( \hat{S}_{KM}(\cdot) - S(\cdot) \right) \Longrightarrow_{n \rightarrow \infty} S(\cdot) W(v(\cdot))$$

sur  $D[t - \varepsilon, t + \varepsilon]$  où  $v(t) \equiv \int_0^t [\pi(s)]^{-1} dH(s)$  et  $W$  est un processus de Wiener et  $D[t - \varepsilon, t + \varepsilon]$  est l'espace de Skorokhod (l'espace des fonctions cadlag.)

Ainsi par la continuité de la fonction sup

$$\sup_{s \in [t-\varepsilon, t+\varepsilon]} \sqrt{n} \left| \hat{S}_{KM}(\cdot) - S(s) \right| \Longrightarrow_{n \rightarrow \infty} \sup_{s \in [t-\varepsilon, t+\varepsilon]} S(s) |W(v(s))|$$

De la propriété du scaling du mouvement brownien on a :

$$S(s)W(v(s)) \stackrel{\mathcal{L}}{=} W(S^2(s)v(s))$$

et que

$$\sup_{s \in [t-\varepsilon, t+\varepsilon]} S(s) |W(v(s))| \stackrel{\mathcal{L}}{=} \sup_{s \in [t-\varepsilon, t+\varepsilon]} |W^*(s)|$$

où  $W^*$  est un autre mouvement brownien. On sait que

$$\sup_{s \in [0, t]} |W^*(s)| \stackrel{\mathcal{L}}{=} \sqrt{t} |W^*(1)|$$

et donc bornée en probabilité. On en déduit que pour  $t$  fixé et  $n \rightarrow \infty$

$$\sup_{s \in [t-\varepsilon, t+\varepsilon]} \sqrt{n} \left| \hat{S}_{KM}(s) - S(s) \right| = O_P(1)$$

et par suite si  $n \rightarrow \infty$  et  $b_n \rightarrow 0$

$$\sqrt{nb_n} \sup_{s \in [t-b_n, t+b_n]} |S_n(s) - S(s)| = O_P(\sqrt{b_n})$$

Donc de (2.15) on conclut que

$$\sqrt{nb_n} \left| \tilde{f}_n(t) - f_n^*(t) \right| \xrightarrow{P}_{n \rightarrow \infty} 0$$

Enfin avec (2.13) et (2.14) on obtient

$$\sqrt{nb_n} \left( \hat{f}_n(t) - f_n^*(t) \right) \Longrightarrow_{n \rightarrow \infty} N(0, \sigma^2(t))$$

Ce qui termine la démonstration.

#### 4) Erreur quadratique intégrée et fenêtre optimale

L'erreur quadratique intégrée permet de choisir une fenêtre optimale. Soit  $\tau$  tel que  $\pi(\tau) > 0$ , et on définit, pour  $\varepsilon > 0$  l'erreur quadratique intégrée de  $\hat{f}_n$  sur l'intervalle  $[0, \tau - \varepsilon]$  par :

$$\mathbf{EQI}(\hat{f}_n) = \mathbf{E} \left( \int_0^{\tau-\varepsilon} [\hat{f}_n(t) - f(t)]^2 dt \right)$$

Le théorème suivant résume quelques propriétés de  $\mathbf{EQI}(\hat{f}_n)$ .

**Théorème 2.5.2.4** *Supposons que  $f$  est continue sur  $[0, \tau]$  où  $\tau$  est tel que  $\pi(\tau) > 0$ .*

(i) *Si  $b_n \rightarrow 0$  on a*

$$\mathbf{EQI}(\hat{f}_n) = \int_0^{\tau-\varepsilon} [f_n^*(t) - f(t)]^2 dt + \int_0^{\tau-\varepsilon} \mathbf{E} [\hat{f}_n(t) - f_n^*(t)]^2 dt + o\left(\frac{1}{n}\right) \quad (2.16)$$

(ii) *Si  $f$  est de classe  $C^2$  sur  $[0, \tau]$ , alors si  $b_n \rightarrow 0$  le premier terme de (2.16) s'écrit*

$$\int_0^{\tau-\varepsilon} [f_n^*(t) - f(t)]^2 dt = \frac{k_2^2}{4} b_n^4 \int_0^{\tau-\varepsilon} [f''(t)]^2 dt + o(b_n^4)$$

où  $k_2$  est donné par **C1**.

(iii) *Si  $\pi$  est continue sur  $[0, \tau]$ , alors si  $b_n \rightarrow 0$  et  $nb_n \rightarrow \infty$ , le deuxième terme de (2.16) s'écrit*

$$\int_0^{\tau-\varepsilon} \mathbf{E} [\hat{f}_n(t) - f_n^*(t)]^2 dt = \frac{1}{nb_n} \int_{-1}^1 K^2(t) dt \int_0^{\tau-\varepsilon} \frac{f(t)}{S_C(t^-)} dt + o\left(\frac{1}{nb_n}\right)$$

(iv) *Si  $f$  est de classe  $C^2$  sur  $[0, \tau]$  et  $\pi$  est continue sur  $[0, \tau]$ , alors si  $b_n \rightarrow 0$  et  $nb_n \rightarrow \infty$ , on a*

$$\mathbf{EQI}(\hat{f}_n) = \frac{k_2^2}{4} b_n^4 \int_0^{\tau-\varepsilon} [f''(t)]^2 dt + \frac{1}{nb_n} \int_{-1}^1 K^2(t) dt \int_0^{\tau-\varepsilon} \frac{f(t)}{S_C(t^-)} dt + o(b_n^4) + o\left(\frac{1}{nb_n}\right) \quad (2.17)$$

On a le résultat suivant :

**Corollaire 2.5.2.5** *La fenêtre optimale asymptotique est*

$$b_n^* = \frac{1}{n^{1/5}} \frac{1}{k_2^{2/5}} \left[ \int_{-1}^1 K^2(t) dt \int_0^{\tau-\varepsilon} \frac{f(t)}{S(t^-)} dt \right]^{1/5} \left[ \int_0^{\tau-\varepsilon} [f''(t)]^2 dt \right]^{-1/5} = O\left(\frac{1}{n^{1/5}}\right)$$

**Remarque :** Si on injecte l'expression de la fenêtre optimale  $b_n^*$  dans (2.17) on obtient l'expression optimale de l' $\mathbf{EQI}(\hat{f}_n)$

$$\begin{aligned} \mathbf{EQI}(\hat{f}_n) &= \frac{5}{4} \frac{k_2^{2/5}}{n^{4/5}} \left[ \int_{-1}^1 K^2(t) dt \int_0^{\tau-\varepsilon} \frac{f(t)}{S(t^-)} dt \right]^{4/5} \left[ \int_0^{\tau-\varepsilon} [f''(t)]^2 dt \right]^{1/5} + o(n^{-4/5}) \\ &= O\left(\frac{1}{n^{4/5}}\right) + o\left(\frac{1}{n^{4/5}}\right) = O\left(\frac{1}{n^{4/5}}\right) \end{aligned}$$

## Chapitre 3

# Simulation

Dans cette partie nous allons illustrer le comportement des estimateurs étudiés dans le chapitre 2 sur des données réelles. Pour cela, nous prenons deux groupes de malades atteints de CGD (chronic granulomatous disease) ([3] p.377-383). L'un est traité par le  $rIFN_\gamma$  et l'autre par effet placebo. Le but étant d'étudier la survie dans les deux groupes et de là conclure sur l'efficacité du traitement.

**traitement 1 ( $rIFN_\gamma$ )**

Notons  $censure=0$  si le patient est perdu de vue et  $censure=1$  si le patient décède au jour  $j$ .

Pour les malades traité par le  $rIFN_\gamma$  sur une durée de 500 jours on a le tableau suivant :

jour	censure
157	1
166	1
198	1
212	1
226	0
235	1
255	1
274	1
277	1
320	1
330	1
331	1
331	1
339	1
345	1
358	0
359	1
360	1
366	1
372	0
376	0
380	1
382	0
383	0
386	0
389	0
389	0
389	0

jour	censure
389	0
389	0
389	0
390	0
390	0
390	0
391	0
394	0
396	0
396	0
397	0
398	1
398	0
400	0
401	0
401	0
403	0
404	0
404	0
407	0
409	1
410	0
411	0
414	0
415	0
415	0
417	0
417	0

jour	censure
420	0
421	0
421	0
424	0
424	0
425	0
426	0
426	0
426	0
426	0
426	0
426	0
430	0
432	0
432	0
432	0
432	0
435	0
436	0
437	0
437	0
443	0
444	0
446	1
449	0
450	0

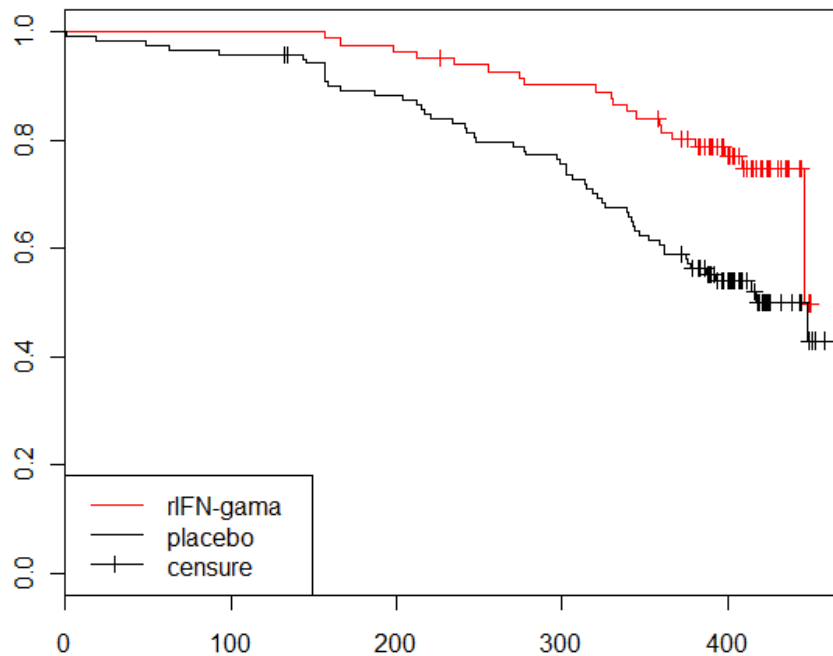
## traitement 2 (placebo)

Pour les malades traité par le effet placebo sur une durée de 500 jours on a le tableau suivant :

jour	censure	jour	censure	jour	censure	jour	censure
1	1	278	0	382	0	409	0
19	1	297	1	383	0	411	0
49	1	299	1	383	0	414	1
63	1	302	1	386	0	416	0
93	1	302	1	386	0	417	1
132	0	306	1	386	0	418	0
134	0	314	1	386	0	418	0
144	1	315	1	387	1	419	0
145	1	318	1	388	0	421	0
157	1	321	1	388	0	421	0
157	1	324	1	389	0	422	0
157	1	326	1	389	0	422	0
157	1	339	1	389	0	423	0
159	1	340	1	390	0	424	0
166	1	342	1	392	0	425	0
187	1	343	1	393	1	425	0
204	1	344	1	394	0	425	0
212	1	347	1	396	0	425	0
215	1	352	1	397	0	426	0
217	1	359	1	400	0	432	0
221	1	362	1	401	0	439	0
234	1	362	1	401	0	443	0
241	1	362	0	401	0	444	0
242	1	372	0	401	0	448	1
247	1	375	1	402	0	449	0
248	1	376	1	402	0	451	0
270	1	378	1	403	0	453	0
277	1	379	0	404	0	453	0
278	1	382	0	407	0	458	0
278	0	382	0	408	0	472	0

### 3.1 Estimation de la fonction de survie S

On calcule l'estimateur de Kaplan-Meier  $\hat{S}_{KM}(t) = \prod_{T_{(i)} \leq t} \left(1 - \frac{1}{n-i+1}\right)^{D_{(i)}}$  (cf chap.2) pour les données précédents. On a le graphe suivant :

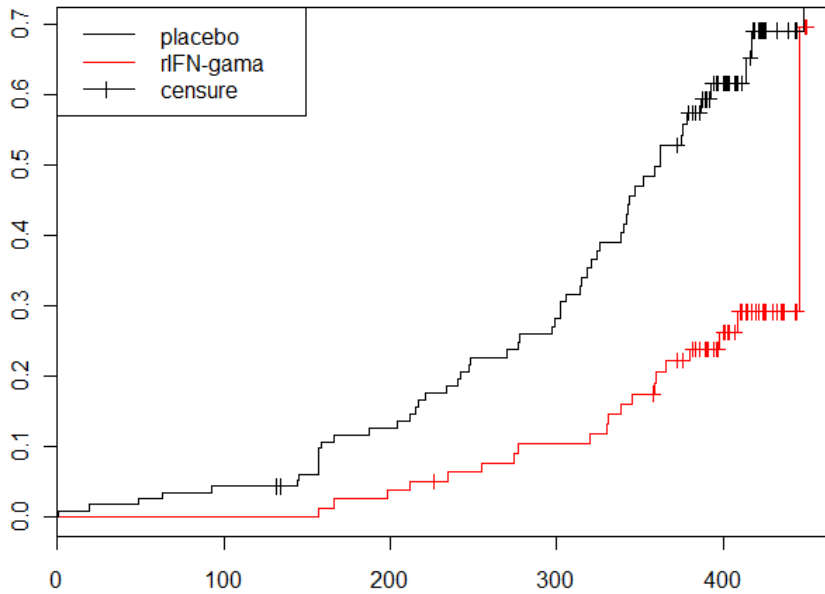


Nous remarquons que la courbe du traitement est nettement "superieure" à celle de l'effet placebo. Nous suspectons donc l'efficacité du traitement.

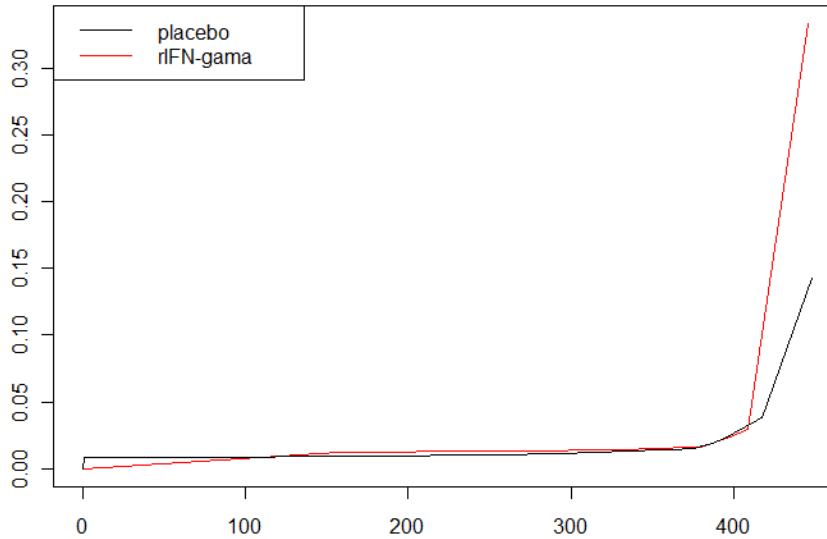


### 3.2 Estimation du taux de hasard cumulé H

Calculons l'estimateur de Nelson-Aalen  $\hat{H}_{NA}(t) = \int_0^t \frac{d\bar{N}_n(s)}{\bar{Y}_n(s)}$  (cf chap.2). On a le graphe suivant :



Pour l'estimateur  $\hat{h}_n(t_i) = \frac{m(t_i)}{r(t_i)}$  on a le graphe suivant :



L'estimateur  $\hat{h}_n$  de  $h$  des deux groupes est une courbe croissante ce qui s'interprète pas un risque croissant de décès par cette maladie.

Nous remarquons que la courbe du risque pour l'effet placebo est "superieur" à celui du traitement.

### 3.3 Estimation de la fonction de densité

#### 3.3.1 Données non censurées

Nous allons estimer les mêmes données en enlevant les cas censurés. L'estimateur à noyau pour les données non censurées est donné par (cf chap.2) :

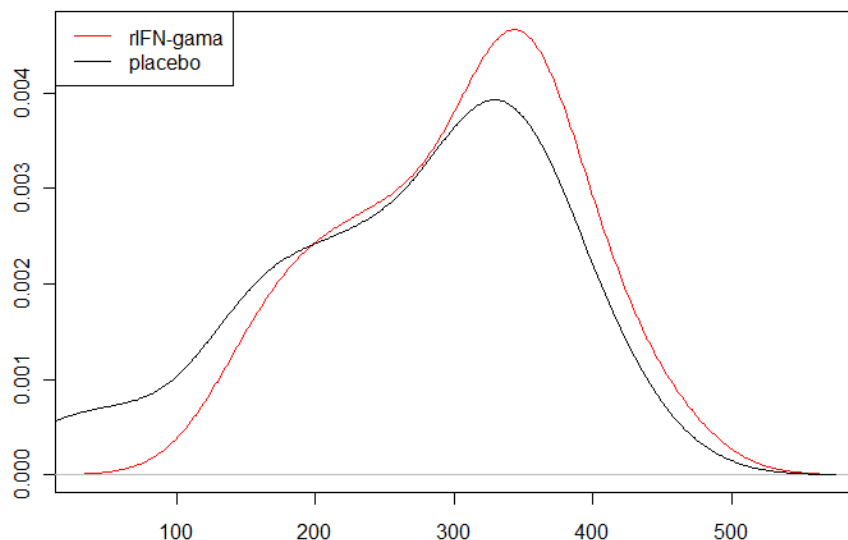
$$f_n(t) = \frac{1}{nb_n} \sum_{i=1}^n K\left(\frac{t - T_i}{b_n}\right)$$

Nous allons faire les calculs pour le noyau gaussien et le noyau d'epanechnikov.

1.Noyau gaussien : il est défini par :

$$\forall u \in \mathbb{R}, K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

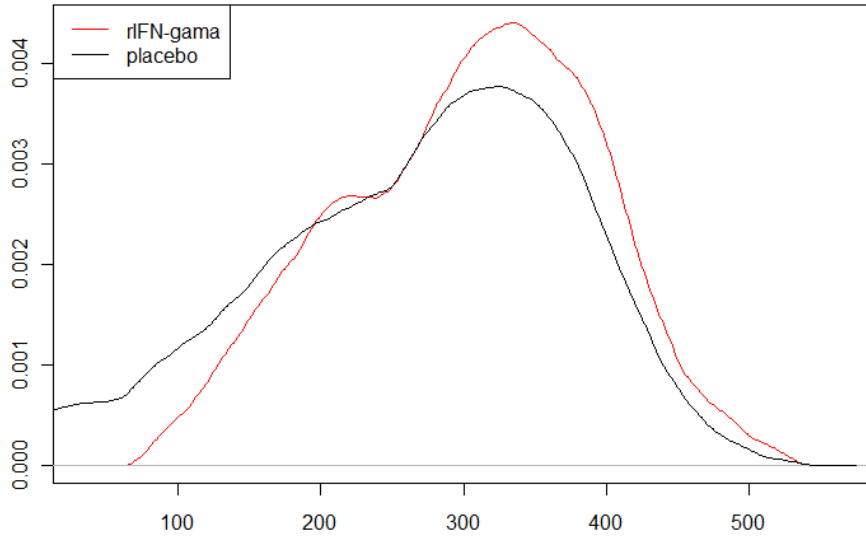
On calcule  $f_n$  avec  $b_n = 40.89$  pour le traitement et  $b_n = 42.36$  pour l'effet placebo. On a le graphe suivant :



2.Noyau d'epanechnikov : il est défini par :

$$\forall u \in \mathbb{R}, K(u) = \frac{3}{4} (1 - u^2) \mathbf{1}_{|u| \leq 1}$$

On calcule  $f_n$  avec  $b_n = 40.89$  pour le traitement et  $b_n = 42.36$  pour l'effet placebo. On a le graphe suivant :

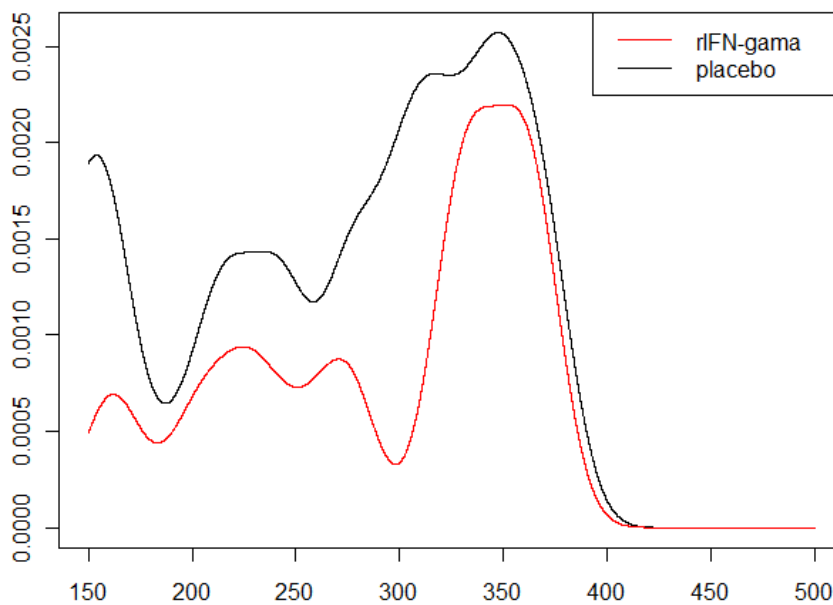


### 3.3.2 Données censurées

L'estimateur à noyau pour les données censurées est donné par (cf chap.2) :

$$f_n(t) = \frac{1}{b_n} \int_0^\infty K\left(\frac{t-s}{b_n}\right) \widehat{S}_{KM}(s^-) d\widehat{H}_{NA}(s)$$

On calcule  $f_n$  avec  $K$  noyau gaussien et  $b_n = 13.28$ . On a le graphe suivant



#### Conclusion :

Nous remarquons que pour chaque estimateur nous pouvons conclure sur l'efficacité du traitement.

# Annexe

**Définition 3.3.1 Mouvement brownien** *Le mouvement brownien standard (ou le processus de Wiener) est un processus stochastique réel  $W = (W_t)_{t \geq 0}$  vérifiant :*

- i)  $W_0 = 0$  P p.s.
- ii)  $\forall s \in [0, t]$  ,  $W_t - W_s \sim \mathcal{N}(0, t - s)$
- iii)  $\forall n \geq 1, \forall t_1 \leq t_2 \leq \dots \leq t_n$   
 $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$  sont indépendants
- iv) P-p.s., l'application  $t \mapsto W_t$  est continue.

**Proposition 3.3.2** *Les 3 propriétés suivantes sont équivalentes :*

- (a)  $\begin{cases} ii) (W_t; t \in \mathbb{R}_+) \text{ est à accroissements indépendants,} \\ iii) \forall 0 \leq s < t, W_t - W_s \text{ suit une loi normale } \mathcal{N}(0, t - s) \end{cases}$
- (b)  $\begin{cases} ii') \forall t \geq 0, W_t \text{ suit une loi normale } \mathcal{N}(0, t), \\ iii') \forall 0 < s \leq t, W_t - W_s \text{ est indépendant de } F_s \text{ sont indépendants} \end{cases}$
- (c)  $\begin{cases} ii'') (W_t)_{t \geq 0} \text{ est un processus gaussien, centré,} \\ iii'') \forall s, t \in \mathbb{R}_+, \text{cov}(W_t, W_s) = E[W_t W_s] = s \wedge t. \end{cases}$

**Théorème 3.3.3** *Soit*

$$Y(t) = \prod_{]0, t]} (1 + dX)$$

et

$$Y'(t) = \prod_{]0, t]} (1 + dX')$$

Alors

$$Y(t) - Y'(t) = \int_{s \in [0, t]} \prod_{]0, s]} (1 + dX) (X(ds) - X'(ds)) \prod_{]s, t]} (1 + dX')$$

Si  $Y'(t)$  est régulier, alors

$$\begin{aligned} \frac{Y(t)}{Y'(t)} - 1 &= \int_{s \in [0, t]} \prod_{]0, s]} (1 + dX) (X(ds) - X'(ds)) \left[ \prod_{]s, t]} (1 + dX') \right]^{-1} \\ &= \int_0^t \frac{Y(t)(s^-)}{Y'(s)} (X(ds) - X'(ds)) \end{aligned}$$

**Corollaire 3.3.4** *Soit  $N = \{N(t), t \geq 0\}$  un processus ponctuel adapté à une filtration  $\{\mathcal{F}_t, t \geq 0\}$ . Alors il existe un unique processus continu à droite croissant  $A$  tel que  $A(0) = 0$  presque sûrement,  $A(t) < \infty$  presque sûrement pour tout  $t$  et  $M = N - A$  est une  $\mathcal{F}$ -martingale locale. De plus,  $A$  est localement borné et  $M$  est localement de carré intégrable.*

# Conclusion

Dans ce mémoire, nous avons étudié la convergence faible dans l'espace  $C[0, 1]$  et  $D[0, 1]$ .

Comme application, nous avons vu le théorème de Rebolledo et les convergence faibles de certains estimateurs en données de survie.

Nous avons vu aussi des simulations pour les estimateurs sur des données réelles.

# Bibliographie

- [1] Billingsley, P. "*Convergence of Probability Measures*" Wiley (1968).
- [2] Bosq, D. et Lecoutre, J. "*Théorie de l'estimation fonctionnelle*". Economica, Paris (1987).
- [3] Fleming, T.R. & D.P. Harrington. "*Counting processes and survival analysis*". Wiley Series in Probability and Statistics. John Wiley & Sons (1991).
- [4] Sahi, N. "*Modèles de survie Estimation et Applications*" Mémoire de magister en probabilités statistiques, (octobre 2011).
- [5] Zhang Biao "*Some asymptotic results for kernel estimation under random censorship*". Bernoulli 2 (2) (1996), p. 183-198.