Analyse de survie et processus de comptage

Philippe Saint Pierre

IMT, Equipe de Statistique et Probabilités

Université Paul Sabatier - Toulouse III

Ecole CIMPA, Lomé, Septembre 2018

INTRODUCTION

On considère une v.a. positive ou nulle X représentant uner durée de vie. On suppose que X est censurée aléatoirement à droite par la variable C,

$$\begin{cases} T_i = X_i \wedge C_i, & \text{où } X_i \text{ et } C_i \text{ sont } i.i.d \text{ et } X \perp C \\ \delta_i = 1\!\!1_{\{X_i \leq C_i\}} & \end{cases}$$

A partir de ces observations, on peut construire deux processus aléatoires :

$$N(t) = \sum_{i=1}^{n} 1\!\!1_{\{T_i \leq t, \delta_i = 1\}},$$

qui compte le nombre de morts au temps t ,

$$Y(t) = \sum_{i=1}^{n} 1\!\!1_{\{T_i \geq t\}},$$

qui compte le nombre d'individus à risque au temps t.

PROCESSUS ALÉATOIRE

Modélisation des occurences dans le temps d'événements aléatoires qui ont lieu en **temps continu** et sont de nature discrète

Exemples de différents types de processus dans un service de chirurgie :

- Espace d'états dénombrable et temps discret : le nombre d'opérations en attente aux temps de la $t^{ième}$ opération.
- Espace d'états dénombrable et temps continu : le nombre d'opérations au cours du temps.
- Espace d'états continu et temps discret : temps d'attente pour la $t^{ième}$ opération.
- Espace d'états continu et temps continu : temps total cumulé de toutes les opérations au cours du temps.

PROCESSUS ALÉATOIRE

DEFINITION

Soit un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$, une **filtration** (ou histoire) $\{\mathscr{F}_t : t \in \mathscr{T}\}$ est une famille croissante et continue à droite de sous-tribus de $\mathscr{A}$.

- Une filtration permet de définir ce qui appartient au passé du processus (qui est observé et connu) et ce qui appartient au futur (pas encore accessible).
- La tribu engendrée par $(X_s, s \leq t)$ est notée $\mathscr{F}_t = \sigma(X_s, s \leq t)$. Dans ce cas, X est automatiquement adapté à la filtration $\sigma(X_s, s \leq t)$.
- On observe également d'autres processus qui peuvent apporter de l'information sur le processus étudié. Dans ce cas, la filtration est élargie pour intégrer cette information (par exemple, prise en compte de covariables).

PROCESSUS ALÉATOIRE

DEFINITION

Soit un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ où $\mathbb{P}$ est la mesure de probabilité sur $(\Omega, \mathscr{A})$. Un **processus aléatoire**, ou encore une **fonction aléatoire réelle** (f.a.r.) est une fonction à deux variables

$$\begin{array}{rcl} X : \mathscr{T} \times \Omega & \longrightarrow & \mathbb{R} \\ (t, \omega) & \longmapsto & X(t, \omega) \end{array}$$

où t représente le temps et ω le hasard. Pour tout $t \geq 0$, la fonction $X_t : \omega \longmapsto X(t, \omega)$ est une variable aléatoire réelle appelée **coordonnée** à l'instant t. Pour tout $\omega \in \Omega$, la **trajectoire** (chemin) est la fonction $t \longmapsto (X(t, \omega), t \in \mathscr{T})$.

PROCESSUS ALÉATOIRE

DEFINITION

Soient un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et une filtration $\mathscr{F}_t$. Le processus X est dit continu à droite (ou à gauche), à variation bornée, croissant (décroissant), possédant des limites à droite (à gauche) si l'ensemble des trajectoires qui satisfont ces propriétés est de probabilité 1.

DEFINITION

Soient un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et $\mathscr{F}_t$ une filtration. Le processus $X = X(t, \omega)$ est dit $\mathscr{F}_t-$**adapté** si $\forall t$, X_t est $\mathscr{F}_t$-mesurable.

DEFINITION

Un processus $(X_t, t \in \mathscr{T})$ est intégrable si $\sup_{t \in \mathscr{T}} \mathbb{E}(|X_t|) < \infty$.

Un processus $(X_t, t \in \mathscr{T})$ est de carré intégrable si $\sup_{t \in \mathscr{T}} \mathbb{E}(X_t^2) < \infty$.

PROCESSUS ALÉATOIRE

DEFINITION

Un processus stochastique X est dit $\mathscr{F}_t$−**prévisible** si comme fonction de $(t,\omega) \in \mathscr{T} \times \Omega \longmapsto \mathbb{R}$ il est mesurable par rapport à la tribu sur $\mathscr{T} \times \Omega$ engendrée par les processus adaptés et continu à gauche.

En quelque sorte, la valeur de X en t est entiérement déterminé à partir des valeurs passées de $X(s)$, $s < t$. Intuitivement, un processus prévisible est un processus dont la valeur en t découle des valeurs observés avant t.

PROPOSITION

Soit X un processus $\mathscr{F}_t$-prévisible alors pour tout t, X_t est $\mathscr{F}_t$−mesurable.

PROPOSITION

Tout processus adapté à $\mathscr{F}_t$ et continu à gauche est $\mathscr{F}_t$−prévisible.

DÉCOMPOSITION DE DOOB-MEYER

THEOREM

Décomposition de Doob-Meyer

Soit X une sous-martingale positive continue à droite et adaptée à la filtration $\mathscr{F}_t$. Il existe une martingale continue à droite M et un processus prévisible croissant continu à droite Λ vérifiant $\mathbb{E}(\Lambda_t) < \infty$ tel que presque sûrement on a

$$X_t = M_t + \Lambda_t \text{ pour tout } t \geq 0.$$

Il y a unicité de la décomposition).

DEFINITION

Le processus prévisible Λ est appelé **compensateur** de X.

DÉCOMPOSITION DE DOOB-MEYER

Le calcul des variances fait intervenir le carré des martingales. Comme le carré d'une martingale est une sous-martingale (inégalité de Jensen), on a

PROPOSITION

Soit M une martingale adaptée à $\mathscr{F}_t$, continue à droite, telle que $\mathbb{E}(M^2(t)) < \infty$ pour tout t. Alors il existe un unique processus continu à droite prévisible noté $< M >=< M, M >$ tel que $< M > (0) = 0$ presque sûrement, $\mathbb{E}(< M > (t)) < \infty$ pour tout t et

$$M^2(t) - < M > (t) \qquad t \geq 0,$$

est une martingale continue à droite. $< M >$ est le compensateur de M^2.

DEFINITION

$< M >=< M, M >$ est le processus de variation prévisible de M.

Intégrale stochastique

Definition

L'intégrale stochastique d'un processus X_t par rapport à un processus B_t est décrite par l'intégrale :

$$\int_0^t X_t dB_t = \lim_{\delta_n \to 0} \sum_{t_i \in \Delta_n} X_{t_i}(B_{t_{i+1}} - B_{t_i}),$$

où $\Delta_n = (t_1, ..., t_n)$ est une partition de $[0, t]$ de pas δ_n. La notation dB_t signifie un changement infinitésimal dans le processus aléatoire qui est une variable aléatoire.

Produit intégral (ou infini)

Definition

Soit $\mathbf{X}(t)$, une matrice $p \times p$ de processus càdlàg, nul en 0 et à variation localement bornée. Le **produit intégral** de X sur $[s,t]$ est

$$\underset{s \leq u \leq t}{\mathscr{P}} (\mathbf{Id} + d\mathbf{X}(\mathbf{u})) = \lim_{\max|t_i - t_{i-1}| \to 0} \prod_{i=1}^{n} \left[\mathbf{Id} + \mathbf{X}(t_i) - \mathbf{X}(t_{i-1})\right]$$

où $t_0 = s < t_1 < \ldots < t_n = t$ est une partition de $[s,t]$ et $\mathbf{Id}$ représente la matrice identité $p \times p$.

Dans le cas scalaire $(p = 1)$, le produit intégral vérifie

$$\underset{s \leq u \leq t}{\mathscr{P}} (1 + dX(u)) = \prod_{s \leq u \leq t} (1 + \Delta X(u)) \exp\left(\int_s^t dX^C(u)\right)$$

où X^C désigne la partie continue de X et $\Delta X(u) = X(u) - X(u-)$.

PRODUIT INTÉGRAL (OU INFINI)

THEOREM

Soient $\mathbf{Z}$ une matrice $p \times p$ de fonctions càdlàg et $\mathbf{Id}$ la matrice identité. L'unique solution $\mathbf{Z}$ de l'équation

$$\mathbf{Z}(t) = \mathbf{Id} + \int_0^t \mathbf{Z}(s-)\mathbf{X}(ds).$$

est $\underset{0 \leq u \leq t}{\mathscr{P}} (\mathbf{Id} + d\mathbf{X}(\mathbf{u})).$

PROCESSUS DE COMPTAGE

DEFINITION

Soit $\{\mathscr{F}_t : t \in \mathscr{T}\}$ une filtration sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Un **processus de comptage** $N(.)$ sur $\mathscr{T}$ est un processus continu à droite avec une limite à gauche, $\mathscr{F}_t$-adaptée, nul en zéro, croissant, à trajectoire constante par morceaux, ayant des sauts d'amplitude 1 et tel que $N(t)$ est p.s. fini pour tout $t \in \mathscr{T}$.

- Un processus de comptage est l'enregistrement des occurrences dans le temps d'événements disjoints et discrets.
- Soit la **filtration naturelle** engendrée par N, $\mathscr{N} = \sigma(N(s), s \leq t)$.
- Un processus de comptage est une sous-martingale (croissant, nul en zéro et borné) $\Longrightarrow$ Décomposition de Doob-Meyer.

PROCESSUS DE COMPTAGE

PROPOSITION

Soit $N = \{N(t), t \geq 0\}$ un processus de comptage adapté à $\mathscr{F}_t$, càdlàg, croissant et localement borné (c'est une sous-martingale locale). Alors il existe un unique processus $\Lambda(.)$ $\mathscr{F}_t$-prévisible, croissant, càdlàg et nul en zéro p.s. vérifiant $\mathbb{E}(\Lambda(t)) < \infty$ p.s. pour tout t, tel que

$$M(t) = N(t) - \Lambda(t), t \in \mathscr{T} \tag{1}$$

soit une martingale locale continue à droite. $\Lambda(.)$ est le compensateur.de N.

- La décomposition permet de diviser l'information initiale (le processus de comptage) en une composante Λ facile à contrôler et et une martingale M qui s'apparente à un résidu.

PROCESSUS DE COMPTAGE

Dans les cas simples, il est possible de deviner la forme du compensateur Λ, et de montrer que $N-\Lambda$ est une martingale. On peut ensuite conclure que Λ est bien le compensateur de N par unicité de la décomposition de Doob-Meyer.

PROPOSITION

Soit $N=\{N(t),t\geq 0\}$ un processus de comptage et soit $\Lambda(.)$ son compensateur tel que $N-\Lambda=M$. On montre que $M^2-\Lambda$ est une martingale continue à droite. Par conséquent, par unicité de la décomposition et la proposition (3) on a

$$<M,M>=\Lambda.$$

PROCESSUS DE COMPTAGE

De nombreuses quantités rencontrées ne s'écrivent pas immédiatement comme des martingales mais sous la forme $\int HdM$, où M est une martingale et H est un processus prévisible.

PROPOSITION

Soit M une martingale locale de carré intégrable et soit H un processus prévisible et localement borné alors

$$U(t) = \int_0^t H(u)dM(u)$$

est une martingale locale de carré intégrable telle que $\mathbb{E}(U(t)) = 0$ et

$$< \int HdM >= \int H^2 d < M > .$$

Cette proposition s'applique notamment à la martingale de décomposition Doob-Meyer d'un processus de comptage.

PROCESSUS DE COMPTAGE

PROPOSITION

Soient N un processus comptage et Λ son compensateur obtenu par la décomposition de Doob-Meyer. On dit que N a une ***intensité*** *λ si λ est prévisible et si*

$$\Lambda(t) = \int_0^t \lambda(u)\,du, \qquad \forall t.$$

Le paramètre λ représente, sachant le passé, la probabilité infinitésimale d'avoir un saut du processus de comptage.

DEFINITION

Un processus de comptage k-dimensionnel $\mathbf{N} = (N_1, N_2, \ldots, N_k)$ est appelé **processus de comptage multivarié** si chacune de ses composantes est un processus de comptage univarié et s'il ne peut y avoir simultanéité des sauts de deux (ou plus) de ses composantes.

VRAISEMBLANCE D'UN PROCESSUS DE COMPTAGE

Soit $N = (N_1, N_2, ..., N_k)$ un processus de comptage multivarié admettant comme compensateur $\Lambda = (\Lambda_1, \Lambda_2, ..., \Lambda_k)$ et comme intensité $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k)$.

PROPOSITION

On définit la vraisemblance partielle

$$L_p = \mathscr{P}_{t\in[0,\tau]} \left(\prod_{h=1}^{k} d\Lambda_h(t)^{\Delta N_h(t)} (1 - d\Lambda_.(t))^{1-\Delta N_.(t)} \right).$$

Si Λ est presque sûrement absolument continue, alors

$$d\mathbb{P} = d\mathbb{P}|_{\mathscr{F}_0} \mathscr{P}_{t\in[0,\tau]} \left(\prod_{h=1}^{k} \lambda_h(t)^{\Delta N_h(t)} \right) \exp\left[-\Lambda_.(\tau)\right],$$

où $\Lambda_. = \sum_{h=1}^{k} \Lambda_h$ et $N_. = \sum_{h=1}^{k} N_h$.

Analyse des durées de vie et processus de comptage

Processus non censuré

On considère $T_1, \dots, T_n$ des v.a. positives indépendantes telles que tous les T_i aient la même fonction de survie S, fonction de répartition F, la densité f et la fonction de risque instantané $\alpha = \frac{f}{S}$. Les variables $T_1, \dots, T_n$ représentent les durées de vie de n individus.

Pour tout $i = 1, \dots, n$, on peut définir

- le processus de comptage $N_i(t) = \mathbf{1}_{\{T_i \leq t\}}$,
- le processus prévisible (càg et adapté) $Y_i(t) = \mathbf{1}_{\{T_i \geq t\}}$,
- la filtration $\mathscr{F}_i = (\mathscr{F}_{it})_{t \geq 0}$ où $\mathscr{F}_{it} = \sigma(N_i(s), s \leq t)$.

ANALYSE DES DURÉES DE VIE ET PROCESSUS DE COMPTAGE

PROPOSITION

Le compensateur du processus de comptage N_i est Λ_i avec

$$\Lambda_i(t) = \int_0^t \alpha(s)Y_i(s)ds \qquad \forall t \geq 0,$$

et N_i admet pour intensité

$$\lambda_i(t) = \alpha(t)Y_i(t) \qquad \forall t \geq 0,$$

*N_i satisfait un modèle d'**intensité multiplicative**.*

En effet, on peut montrer que Λ_i est un processus càdlàg, prévisible, croissant, nul en zéro et continu et que $M_i = N_i - \Lambda_i$ est une martingale locale.

Processus agrégé

Comme on observe plusieurs processus de comptage, on va travailler sur un processus agrégé.

$$N_.(t) = \sum_{i=1}^{n} N_i(t), \quad Y_.(t) = \sum_{i=1}^{n} Y_i(t), \quad \text{et} \quad \mathscr{F}_t = \sigma(N_.(s), s \leq t).$$

Proposition

$N_.$ est un processus de comptage appelé processus de comptage agrégé.

PROCESSUS AGRÉGÉ

PROPOSITION

Le compensateur du processus de comptage $N_.$ est $\Lambda_.$ avec

$$\Lambda_.(t) = \int_0^t \alpha(s) Y_.(s) ds \qquad \forall t \geq 0,$$

et $N_.$ admet pour intensité

$$\lambda(t) = \alpha(t) Y_.(t) \qquad \forall t \geq 0.$$

$N_.$ satisfait un modèle d'intensité multiplicative.

Notons que $M_. = \sum_{i=1}^n M_. = N_. - \Lambda_.$ est une martingale locale et que $\Lambda_. = \sum_{i=1}^n \Lambda_.$ est un processus càdlàg, prévisible, croissant, nul en zéro et continu.

PROCESSUS DE COMPTAGE ET CENSURE À DROITE

PROCESSUS CENSURÉ À DROITE

Considérons n individus et X_i la durée de vie de l'individu i. On suppose que les durées $X_1, \ldots, X_n$ sont $i.i.d$. On se place dans le cas d'une censure à droite aléatoire de type I. Pour $i = 1, \ldots, n$, on observe les variables

$$\begin{cases} T_i = X_i \wedge C_i, & \text{où les } C_i \text{ sont } i.i.d \\ \delta_i = \mathbb{1}_{\{X_i \leq C_i\}} \end{cases}$$

et on suppose que X_i et C_i sont indépendantes.

PROCESSUS DE COMPTAGE ET CENSURE À DROITE

Dans le cas où la durée de vie X_i est censurée à droite par C_i, l'observation des processus de comptage $N_i(t) = \mathbb{1}_{\{X_i \leq t\}}$ et $Y_i(t) = \mathbb{1}_{\{X_i \geq t\}}$ est stoppée au temps aléatoire C_i : N_i est observé uniquement sur l'ensemble aléatoire $E_i = \{t \leq C_i\}$ où quand

$$\mathscr{C}_i\,(t) = \mathbb{1}_{\{t \in E_i\}} = \mathbb{1}_{\{t \leq C_i\}} = 1.$$

$\mathscr{C}_i\,(t)$ est le processus de censure à droite (prévisible) de l'individu i.

PROCESSUS DE COMPTAGE ET CENSURE À DROITE

DEFINITION

Pour $i = 1, \ldots, n$, les processus réellement observés sont

- $N_i^c(.)$ le processus de comptage censuré à droite représentant la partie observable de $N_i(.)$:

$$N_i^c(t) = \int_0^t \mathscr{C}_i(s)\, dN_i(s) = 1\!\!1_{\{X_i \leq t, \delta_i = 1\}} = 1\!\!1_{\{T_i \leq t, \delta_i = 1\}},$$

- $Y_i^c(t)$ tel que,

$$Y_i^c(t) = 1\!\!1_{\{T_i \geq t\}} = 1\!\!1_{\{C_i \geq t\}} 1\!\!1_{\{X_i \geq t\}} = \mathscr{C}_i(t)\, Y_i(t).$$

Remarque : nécessité d'élargir la filtration pour prendre en compte la censure.

$$\mathscr{F}^c = (\mathscr{F}_t^c)_{t \geq 0} \qquad \text{avec } \mathscr{F}_t^c = \sigma\left((N_i^c(s), Y_i^c(s)), i = 1, \ldots, n, s \leq t\right).$$

CENSURE À DROITE INDÉPENDANTE

DEFINITION

Soit $N(.)$ un processus de comptage de compensateur $\Lambda(.)$ par rapport à la filtration $\mathscr{F}_t$. Soit $\mathscr{C}(.)$ un processus de censure à droite prévisible par rapport à $\mathscr{F}_t^c$. La censure à droite générée par $\mathscr{C}(.)$ est **indépendante** si le compensateur de $N(.)$ par rapport à $\mathscr{F}_t^c$ est aussi $\Lambda(.)$.

Autrement dit, la connaissance des temps de censure juste avant t ne modifie pas l'intensité du processus N au temps t. Lorsque la censure est indépendante, la répartition des temps de décès est la même pour les patients censurés et pour les patients non censurés.

CENSURE À DROITE INDÉPENDANTE

Considérons les processus agrégés censurées

$$N_{.}^{c}(t) = \sum_{i=1}^{n} N_{i}^{c}(t) \quad \text{et} \quad Y_{.}^{c}(t) = \sum_{i=1}^{n} Y_{i}^{c}(t).$$

PROPOSITION

Sous l'hypothèse de censure à droite indépendante,
Le compensateur $\Lambda_{.}^{c}(.)$ par rapport à $\mathscr{F}_{t}^{c}$ de $N_{.}^{c}(.)$ est

$$\Lambda_{.}^{c}(t) = \sum_{i=1}^{n} \int_{0}^{t} \mathscr{C}_{i}(s)\, d\Lambda_{i}(s) = \int_{0}^{t} \alpha(s)\, Y_{.}^{c}(s) ds.$$

$N_{.}^{c}(.)$ satisfait un modèle d'intensité multiplicative par rapport à $\mathscr{F}_{t}^{c}$ avec

$$\lambda^{c}(t) = \alpha(t) Y_{.}^{c}(t) \qquad \forall t \geq 0,$$

et $Y_{i}^{c}(t) = \mathscr{C}_{i}(t)\, Y_{i}(t)$.

CENSURE À DROITE INDÉPENDANTE

CENSURE INDÉPENDANTE

Sous hypothèse de censure indépendance

- un processus de comptage censuré à droite satisfait toujours un modèle d'intensité multiplicative.
- la vraisemblance partielle associé à ce processus a la même forme que la vraisemblance du processus non censuré
- Ces résultats impliquent que l'inférence et la théorie asymptotique s'appliquent de la même manière pour les données censurées.

Désormais, nous nous placerons dans le cadre d'un **mécanisme de censure indépendante**. Dans tout ce qui suit, les processus considérés seront des versions censurées mêmes si les notations ne le feront plus apparaître.

ESTIMATEUR DE NELSON-AALEN : NOTATIONS

Soient $X_1, \dots, X_n$ les durées de vie de n individus. On se place dans le cas d'une censure à droite aléatoire de type I, pour $i = 1, \dots, n$ on observe

$$\begin{cases} T_i = X_i \wedge C_i, \\ \delta_i = \mathbb{1}_{\{X_i \leq C_i\}} \end{cases}$$

On fait les hypothèses suivantes,

- les variables X_i et C_i sont indépendantes,
- les variables $X_1, \dots, X_n$ sont $i.i.d$ de fonction de répartition F, de fonction de survie $S = \bar{F}$, de densité f, de fonction de risque instantané $\alpha = \frac{f}{S}$ et de risque cumulé $A = \int \alpha$,
- les variables $C_1, \dots, C_n$ sont $i.i.d$ de fonction de répartition G et de fonction de survie $\bar{G}$,
- on travaille sur $[0, \tau_H]$ où $\tau_H = \tau_F \wedge \tau_G = \sup\{t : \bar{F}(t)\bar{G}(t) > 0\}$, $\tau_F = \sup\{t : \bar{F}(t) > 0\}$ et $\tau_G = \sup\{t : \bar{G}(t) > 0\}$.

ESTIMATEUR DE NELSON-AALEN : NOTATIONS

Sous une hypothèse de censure indépendante, le processus de comptage

$$N(t) = \sum_{i=1}^{n} N_i(t) = \sum_{i=1}^{n} 1\!\!1_{\{T_i \leq t, \delta_i = 1\}}$$

vérifie

$$N(t) = \Lambda(t) + M(t)$$

où $M(.)$ est une martingale locale de carré intégrable et où $\Lambda(.)$ est le compensateur de $N(.)$ tel que

$$\Lambda(t) = \sum_{i=1}^{n} \Lambda_i(t) = \int_0^t \alpha(s) Y(s) ds \quad \text{où} \quad Y(t) = \sum_{i=1}^{n} Y_i(t) = \sum_{i=1}^{n} 1\!\!1_{\{T_i \geq t\}}.$$

De plus, le processus de comptage $N(.)$ a pour intensité multiplicative

$$\lambda(t) = \alpha(t) Y(t) \qquad \forall t \geq 0,$$

avec $Y(t)$ est un processus prévisible et α est le risque instantané.

ESTIMATEUR DE NELSON-AALEN : NOTATIONS

On peut notamment remarquer que $\mathbb{E}(M) = 0$ car, $\forall i = 1, \dots, n$
Démonstration :

$$
\begin{aligned}
\mathbb{E}(N_i(t)) &= \mathbb{P}(T_i \le t, \delta_i = 1) = \mathbb{P}(X_i \le t, X_i \le C_i) = \mathbb{E}(\mathbb{E}\left(\mathbb{1}_{\{X_i \le t\}} \mathbb{1}_{\{X_i \le C_i\}} \mid X_i\right)) \\
&= \mathbb{E}(\mathbb{1}_{\{X_i \le t\}} \bar{G}(X_i-)) = \int_0^t \bar{G}(s-) f(s) ds = \int_0^t \bar{G}(s-) S(s-) \frac{f(s)}{S(s)} ds \\
&= \int_0^t \mathbb{P}(C_i \ge s, X_i \ge s) \frac{f(s)}{S(s)} ds = \int_0^t \mathbb{P}(T_i \ge s) \alpha(s) ds \\
&= \int_0^t \mathbb{E}(Y_i(s)) \alpha(s) ds = \mathbb{E}\left(\int_0^t Y_i(t) \alpha(s) ds\right) = \mathbb{E}\left(\Lambda_i(t)\right).
\end{aligned}
$$

ESTIMATEUR DE NELSON-AALEN

On cherche à estimer le risque cumulé A

$$A(t) = \int_0^t \alpha(s)ds \qquad \forall t \geq 0.$$

En effet, d'après la décomposition de Doob-Meyer

$$M(t) = N(t) - \Lambda(t) = N(t) - \int_0^t \alpha(s)Y(s)ds$$

est une martingale locale de carré intégrable.

DEFINITION

L'estimateur de **Nelson-Aalen** du risque cumulé est

$$\hat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s), \tag{2}$$

où $J(t) = \mathbb{1}_{\{Y(t)>0\}}$. On utilise la convention $J(t)/Y(t) = 0$ si $J(t) = 0$.

ESTIMATEUR DE NELSON-AALEN

REMARQUE

Soient $\begin{pmatrix} T_{(1)} \\ \delta_{(1)} \end{pmatrix} < \cdots < \begin{pmatrix} T_{(n)} \\ \delta_{(n)} \end{pmatrix}$ *les statistiques d'ordre de l'échantillon ordonnées selon les valeurs de* T, *alors*

$$\hat{A}(t) = \sum_{j:T_{(j)} \leq t} \frac{\delta_{(j)}}{Y\left(T_{(j)}\right)} = \sum_{j:T_{(j)} \leq t} \frac{\delta_{(j)}}{n+1-j}.$$

La fonction $\hat{A}$ *est constante par morceaux, croissante, continue à droite, de saut* $1/Y\left(T_{(j)}\right)$ *au temps de saut* $T_{(j)}$ *de* N.

ESTIMATEUR DE NELSON-AALEN

REMARQUE

On peut montrer que l'estimateur de Nelson-Aalen peut être vu comme un estimateur non-paramétrique du maximum de vraisemblance. La vraisemblance peut s'écrire

$$\mathscr{L} = \prod_{t \leq \tau} \left(\prod_{i=1}^{n} \left(Y_i(t) \Delta A(t) \right)^{\Delta N_i(t)} \right) \exp \left(- \sum_{i=1}^{n} \sum_{t \leq \tau} Y_i(t) \Delta A(t) \right).$$

Le maximum de la vraisemblance est obtenu pour

$$\Delta \hat{A}(t) = \frac{\sum_{i=1}^{n} \Delta N_i(t)}{\sum_{i=1}^{n} Y_i(t)} = \frac{\Delta N(t)}{Y(t)}.$$

On en déduit l'estimateur de Nelson-Aalen,

$$\hat{A}(t) = \sum_{s \leq t} \Delta \hat{A}(s) = \sum_{s \leq t} \frac{\Delta N(s)}{Y(s)}.$$

BIAIS ET VARIANCE DE NELSON-AALEN

On définit $A^*(t)$ la limite de $\hat{A}(t)$

$$A^*(t) = \int_0^t J(s)\alpha(s)ds$$

où $J(t) = \mathbb{1}_{\{Y(t)>0\}}$. On a ensuite la proposition suivante

PROPOSITION

On a pour tout $t \geq 0$

$$\begin{aligned} \hat{A}(t) - A^*(t) &= \int_0^t \frac{J(s)}{Y(s)} dN(s) - \int_0^t J(s)\alpha(s)ds \\ &= \int_0^t \frac{J(s)}{Y(s)} \left(dN(s) - \alpha(s)Y(s)ds\right) = \int_0^t \frac{J(s)}{Y(s)} dM(s), \end{aligned}$$

où $\frac{J}{Y}$ est un processus prévisible localement borné (car Y est prévisible) et M une martingale locale de carré intégrable. De plus, on a

Biais et variance de Nelson-Aalen

Proposition

*Le **biais** de $\hat{A}$ est*

$$\mathbb{E}(\hat{A}(t)) - A(t) = -\int_0^t \alpha(s)\mathbb{P}(Y(s)=0)ds$$
$$= -\int_0^t \alpha(s)\left(1-\bar{F}(s^-)\bar{G}(s^-)\right)^n ds.$$

Le biais de $\hat{A}$ est asymptotiquement nul.

On a vu que $\mathbb{E}(\hat{A}(t)) = \mathbb{E}(A^*(t))$. De plus $\mathbb{E}(A^*(t)) = \int_0^t \mathbb{E}(J(s))\alpha(s)ds = \int_0^t \mathbb{P}(Y(s)>0)\alpha(s)ds$, on peut donc écrire

$$\mathbb{E}(\hat{A}(t)) - A(t) = \int_0^t \left(\mathbb{P}(Y(s)>0)-1\right)\alpha(s)ds = -\int_0^t \mathbb{P}(Y(s)=0)\alpha(s)ds.$$

De plus

$$\mathbb{P}(Y(s)=0) = \mathbb{P}(T_i < s\ \forall i=1,\ldots,n) = \mathbb{P}(T < s)^n = (1-\mathbb{P}(T \geq s))^n$$

Biais et variance de Nelson-Aalen

Proposition

$< \hat{A} - A^* > (t) = \int_0^t \frac{J(s)}{Y^2(s)} d < M > (s) = \int_0^t \frac{J(s)}{Y^2(s)} d\Lambda(s) = \int_0^t \frac{J(s)}{Y(s)} \alpha(s) ds.$

Proposition

$\mathbb{E}\left((\hat{A}(t) - A^*(t))^2\right) = \int_0^t \mathbb{E}\left(\frac{J(s)}{Y(s)}\right) \alpha(s) ds.$

Proposition

Un estimateur de la variance de $\hat{A}$ est

$$\hat{\sigma}^2_{\hat{A}}(t) = \int_0^t \frac{J(s)}{[Y(s)]^2} dN(s).$$

Biais et variance de Nelson-Aalen

Démonstration :
En effet, $<\hat{A}-A^*>$ est le compensateur de $(\hat{A}-A^*)^2$. On a donc
$\mathbb{E}\left((\hat{A}(t)-A^*(t))^2\right)=\mathbb{E}\left(<\hat{A}-A^*>(t)\right)=\int_0^t \mathbb{E}\left(\frac{J(s)}{Y(s)}\right)\alpha(s)ds.$

On a montré que $\mathbb{E}(\hat{A}(t))\simeq A(t)$ quand n est grand. De plus, on peut montrer que $\forall t\in[0,\tau_H]$, $A^*(t)-A(t)\underset{n\to\infty}{\longrightarrow}0$ car $\mathbb{P}(J(t)=1)\underset{n\to\infty}{\longrightarrow}1$. Finalement, la variance de $\hat{A}$ vérifie

$$\mathbb{V}(\hat{A}(t))=\mathbb{E}\left((\hat{A}(t)-\mathbb{E}(\hat{A}(t))^2\right)\underset{n\to\infty}{\simeq}\mathbb{E}\left((\hat{A}(t)-A(t))^2\right)\underset{n\to\infty}{\simeq}\mathbb{E}\left((\hat{A}(t)-A^*(t))^2\right)$$
$$=\int_0^t \mathbb{E}\left(\frac{J(s)}{Y(s)}\right)\alpha(s)ds=\int_0^t \mathbb{E}\left(\frac{J(s)}{Y(s)}\right)\frac{1}{Y(s)}d\Lambda(s).$$

Comme $\int_0^t d\Lambda(s)=\int_0^t \alpha(s)Y(s)ds=N(t)-M(s)$ où M est une martingale, on propose l'estimateur suivant

PROPRIÉTÉS ASYMPTOTIQUES DE NELSON-AALEN

THEOREM (CONSISTANCE UNIFORME DE $\hat{A}$)

On a

$$\sup_{s\in[0,t]} \left|\hat{A}(s) - A(s)\right| \xrightarrow{\mathbb{P}} 0,$$

car les conditions suivantes sont vérifiées

$$\int_0^t \frac{J(s)}{Y(s)}\alpha(s)ds \xrightarrow[n\to\infty]{\mathbb{P}} 0 \qquad (1)$$

$$\int_0^t (1 - J(s))\alpha(s)ds \xrightarrow[n\to\infty]{\mathbb{P}} 0 \qquad (2)$$

On peut montrer, en appliquant l'inégalité de Lenglart, que les conditions (1) et (2) impliquent la consistance uniforme.

Propriétés asymptotiques de Nelson-Aalen

Theorem (Convergence faible de $\hat{A}$)

Soit $t \in [0, \tau_H]$, *et* $\sigma^2(s) = \int_0^s \frac{\alpha(u)}{\bar{F}(u)\bar{G}(u)} du.$

$$\forall s \in [0,t], \qquad n \int_0^s \frac{J(u)}{Y(u)} \alpha(u) du \xrightarrow[n\to\infty]{\mathbb{P}} \sigma^2(s) \tag{1}$$

$$\forall s \in [0,t], \qquad \sqrt{n} \int_0^s (1 - J(u)) \alpha(u) du \xrightarrow[n\to\infty]{\mathbb{P}} 0 \tag{2}$$

$$\forall \varepsilon > 0, \qquad n \int_0^s \frac{J(u)}{Y(u)} \alpha(u) \mathbb{1}_{\left\{\sqrt{n}\frac{J(u)}{Y(u)} > \varepsilon\right\}} du \xrightarrow[n\to\infty]{\mathbb{P}} 0 \tag{3}$$

si les conditions précédentes, sont vérifiées, on a

$$\sqrt{n}(\hat{A} - A) \overset{D}{\rightsquigarrow} U$$

où U *est une martingale gaussienne telle que* $U(0) = 0$ *et*
$Cov(U(s), U(t)) = \sigma^2(s \wedge t)$

ESTIMATEUR DE KAPLAN-MEIER

PROPOSITION

Soit $\forall t \geq 0,\ S(t) = \bar{F}(t) = \mathbb{P}(X > t)$ *et* $A(t) = -\int_0^t \frac{d\bar{F}(s)}{\bar{F}(s-)}$, *alors*

$$\bar{F}(t) = \mathop{\mathscr{P}}_{0 \leq s \leq t} (1 - dA(s)), \qquad \forall t \text{ tel que } A(t) < \infty \text{ (i.e. } \forall t \in [0, \tau_F[).$$

DÉMONSTRATION.

On remarque que $\forall t \in [0, \tau_F[,\ \bar{F}(t-) > 0$ et $A(t) < \infty$. En différenciant A,

$$dA(t) = -\frac{d\bar{F}(t)}{\bar{F}(t-)} \Longleftrightarrow \bar{F}(t-)dA(t) = -d\bar{F}(t) \Longrightarrow \int_0^t d\bar{F}(s) = -\int_0^t \bar{F}(s-)dA$$

$$\Longrightarrow \bar{F}(t) = 1 - \int_0^t \bar{F}(s-)dA(s).$$

L'équation précedente est une équation de Volterra où $Z = \bar{F}$ et $X = -A$. Donc $\forall t \in [0, \tau_F[,\ \bar{F}(t-) > 0$, la solution de cette équation est donnée par le produit

ESTIMATEUR DE KAPLAN-MEIER

REMARQUE

Si T est une variable continue, elle admet une densité f. Dans ce cas, $d\bar{F}(s) = -f(s)ds$ et $\bar{F}(s-) = \bar{F}(s)$, on a alors que

$$A(t) = \int_0^t \frac{f(s)}{\bar{F}(s-)} ds = \int_0^t \alpha(s) ds$$

où A est la fonction de risque cumulé. De plus, d'après la définition du produit intégral, on a

$$\bar{F}(t) = \underset{0 \leq s \leq t}{\mathscr{P}} (1 - dA(s)) = \exp\left(-\int_0^t dA^C(s)\right) = \exp\left(-\int_0^t \alpha(s) ds\right).$$

ESTIMATEUR DE KAPLAN-MEIER

On va utiliser l'estimateur de Nelson-Aalen et la proposition précédente pour obtenir l'estimateur de Kaplan-Meier. En effet, d'après la définition du produit intégral, comme $\hat{A}$ est une fonction càdlàg, croissante et à variation localement bornée $\left(\hat{A}(t) - \hat{A}(t-) \leq n\right)$, on peut donc définir

DEFINITION

L'estimateur de **Kaplan-Meier** de la fonction de survie est

$$\hat{S}(t) = \underset{0 \leq s \leq t}{\mathscr{P}} (1 - d\hat{A}(s)), \qquad \forall t \in [0, \tau_F[\tag{3}$$

où $\hat{A}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s)$ est l'estimateur de Nelson-Aalen ; $J(t) = \mathbb{1}_{\{Y(t)>0\}}$. On utilise la convention $J(t)/Y(t) = 0$ si $J(t) = 0$.

ESTIMATEUR DE KAPLAN-MEIER

REMARQUE

Soit $\begin{pmatrix} T_{(1)} \\ \delta_{(1)} \end{pmatrix} < \cdots < \begin{pmatrix} T_{(n)} \\ \delta_{(n)} \end{pmatrix}$ *les statistiques d'ordre de l'échantillon ordonnées selon les valeurs de* T*. L'estimateur* $\hat{A}$ *etant une fonction constante par morceaux, l'estimateur de Kaplan-Meier devient un produit fini*

$$\hat{S}(t) = \mathscr{P}_{0 \leq s \leq t}(1 - \Delta\hat{A}(s)) = \prod_{0 \leq s \leq t}\left(1 - \frac{\Delta N(s)}{Y(s)}\right)$$

$$= \prod_{j:T_{(j)} \leq t}\left(1 - \frac{\Delta N(T_{(j)})}{Y(T_{(j)})}\right) = \prod_{j:T_{(j)} \leq t}\left(1 - \frac{\delta_{(j)}}{Y\left(T_{(j)}\right)}\right) = \prod_{j:T_{(j)} \leq t}\left(1 - \frac{\delta_{(j)}}{n+1-j}\right),$$

Estimateur de Kaplan-Meier

Remarque

On montre que l'estimateur de Kaplan-Meier est un estimateur non-paramétrique du maximum de vraisemblance. La modélisation des durées de vie peut s'exprimer en fonction des deux paramètres suivants qui sont liés par les relations suivantes

$$A(t) = -\int_0^t \frac{dS(s)}{S(s-)},$$

$$S(t) = \mathop{\mathscr{P}}_{0\le s\le t}(1 - dA(s)).$$

Comme $\hat{A}$ est un estimateur non paramétrique du maximum de vraisemblance de A, la propriété d'invariance des estimateurs du maximum de vraisemblance permet de conclure que $\hat{S}$ est également un estimateur non paramétrique du maximum de vraisemblance de S.

Biais et variance de Kaplan-Meier

On introduit la quantité

$$\bar{F}^*(t) = S^*(t) = \underset{0\le s\le t}{\mathscr{P}}(1 - dA^*(s)), \qquad \forall t \ge 0,$$

où $A^*(s) = \int_0^s J(u)\alpha(u)du$.

Proposition

On a pour tout $t \ge 0$

$$\begin{aligned} \frac{\hat{S}(t)}{S^*(t)} - 1 &= -\int_0^t \frac{\hat{S}(s-)}{S^*(s)} d(\hat{A} - A^*)(s) \\ &= -\int_0^t \frac{\hat{S}(s-)}{S^*(s)} \frac{J(s)}{Y(s)} dM(s), \end{aligned}$$

où M *est une martingale locale de carré intégrable telle que pour tout* $s \ge 0$, $M(t) = N(t) - \int_0^t \alpha(s)Y(s)ds$. *De plus,* $\frac{\hat{S}(t)}{S^*(t)} - 1$ *est une martingale locale de carré intégrable sur* $[0, \tau_F[$ *telle que*

BIAIS ET VARIANCE DE KAPLAN-MEIER

PROPOSITION

*Le **biais** de $\hat{S}$ vérifie les inégalités suivantes*

$$0 \leq \mathbb{E}(\hat{S}(t)) - S(t) \leq F(t)\left(1 - S(t)\bar{G}(t)\right)^n \leq F(t)\exp(-\mathbb{E}(Y(t))).$$

La dernière inégalité est obtenu en remarquant que

$$\mathbb{E}(Y(t)) = n\mathbb{P}(T_i \geq t) = nS(t)\bar{G}(t)$$

et que $\log(1-x) \leq -x$.
Le biais de l'estimateur de Kaplan-Meier est faible quand $\mathbb{E}(Y(t))$ est grande , $i.e.$ au niveau des premières observations.

BIAIS ET VARIANCE DE KAPLAN-MEIER

On cherche à obtenir un estimateur de la variance de $\hat{S}$. Comme $\frac{\hat{S}(.)}{S^*(.)} - 1$ est une martingale locale de carré intégrable, on peut considèrer la décomposition suivante

$$\left(\frac{\hat{S}(t)}{S^*(t)} - 1\right)^2 = \left\langle \frac{\hat{S}(t)}{S^*(t)} - 1 \right\rangle + \xi_t,$$

où ξ_t est une martingale locale de carré intégrable et

$$\mathbb{E}\left(\left(\frac{\hat{S}(t)}{S^*(t)} - 1\right)^2\right) = \mathbb{E}\left(\left\langle \frac{\hat{S}(t)}{S^*(t)} - 1 \right\rangle\right).$$

De plus,

$$\mathbb{V}\left(\frac{\hat{S}(t)}{S^*(t)} - 1\right) = \mathbb{E}\left(\left(\frac{\hat{S}(t)}{S^*(t)} - 1\right)^2\right) = \mathbb{V}\left(\frac{\hat{S}(t)}{S^*(t)}\right)$$

.

BIAIS ET VARIANCE DE KAPLAN-MEIER

De plus on a,

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{\hat{S}(t)}{S^*(t)}-1\right)^2\right) &= \mathbb{E}\left(\left\langle -\int_0^t \frac{\hat{S}(s-)}{S^*(s)}\frac{J(s)}{Y(s)}dM(s)\right\rangle\right) \\
&= \mathbb{E}\left(\int_0^t \left(\frac{\hat{S}(s-)J(s)}{S^*(s)Y(s)}\right)^2 d < M > (s)\right) \\
&= \mathbb{E}\left(\int_0^t \left(\frac{\hat{S}(s-)}{S^*(s)Y(s)}\right)^2 J(s)Y(s)\alpha(s)ds\right).
\end{aligned}$$

Biais et variance de Kaplan-Meier

En remplaçant $\hat{S}(s-)$ par $\hat{S}(s)$ et $\alpha(s)Y(s)ds$ par $dN(s)$ on montre que

$$\mathbb{V}\left(\hat{S}(t)\right) \simeq (S(t))^2 \int_0^t \frac{J(s)}{(Y(s))^2} dN(s).$$

Proposition

Un estimateur de la variance de $\hat{S}$ est

$$\hat{\sigma}^2_{\hat{S}}(t) = \left(\hat{S}(t)\right)^2 \int_0^t \frac{J(s)}{(Y(s))^2} dN(s).$$

Propriétés asymptotiques de Kaplan-Meier

Theorem (Consistance uniforme de $\hat{S}$)

Soit $t \in [0, \tau_H]$, *on a*

$$\sup_{s\in[0,t]} \left|\hat{S}(s) - S(s)\right| \xrightarrow{\mathbb{P}} 0,$$

car les conditions suivantes sont vérifiées

$$\int_0^t \frac{J(s)}{Y(s)}\alpha(s)ds \xrightarrow[n\to\infty]{\mathbb{P}} 0 \qquad (1)$$

$$\int_0^t (1 - J(s))\alpha(s)ds \xrightarrow[n\to\infty]{\mathbb{P}} 0 \qquad (2)$$

On peut montrer, en appliquant l'inégalité de Lenglart, que les conditions (1) et (2) impliquent la consistance uniforme.

PROPRIÉTÉS ASYMPTOTIQUES DE KAPLAN-MEIER

THEOREM (CONVERGENCE FAIBLE DE $\hat{S}$)

Soit $t \in [0, \tau_H]$, *on pose* $\sigma^2(s) = \int_0^s \frac{\alpha(u)}{S(u)\bar{G}(u)} du$. *Comme les conditions suivantes, sont vérifiées*

$$\forall s \in [0,t], \qquad n \int_0^s \frac{J(u)}{Y(u)} \alpha(u) du \xrightarrow[n\to\infty]{\mathbb{P}} \sigma^2(s) \qquad (1)$$

$$\forall s \in [0,t], \qquad \sqrt{n} \int_0^s (1 - J(u)) \alpha(u) du \xrightarrow[n\to\infty]{\mathbb{P}} 0 \qquad (2)$$

$$\forall \varepsilon > 0, \qquad n \int_0^s \frac{J(u)}{Y(u)} \alpha(u) 1\!\!1_{\left\{\sqrt{n}\frac{J(u)}{Y(u)} > \varepsilon\right\}} du \xrightarrow[n\to\infty]{\mathbb{P}} 0 \quad (3)$$

on a

$$\sqrt{n}(\hat{S} - S) \overset{D}{\rightsquigarrow} -SU$$

où U *est une martingale gaussienne telle que* $U(0) = 0$ *et* $Cov(U(s), U(t)) = \sigma^2(s \wedge t)$.

ESTIMATION SEMI-PARAMÉTRIQUE

Modèle de régression permettant d'ajuster l'intensité sur des covariables.
On introduit pour $i \in \{1, \ldots, n\}$,

- $\mathbf{Z}_i = (Z_{1i}, \ldots, Z_{pi})$, le vecteur de covariables de dimension p,
- $\beta = (\beta_1, \ldots, \beta_p)$, le vecteur de dimension p des coefficients de régression,

Pour tout $i = 1, \ldots, n$, l'intensité λ_i du processus N_i est

$$\lambda_i(t) = Y_i(t)\,\alpha_i(t),$$

où α_i est le rique instantané spécifique à l'individu i.

ESTIMATION SEMI-PARAMÉTRIQUE

Pour prendre en compte les covariables $\mathbf{Z}_i$ on suppose que les riques instantanés α_i suivent un modèle semi-paramétrique à risque multiplicatif de "Cox", c'est-à-dire,

$$\alpha_i(t \mid \mathbf{Z}_i) = \alpha_0(t) \exp(\beta^T \mathbf{Z}_i), \qquad i = 1, \ldots, n.$$

où $\alpha_0(t)$ est le risque de base. Notons que le modèle fait des hypothèses de risques proportionnels et de log-linéarité.

ESTIMATION SEMI PARAMÉTRIQUE : RISQUE CUMULÉ

La vraisemblance partielle peut s'écrire

$$\mathscr{L}(\beta) = \prod_{t \leq \tau} \prod_{i=1}^{n} \left(\Delta A_0(t) Y_i(t) \exp(\beta^T \mathbf{Z}_i)\right)^{\Delta N_i(t)} \exp\left[-\sum_{i=1}^{n} \sum_{t \leq \tau} Y_i(t) \exp(\beta^T \mathbf{Z}_i) \Delta A_0(t)\right].$$

où $A_0(t) = \int_0^t \alpha_0(s) ds$. En considérant β fixé, la maximisation de la vraisemblance par rapport à $\Delta A_0(.)$ conduit à

$$\Delta \hat{A}_0(t) = \frac{\Delta N(t)}{\sum_{i=1}^{n} Y_i(t) \exp(\beta^T \mathbf{Z}_i)}.$$

PROPOSITION

*Pour β fixé, $A_0(t) = \int_0^t \alpha_0(s) ds$ est estimé par l'**estimateur de Breslow**,*

$$\hat{A}_0(t) = \int_0^t \frac{J(s)}{\sum_{i=1}^{n} Y_i(s) \exp(\beta^T \mathbf{Z}_i)} dN(s), \tag{4}$$

avec $J(u) = 1_{\{Y(t) > 0\}}$.

ESTIMATION SEMI PARAMÉTRIQUE : COEFFICIENTS DE RÉGRESSION

En remplaçant, $A_0(t)$ par son estimateur, la vraisemblance partielle devient,

$$\mathscr{L}(\beta) = \prod_{t\leq\tau}\prod_{i=1}^{n}\left(\frac{Y_i(t)\exp(\beta^T\mathbf{Z}_i)}{\sum_{j=1}^{n} Y_j(t)\exp(\beta^T\mathbf{Z}_j)}\right)^{\Delta N_i(t)} \prod_{t\leq\tau}\prod_{i=1}^{n}(J(t)dN(t))^{\Delta N_i(t)}\exp\left[-\int_0^\tau J(s)dN(s)\right]$$
$$= \mathscr{L}_{Cox}(\beta)\times\prod_{t\leq\tau}\prod_{i=1}^{n}(J(t)dN(t))^{\Delta N_i(t)}\exp\left[-\int_0^\tau J(s)dN(s)\right].$$

où par définition, $\mathscr{L}_{Cox}(\beta)$ est **la vraisemblance partielle de Cox,**

$$\mathscr{L}_{Cox}(\beta) = \prod_{t\leq\tau}\prod_{i=1}^{n}\left(\frac{Y_i(t)\exp(\beta^T\mathbf{Z}_i)}{\sum_{j=1}^{n} Y_j(t)\exp(\beta^T\mathbf{Z}_j)}\right)^{\Delta N_i(t)}. \quad (5)$$

Vraisemblance introduite par Cox (1972) par une approche différente.
La vraisemblance partielle de Cox n'est pas une vraisemblance dans le sens statistique du terme, mais il est établi qu'elle peut être utilisée comme telle pour estimer les coefficients de régression.

ESTIMATION SEMI PARAMÉTRIQUE : COEFFICIENTS DE RÉGRESSION

La fonction de log-vraisemblance partielle de Cox est,

$$\log \mathscr{L}_{Cox}(\beta) = \sum_i \int_0^\tau \left[\beta^T \mathbf{Z}_i - \log \left(\sum_{j=1}^n Y_j(t) \exp(\beta^T \mathbf{Z}_j) \right) \right] dN_i(t) + cste.$$

Le vecteur score (dérivées de la Log-Vraisemblance par rapport à β) est

$$U(\beta) = \frac{\partial \log \mathscr{L}_{Cox}(\beta)}{\partial \beta} = \sum_i \int_0^\tau \left[\mathbf{Z}_i - \frac{\sum_{j=1}^n Y_j(t) \mathbf{Z}_j \exp(\beta^T \mathbf{Z}_j)}{\sum_{j=1}^n Y_j(t) \exp(\beta^T \mathbf{Z}_j)} \right] dN_i(t),$$

PROPOSITION

L'estimateur $\hat{\beta}$ du maximum de la vraisemblance partielle de Cox vérifie

$$U(\hat{\beta}) = 0.$$

ESTIMATION SEMI PARAMÉTRIQUE : COEFFICIENTS DE RÉGRESSION

En pratique, les estimations des coefficients de régression sont obtenues par maximisation de la log-vraisemblance à l'aide d'algorithmes itératifs, comme par exemple l'algorithme de quasi-Newton.

REMARQUE

Pour simplifier les écritures, on considère des covariables indépendantes du temps néanmoins le modèle de Cox permet de prendre en compte des covariables dépendante du temps. En effet, les raisonnements précédents peuvent facilement être adaptés à des covariables dépendantes du temps.

ESTIMATION SEMI PARAMÉTRIQUE : COEFFICIENTS DE RÉGRESSION

REMARQUE

Les vraisemblances successivement utilisées sont résumées par le schéma suivant :

Vraisemblance partielle
(relative à un processus non censuré)
$$L_p(\beta)$$
$$\downarrow$$
Vraisemblance partielle
(relative à un processus censuré)
$$\mathscr{L}^c(\beta)$$
$$\downarrow$$
Vraisemblance partielle de Cox
$$\mathscr{L}_{Cox}(\beta)$$

Estimation semi paramétrique : survie

Soit $\mathbf{Z}_0$, la valeur des covariables pour un individu.

Proposition

Un estimateur de la fonction de survie est donné par le produit intégral

$$\hat{S}(t \mid \mathbf{Z}_0) = \underset{s \in]0,t]}{\mathscr{P}} \left(1 - d\hat{A}(s \mid \mathbf{Z}_0)\right), \qquad \forall t \in [0, \tau[,$$

avec

$$\hat{A}(t \mid \mathbf{Z}_0) = \int_0^t \frac{J(s) \times \exp(\beta^T \mathbf{Z}_0)}{\sum_{i=1}^n \exp(\beta^T \mathbf{Z}_i) Y_i(s)} dN(s).$$

Propriétés asymptotiques : estimation semi paramétrique

Proposition

L'estimateur $\hat{\beta}$ du maximum de la vraisemblance partielle de Cox vérifie

$$\sqrt{n}\,(\hat{\beta}-\beta_0) \xrightarrow{\mathscr{L}} N(0, nI^{-1}(\hat{\beta})),$$

où β_0 est le vrai paramètre et $I(\beta)$ est la matrice d'information de Fisher :

$$I(\beta) = -\mathbb{E}\left(\frac{\partial^2 \log \mathscr{L}_{Cox}(\beta)}{\partial \beta^2}\right) = -\left\{\mathbb{E}\left(\frac{\partial^2 \log \mathscr{L}_{Cox}(\beta)}{\partial \beta_i \partial \beta_j}\right)\right\}_{p\times p}$$
$$= -\mathbb{E}\left(\int_0^\tau \frac{\left[\sum_{i=1}^n Y_i(t)\mathbf{Z}_i\mathbf{Z}_i^T \exp(\beta^T\mathbf{Z}_i)\right] - \left[\sum_{i=1}^n Y_i(t)\mathbf{Z}_i \exp(\beta^T\mathbf{Z}_i)\right]\left[\sum_{i=1}^n Y_i(t)\mathbf{Z}_i \exp(\beta^T\mathbf{Z}_i)\right]^T}{\left[\sum_{i=1}^n Y_i(t)\exp(\beta^T\mathbf{Z}_i)\right]^2} dN(t)\right)$$

PROPRIÉTÉS ASYMPTOTIQUES : ESTIMATION SEMI PARAMÉTRIQUE

- Pour démontrer les propriétés asymptotiques, on utilise le fait que $M_i(t) = N_i(t) - \int_0^t Y_i(s)\alpha_o(s)\exp(\beta^T \mathbf{Z}_i)ds$ est une martingale par rapport à $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$ avec $\mathscr{F}_t = \sigma\left((N_i(s), Y_i(s), \mathbf{Z}_i), i = 1, \ldots, n, s \leq t\right)$.

- L'inverse de la matrice d'information de Fisher étant égale à la matrice de variance-covariance, elle fournit une estimation de la variance de $\hat{\beta}$ en considérant la matrice d'information observée au point $\beta = \hat{\beta}$, $i.e.$ $\mathscr{I}(\hat{\beta}) = -\left[\frac{\partial^2 \log \mathscr{L}_{Cox}(\beta)}{\partial \beta^2}\right]_{\beta=\hat{\beta}}$.

- On peut ensuite obtenir un estimateur de la variance de $\hat{S}$ et les propriétés asymptotiques de $\hat{S}$ à partir de celles de $\hat{\beta}$ et $\hat{A}_0(t)$ et de l'utilisation de la delta-méthode.