

Construction et utilisation de lois d'expérience : approches avancées

Les approches multi-états

Quentin Guibert

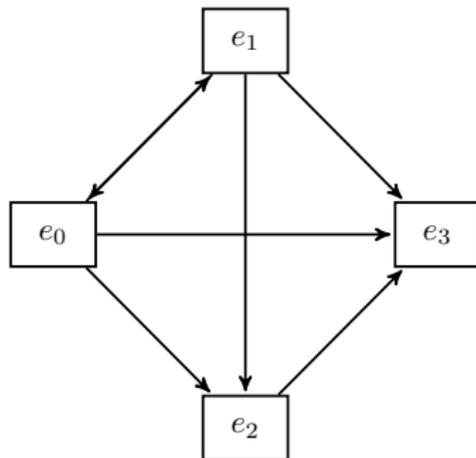
Institut de Science Financière et d'Assurances - UCB Lyon 1
Laboratoire de recherche de Sciences Actuarielle et Financière

Modèles de durée - Applications actuarielles
ISFA - Lyon - 12 Décembre 2014



Assurance de personnes et modèle multi-états

En assurance de personnes, le suivi du risque d'un assuré se traduit par une succession d'évènements ouvrant droit au versement de prestations ou de primes.



Par exemple, en arrêt de travail, 4 états permettent de définir la situation de l'assuré :

- e_0 , l'état sain de cotisant ;
- e_1 , l'état incapable ;
- e_2 , l'état invalide ;
- e_3 , l'état décès.

Assurance de personnes et modèle multi-états

Chaque changement d'état donne lieu à des flux financiers différents soit pour l'assureur, soit pour l'assuré.

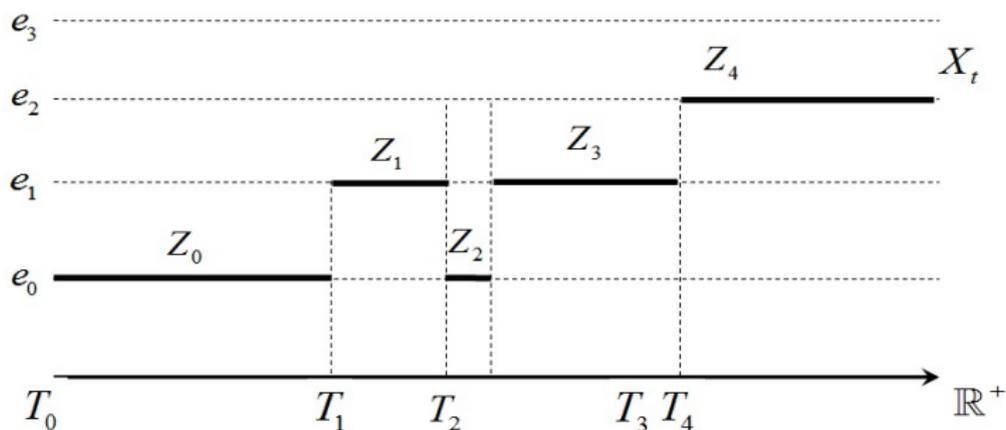


Figure: Exemple de "chemin" parcouru au cours de la durée de vie d'un assuré.

Type de données observées et méthodes d'estimation

- Dans un contexte d'assurance, les données observées sont généralement des **données transversales** ou des **données longitudinales**.
- Avec les premières, il est naturel de se tourner d'avantage vers des **méthodes de régularisation** (CMIR12, 1991; Denuit and Robert, 2007).
- Le second type de données permet une analyse plus fine et la construction de lois plus fidèles au phénomène observé. Toutefois, ces données sont généralement incomplètes :
 - **Censure à droite et troncature à gauche** ;
 - Censure par intervalle.
- Le cas des observations continus avec censure par intervalle est d'avantage rencontré en biostatistique (Commenges, 2002) et ne sera pas traité ici.

Motivation

- Dans la littérature, ces modèles sont classiquement utilisés dans les sciences actuarielles en arrêt de travail, pour la mortalité accidentel et en dépendance (Haberman and Pitacco, 1998; Denuit and Robert, 2007; Christiansen, 2012).
- Les praticiens ne mettent pas nécessairement en œuvre ces techniques et adoptent des méthodes de construction de tables dans le prolongement des celles utilisées pour des données de survie. Il en résulte un traitement marginale de chaque lois de transition.
- Solvabilité 2 requiert de mettre en place des techniques de construction de lois biométriques réalistes. Si elles sont reconstruites, l'interdépendance entre chaque état devrait être considérée dans la phase de construction.

Objectif

Décrire un cadre de travail pour l'estimation de lois d'expérience permettant de tenir compte de ces interactions. Celui-ci s'appuie sur l'utilisation de modèles multi-états qui permettent d'avoir une représentation détaillée de chaque loi de transition.

Notations et hypothèses

- Le processus $(X_t)_{t \geq 0}$ est caractérisé à partir des probabilités de transition entre chaque état et d'une loi initiale. En assurance, on distingue généralement **deux cas**.
- Cas markovien** : l'évolution du processus $(X_t)_{t \geq 0}$ dépend uniquement de l'information au temps présent $\mathcal{F}_s = \sigma(X_s)$. On a alors pour $j \in \mathcal{S}$

$$\mathbb{P}(X_t = j \mid \mathcal{F}_s) = \mathbb{P}(X_t = j \mid X_s) := p_{X_s, j}(s, t).$$

- Cas semi-markovien** : le processus $(X_t, U_t)_{t \geq 0}$ est markovien (Helwich, 2008). On écrit alors $i, j \in \mathcal{S}$

$$p_{ij}(s, t, u, v) := \mathbb{P}(X_t = j, U_t \leq v \mid X_s = i, U_s = u).$$

- On utilise généralement des notations matricielles $P(s, t) = (p_{ij}(s, t))_{i, j \in \mathcal{S}}$ et $P(s, t, u, v) = (p_{ij}(s, t, u, v))_{i, j \in \mathcal{S}}$.
- Dans ce contexte, les probabilités de survie s'écrivent (sans le terme $U_s = u$ dans le cas markovien) pour $i \in \mathcal{S}$

$$1 - \bar{p}_{ii}(s, t, u) := \mathbb{P}(T_{N(s)+1} > t \mid X_s = i, U_s = u).$$

Notations et hypothèses

- De même que pour un modèle de durée, on définit usuellement des fonctions d'intensités de transition.
- Cas markovien** : les intensités de transition sont définies par

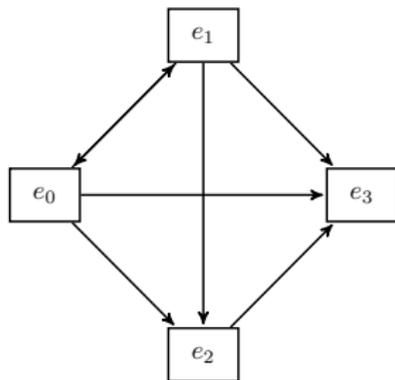
$$\mu_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{p_{ij}(t, t + \Delta t)}{\Delta t} \text{ pour } i \neq j \text{ et } \mu_{ii}(t) = - \sum_{i \neq j} \mu_{ij}(t).$$

- Cas semi-markovien** : de manière équivalente, on a

$$\mu_{ij}(t, u) = \lim_{\Delta t \rightarrow 0} \frac{p_{ij}(t, t + \Delta t, u, \infty)}{\Delta t} \text{ pour } i \neq j \text{ et } \mu_{ii}(t, u) = - \sum_{i \neq j} \mu_{ij}(t, u).$$

- Dans le cas semi-markovien, il est également possible de travailler à partir du noyau semi-markovien (Janssen and Manca, 2006).
- Lorsque les intensités de transition sont constantes en fonction du temps t , on parle de modèle **homogène**.

Exemple 2 : Incapacité-Invalidité (Semi-Markov)



Le calcul de la provision technique (e.g. Christiansen, 2012) nécessite de considérer l'engagement produit par l'entrée dans un nouvel état :

- $V_0(t, u)$, provisions techniques en e_0 ;
- $V_1(t, u)$, provisions techniques en e_1 ;
- $V_2(t, u)$, provisions techniques en e_2 .

On note $b_i(t)$ les flux nets versés dans l'état i .

Provisions techniques

L'engagement vu de e_1 vaut en t (expression similaire pour les autres états)

$$V_1(t, u) = \int_t^\infty \left(b_1(\tau) + \mu_{10}(\tau, \tau - t + u) V_0(\tau, 0) + \mu_{12}(\tau, \tau - t + u) V_2(\tau, 0) + \mu_{13}(\tau, \tau - t + u) b_3(\tau) \right) d\tau$$

Lien entre probabilités et intensités de transition

- Dans la pratique, les actuaires utilisent d'avantage des tables de probabilités de transition.
- **Cas markovien** : les probabilités de transition sont simples à manipuler grâce à la relation de **Chapman-Kolmogorov**

$$P(s, t) = P(s, \tau) P(\tau, t) , s \leq \tau \leq t.$$

- **Cas semi-markovien** : la relation de **Chapman-Kolmogorov** existe mais elle est cette fois beaucoup plus délicate à manipuler (Helwich, 2008)

$$p_{ij}(s, t, u, v) = \sum_{k \in \mathcal{S}} \int_0^{\infty} p_{kj}(r, t, \tau, v) p_{ik}(s, r, u, d\tau) , s \leq r \leq t.$$

Lien entre probabilités et intensités de transition

- Toutefois, il existe un **lien** entre intensités et probabilités de transition qu'il est utile de considérer pour l'inférence. Ce lien découle des équations de Chapman-Kolmogorov.
- **Cas markovien** : Les probabilités de transition sont solutions des équations de forward et backward de Kolmogorov (matricielle)

$$\frac{\partial}{\partial t} P(s, t) = P(s, t) \mu(t) \text{ et } \frac{\partial}{\partial s} P(s, t) = P(s, t) \mu(s) \text{ avec } P(s, s) = \mathbf{Id}.$$

- L'équation forward admet une unique solution s'exprimant sous la forme d'un produit intégral

$$\begin{aligned} P(s, t) &= \lim_{\max_{i=1, \dots, n} |\tau_i - \tau_{i-1}| \rightarrow 0} \prod_{k=1}^n (\mathbf{Id} + \Lambda(\tau_k) - \Lambda(\tau_{k-1})) \\ &= \mathcal{P}_{\tau \in]s, t]} (\mathbf{Id} + d\Lambda(\tau)), \end{aligned}$$

avec $0 < \tau_1 < \dots < \tau_n$ une partition de $]s, t]$ et $\Lambda(t) = (\int_0^t \mu_{ij}(\tau) d\tau)_{i,j \in \mathcal{S}}$ la matrice des fonctions **d'intensités cumulées**.

Estimation non-paramétrique - Cas général

Estimateur d'Aalen-Johansen (AJ)

- L'estimateur d'AJ des **probabilités de transition** s'appuie sur la relation entre probabilités et intensités de transition i.e.

$$\widehat{P}(s, t) = \prod_{l: T_l \leq t} (\mathbf{Id} + \Delta \widehat{\Lambda}(\tau_l)),$$

avec $s < T_1 < \dots < T_l \leq t$ les temps de saut situés entre s et t .

- Avec 2 états, on retrouve l'estimateur de Kaplan-Meier

$$\widehat{S}(t) = \prod_{s: t \leq t} \left(1 - \frac{\Delta N(s)}{Y(s)} \right).$$

- Le covariance et donc la **variance** de l'estimateur AJ prennent la forme

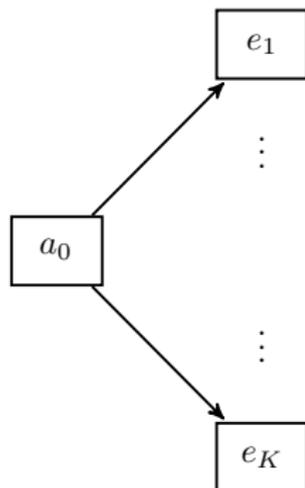
$$\begin{aligned} & \widehat{Cov}(\widehat{p}_{ij}(s, t), \widehat{p}_{lm}(s, t)) \\ &= \sum_{k \in \mathcal{S}} \sum_{h \neq k} \int_s^t \widehat{p}_{ih}(s, \tau) \widehat{p}_{lh}(s, \tau) (\widehat{p}_{kj}(\tau, t) - \widehat{p}_{hj}(\tau, t)) \\ & \quad \times (\widehat{p}_{km}(\tau, t) - \widehat{p}_{hm}(\tau, t)) \frac{\mathbb{1}_{\{Y_i(\tau) > 0\}}}{Y_i^2(\tau)} dN_{ij}(\tau). \end{aligned}$$

Estimation non-paramétrique - Cas général

Estimateur d'Aalen-Johansen

- Il existe également un estimateur de type Greenwood pour la variance.
- L'estimateur d'AJ est biaisé positivement, consistant et **asymptotiquement normal**.
- La manière la plus simple d'obtenir des intervalles de confiance asymptotiques est de procéder par bootstrap non paramétrique.

Application 1 : Modèle à risques concurrents



- Dans le cas d'un modèle à risques concurrents, on distingue plusieurs causes de sortie d'un état initial (e.g. mortalité par cause, entrée en dépendance, rachats, etc.)
- Une littérature extensive est disponible pour l'inférence de ce type de modèle notamment :
 - approches basées sur les durées de survie marginales (Prentice *et al.*, 1978),
 - approches de type "graphic-copula" (Carrière, 1994),
 - modèles à hasard proportionnel (Fine and Gray, 1999).
- Un exemple pour **l'incidence en dépendance** est disponible dans Guibert and Planchet (2014a).

Application 1 : Modèle à risques concurrents

- Les praticiens s'appuient généralement sur les **durées marginales** associées à chaque cause. Ces durées T_{01}, \dots, T_{0K} sont **latentes** et la durée de vie s'écrit

$$T = \min_{j=1, \dots, K} (T_{0j}) \text{ avec } \mathbb{P}(T_{0j} > t) = \exp\left(-\int_0^t \alpha_j(\tau) d\tau\right).$$

- Les intensités des durées latentes $\alpha_j(t)$ sont équivalentes aux intensités de transition $\mu_{0k}(t)$ **si ces durées sont indépendantes**.
- Cette hypothèse est forte mais relativement reprise en assurance. Elle peut conduire à des interprétations hasardeuses (Andersen and Keiding, 2012).

Application 1 : Modèle à risques concurrents

- Sous l'hypothèse d'indépendance, les taux de d'incidence

$$q_j^*(t) = \mathbb{P}(T_{0j} \leq t + 1 \mid T_{0j} > t),$$

peuvent être estimés à partir de l'estimateur de Kaplan-Meier.

- Les praticiens corrigent généralement les taux de d'incidence estimés marginalement en introduisant un structure de dépendance arbitraire.
- Algorithme d'ajustement** pour un ordre quelconque (j_1, \dots, j_n) est :

- $\check{q}_{j_1}(t) = \hat{q}_{j_1}^*(t)$
- $\check{q}_{j_2}(t) = \hat{q}_{j_2}^*(t) (1 - \hat{q}_{j_1}^*(t))$,
- ...
- $\check{q}_{j_n}(t) = \hat{q}_{j_n}^*(t) \prod_{k=1}^{n-1} (1 - \hat{q}_{j_k}^*(t))$.

- Quelque soit l'ordre, l'estimateur ajusté $\check{q}_j(t)$ est compris entre 2 bornes fixes

$$b_j^-(t) = \hat{q}_j^*(t) \prod_{k \neq j} (1 - \hat{q}_k^*(t)) \leq \check{q}_j(t) \leq \hat{q}_j^*(t) = b_j^+(t).$$

Application 1 : Modèle à risques concurrents

- Sous l'hypothèse d'indépendance, les taux de d'incidence

$$q_j^*(t) = \mathbb{P}(T_{0j} \leq t + 1 \mid T_{0j} > t),$$

peuvent être estimés à partir de l'estimateur de Kaplan-Meier.

- Les praticiens corrigent généralement les taux de d'incidence estimés marginalement en introduisant un structure de dépendance arbitraire.
- **Algorithme d'ajustement** pour un ordre quelconque (j_1, \dots, j_n) est :

- $\check{q}_{j_1}(t) = \hat{q}_{j_1}^*(t)$
- $\check{q}_{j_2}(t) = \hat{q}_{j_2}^*(t) (1 - \hat{q}_{j_1}^*(t))$,
- ...
- $\check{q}_{j_n}(t) = \hat{q}_{j_n}^*(t) \prod_{k=1}^{n-1} (1 - \hat{q}_{j_k}^*(t))$.

- Quelque soit l'ordre, l'estimateur ajusté $\check{q}_j(t)$ est compris entre 2 bornes fixes

$$b_j^-(t) = \hat{q}_j^*(t) \prod_{k \neq j} (1 - \hat{q}_k^*(t)) \leq \check{q}_j(t) \leq \hat{q}_j^*(t) = b_j^+(t).$$

Application 1 : Modèle à risques concurrents

- Sous l'hypothèse d'indépendance, les taux de d'incidence

$$\hat{q}_j^*(t) = \mathbb{P}(T_{0j} \leq t + 1 \mid T_{0j} > t),$$

peuvent être estimés à partir de l'estimateur de Kaplan-Meier.

- Les praticiens corrigent généralement les taux de d'incidence estimés marginalement en introduisant un structure de dépendance arbitraire.
- Algorithme d'ajustement** pour un ordre quelconque (j_1, \dots, j_n) est :

- $\check{q}_{j_1}(t) = \hat{q}_{j_1}^*(t)$
- $\check{q}_{j_2}(t) = \hat{q}_{j_2}^*(t) \left(1 - \hat{q}_{j_1}^*(t)\right),$
- ...
- $\check{q}_{j_n}(t) = \hat{q}_{j_n}^*(t) \prod_{k=1}^{n-1} \left(1 - \hat{q}_{j_k}^*(t)\right).$

- Quelque soit l'ordre, l'estimateur ajusté $\check{q}_j(t)$ est compris entre 2 bornes fixes

$$b_j^-(t) = \hat{q}_j^*(t) \prod_{k \neq j} (1 - \hat{q}_k^*(t)) \leq \check{q}_j(t) \leq \hat{q}_j^*(t) = b_j^+(t).$$

Application 1 : Modèle à risques concurrents

- Sous l'hypothèse d'indépendance, les taux de d'incidence

$$q_j^*(t) = \mathbb{P}(T_{0j} \leq t + 1 \mid T_{0j} > t),$$

peuvent être estimés à partir de l'estimateur de Kaplan-Meier.

- Les praticiens corrigent généralement les taux de d'incidence estimés marginalement en introduisant un structure de dépendance arbitraire.
- Algorithme d'ajustement** pour un ordre quelconque (j_1, \dots, j_n) est :

- $\check{q}_{j_1}(t) = \hat{q}_{j_1}^*(t)$
- $\check{q}_{j_2}(t) = \hat{q}_{j_2}^*(t) (1 - \hat{q}_{j_1}^*(t))$,
- ...
- $\check{q}_{j_n}(t) = \hat{q}_{j_n}^*(t) \prod_{k=1}^{n-1} (1 - \hat{q}_{j_k}^*(t))$.

- Quelque soit l'ordre, l'estimateur ajusté $\check{q}_j(t)$ est compris entre 2 bornes fixes

$$b_j^-(t) = \hat{q}_j^*(t) \prod_{k \neq j} (1 - \hat{q}_k^*(t)) \leq \check{q}_j(t) \leq \hat{q}_j^*(t) = b_j^+(t).$$

Application 1 : Modèle à risques concurrents

- Sous l'hypothèse d'indépendance, les taux de d'incidence

$$q_j^*(t) = \mathbb{P}(T_{0j} \leq t + 1 \mid T_{0j} > t),$$

peuvent être estimés à partir de l'estimateur de Kaplan-Meier.

- Les praticiens corrigent généralement les taux de d'incidence estimés marginalement en introduisant un structure de dépendance arbitraire.
- Algorithme d'ajustement** pour un ordre quelconque (j_1, \dots, j_n) est :

- $\check{q}_{j_1}(t) = \hat{q}_{j_1}^*(t)$
- $\check{q}_{j_2}(t) = \hat{q}_{j_2}^*(t) (1 - \hat{q}_{j_1}^*(t))$,
- ...
- $\check{q}_{j_n}(t) = \hat{q}_{j_n}^*(t) \prod_{k=1}^{n-1} (1 - \hat{q}_{j_k}^*(t))$.

- Quelque soit l'ordre, l'estimateur ajusté $\check{q}_j(t)$ est compris entre 2 bornes fixes

$$b_j^-(t) = \hat{q}_j^*(t) \prod_{k \neq j} (1 - \hat{q}_k^*(t)) \leq \check{q}_j(t) \leq \hat{q}_j^*(t) = b_j^+(t).$$

Application 1 : Modèle à risques concurrents

- Application réalisée **sous R sur des données simulées** avec un modèle à risques concurrents avec 3 causes de sortie.
- De nombreux packages sous R sont dédiés à l'inférence de modèles à risques compétitifs et/ou multi-états (Beyersmann *et al.*, 2011; Putter *et al.*, 2007). Nous utilisons ici le package **mstate** (De Wreede *et al.*, 2011).

Data:

	id	from	to	trans	Tstart	Tstop	time	status	clus	seg
1	1	a	i_1	1	0	10.13740	10.13740	0	1	1
2	1	a	i_2	2	0	10.13740	10.13740	1	1	1
3	1	a	d	3	0	10.13740	10.13740	0	1	1
5	2	a	i_1	1	0	21.05648	21.05648	0	1	1
6	2	a	i_2	2	0	21.05648	21.05648	1	1	1
7	2	a	d	3	0	21.05648	21.05648	0	1	1

Application 1 : Modèle à risques concurrents

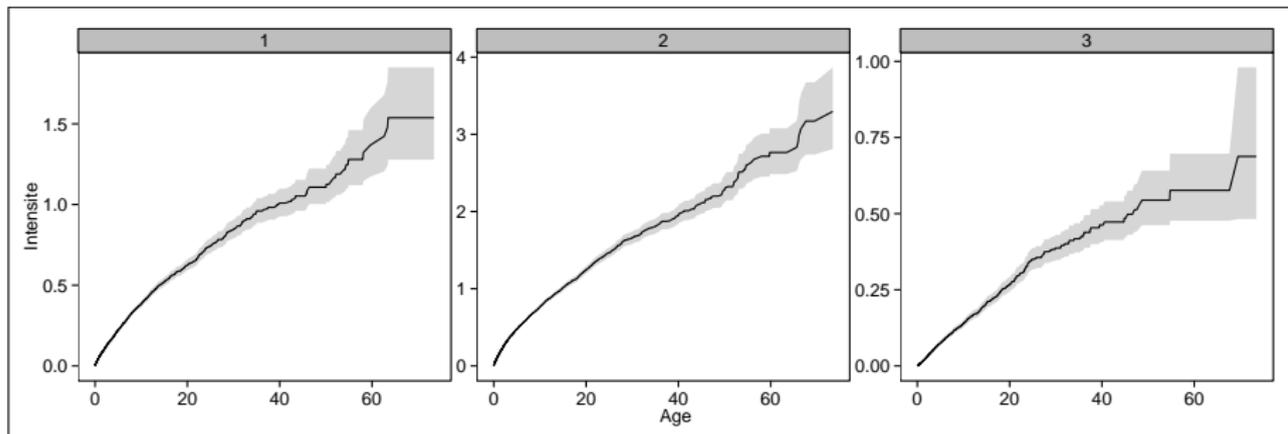


Figure: Estimation des intensités cumulées de transition et intervalles de confiance à 95%.

```
> library(mstate)
> c=coxph(Surv(Tstart, Tstop, status)~strata(trans), data=data.2,
method = "breslow")
> msf=msfit(object = c, vartype = "aalen", trans = tmat, variance=T)
```

Application 1 : Modèle à risques concurrents

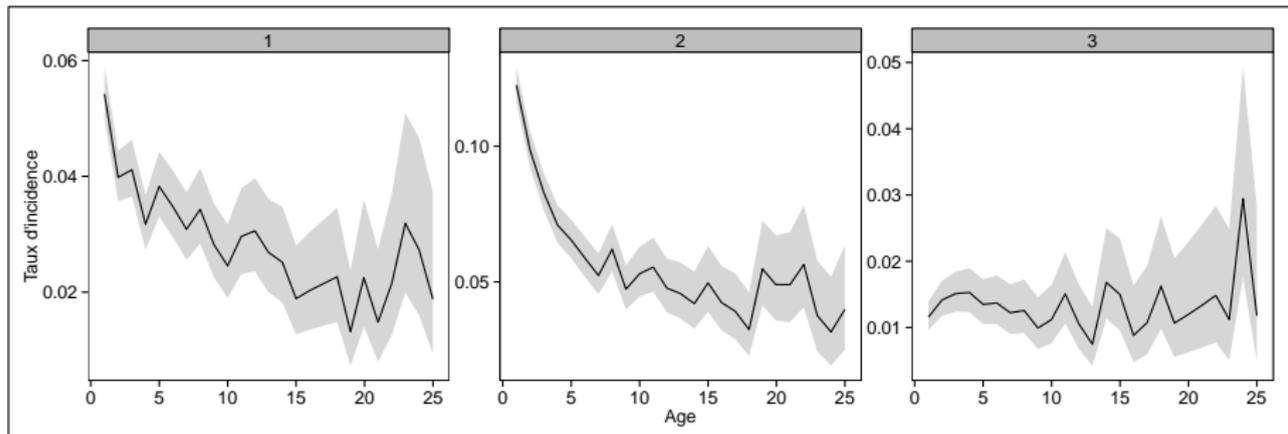


Figure: Estimation des taux d'incidence $q_{0k}(t) = p_{0k}(t, t+1)$ via l'estimateur d'AJ et intervalles de confiance à 95%.

```
> p=probtrans2(msf, predt=0, pas=pas, method="aalen", variance=T,
direction="ratecurve")
```


Modèle à risques concurrents - Cas réel pour l'incidence en dépendance lourde

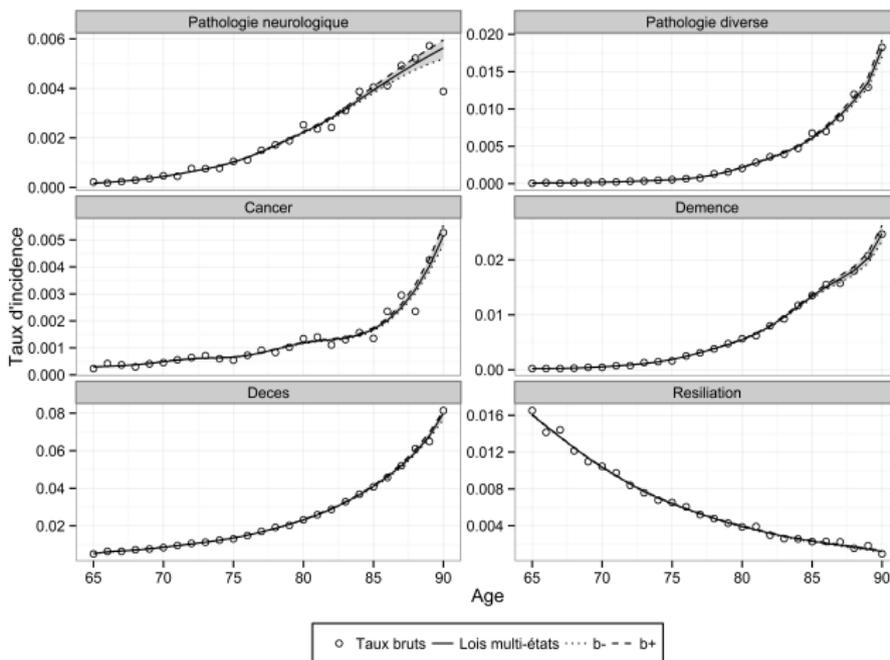


Figure: Comparaison des deux approches après lissage par Whittaker-Henderson.

Estimation paramétrique

- Dans le cadre markovien avec censure à droite, l'utilisation d'un modèle paramétrique simple (Weibull, Gompertz-Makeham, ...) peut apparaître trop restrictive.
- Ainsi, le cadre le plus souvent repris consiste à supposer que les intensités de transition sont **constantes par morceaux**.
- Les approches paramétriques sont d'avantage utiliser dans la cas de modèle à censure par intervalle.

Estimation paramétrique

- L'estimation d'un paramètre θ tel que $\mu_{ij}(t) = \mu_{ij}(t, \theta)$ se fait par **maximum de vraisemblance**.
- La **vraisemblance partielle individuelle** (i.e. associée au cas avec censure à droite) du modèle s'écrit

$$\mathcal{L}(\theta) = \left(\prod_{t \in \mathcal{T}} \prod_{i \neq j} (\mu_{ij}(t, \theta) Y_i(t))^{\Delta N_{ij}(t)} \right) \exp \left(- \int_{t \in \mathcal{T}} \sum_{i \neq j} \mu_{ij}(t, \theta) Y_i(t) dt \right)$$

- Cette expression se réécrit plus simplement en considérant les temps de saut individuels.

Estimation paramétrique - Intensités constantes par morceaux

- Les modèles à intensités constantes par morceaux sont été très utilisés dans la pratique (modèle de Poisson)

$$\mu_{ij}(t) = \mu_{ij,m}, \quad t_{ij,m-1} < t \leq t_{ij,m}.$$

- L'estimateur du maximum de vraisemblance est facilement obtenue et est asymptotiquement gaussien

$$\sqrt{n} \{ \hat{\mu}_{ij,m} - \mu_{ij,m} \} \xrightarrow{d} \mathcal{N}(0, \sigma_{ij,m}),$$

$$\hat{\mu}_{ij,m} = \frac{N_{ij}(t_{ij,m}) - N_{ij}(t_{ij,m-1})}{\sum_{k=1}^n \int_{t_{ij,m-1}}^{t_{ij,m}} Y_i^k(t) dt} \quad \text{et} \quad \sigma_{ij,m} = \frac{\sqrt{N_{ij}(t_{ij,m}) - N_{ij}(t_{ij,m-1})}}{\sum_{k=1}^n \int_{t_{ij,m-1}}^{t_{ij,m}} Y_i^k(t) dt}.$$

- Les estimateurs des probabilités de transition sont alors obtenus par des méthodes numériques. Leur variance asymptotique est obtenue par **méthode Delta**.
- Une phase de **régularisation** est ensuite nécessaire (e.g. GLM).

Estimation paramétrique

- De nombreux outils statistiques permettent de valider les paramètres estimés. On retrouve par exemple les tests classiques de **Wald, du score ou du rapport de vraisemblance** pour tester l'hypothèse $\mathcal{H}_0 : \theta = \theta_0$ s'appuyant sur une statistique χ^2 .
- Dans le cadre du **modèle de Poisson**, le score $U_{ij,m}$ et l'information observée $I_{ij,m}$, utiles à ces tests, sont données par

$$U_{ij,m} = \frac{N_{ij}(t_{ij,m}) - N_{ij}(t_{ij,m-1})}{\mu_{ij,m}} - \sum_{k=1}^n \int_{t_{ij,m-1}}^{t_{ij,m}} Y_i^k(t) dt,$$

$$I_{ij,m} = \frac{N_{ij}(t_{ij,m}) - N_{ij}(t_{ij,m-1})}{\mu_{ij,m}^2}.$$

Prise en compte de l'hétérogénéité

- Pour prendre en compte une **source d'hétérogénéité**, les volumes de données sont souvent insuffisants pour permettre des estimations indépendantes (stratification).
- Les sources d'hétérogénéité **explicites** sont alors prise en compte à partir de modèles régressifs et les sources **implicites** à partir de modèle de fragilité.
- Dans ces conditions, les différents sous-groupes étudiés sont positionnés par rapport à une population de référence.
- Dans le cadre des modèles de durée, ce sont essentiellement des modèles de type **Cox (multiplicatif)** ou **Aalen (additif)** qui sont utilisés du fait de leur facilité d'utilisation et de leurs qualités naturelles.
- Ces approches s'intègrent au cadre multi-états.

Prise en compte de l'hétérogénéité - Modèle de Cox

- Le modèle semi-paramétrique de Cox (Cox, 1972) constitue le modèle le plus utilisé en pratique

$$\mu_{ij}(t | \mathbf{Z}_{ij}) = \mu_{ij,0}(t) \exp(\boldsymbol{\beta}_{ij}^\top \mathbf{Z}_{ij}).$$

- L'estimation des paramètres s'appuie comme pour les modèles de durée classiques sur la **vraisemblance partielle de Cox**

$$\mathcal{L}_{Cox}(\boldsymbol{\beta}) = \prod_{i \in \mathcal{T}} \prod_{i \neq j} \prod_{k=1}^n \left(\frac{Y_i^k(t) \exp(\boldsymbol{\beta}_{ij}^\top \mathbf{Z}_{ij}^k)}{\sum_{i=1}^n \exp(\boldsymbol{\beta}_{ij}^\top \mathbf{Z}_{ij}^k) Y_i^k(t)} \right)^{\Delta N_{ij}^{(k)}(t)}.$$

- Les intensités cumulées de base $\Lambda_{ij,0|z}(t) = \int_0^t \mu_{ij,0}(\tau) d\tau$ sont obtenues à partir de **l'estimateur de Breslow**

$$\hat{\Lambda}_{ij,0|z}(t) = \int_0^t \frac{\mathbb{1}_{\{Y_i^k(\tau) > 0\}}}{\sum_{i=1}^n \exp(\boldsymbol{\beta}_{ij}^\top \mathbf{Z}_{ij}^k) Y_i^k(\tau)} dN_{ij}(\tau).$$

Prise en compte de l'hétérogénéité - Modèle de Cox

- L'estimateur du maximum de vraisemblance $\hat{\beta}$ est **asymptotiquement gaussien** et s'obtient à l'aide d'une méthode numérique

$$\sqrt{n} \{ \hat{\beta} - \beta \} \xrightarrow{d} \mathcal{N} \left(0, nI^{-1} \left(\hat{\beta} \right) \right),$$

$$I \left(\hat{\beta} \right) = - \frac{\partial^2 \log \left(\mathcal{L}_{Cox} \left(\beta \right) \right)}{\partial \beta^2}.$$

- Remarque** : une approche plus parcimonieuse consiste à considérer un vecteur de coefficients $\beta = \beta_{ij}, \forall i \neq j$, relativement long, applicable à toutes les transitions de manière à avoir

$$\mu_{ij} (t | \mathbf{Z}_{ij}) = \mu_{ij,0} (t) \exp \left(\beta^\top \mathbf{Z}_{ij} \right).$$

Prise en compte de l'hétérogénéité - Modèle de Cox

- Les **probabilités de transition** individuelles sont enfin déduites à partir de l'estimateur d'AJ appliqué aux estimateurs des intensités cumulés et en prenant en compte les covariables.

$$\hat{P}(s, t | \mathbf{Z}) = \prod_{l: T_l \leq t} \left(\mathbf{Id} + \Delta \hat{\Lambda}(T_l | \mathbf{Z}) \right),$$

$$\Delta \hat{\Lambda}_{ij}(t | \mathbf{Z}_{ij}) = \Delta \hat{\Lambda}_{ij,0}(t | \mathbf{Z}_{ij}) \exp \left(\beta_{ij}^\top \mathbf{Z}_{ij} \right), \quad i \neq j,$$

et $s < T_1 < \dots < T_l \leq t$ les temps de saut situés entre s et t .

- Cet estimateur est **asymptotiquement normal**.

Prise en compte de l'hétérogénéité - Modèle de Cox

- La significativité des coefficients est estimée à partir des **tests classiques de Wald, du score ou du rapport de vraisemblance**.
- L'hypothèse de proportionnalité des risques est quant à elle testée de manière classique à partir des **résidus de Shoenfeld** (Martinussen and Scheike, 2006).
- Par ailleurs, si la durée est prise en compte via un modèle à hasard proportionnel dépendant de l'âge, l'ensemble des résultats vus ci-avant restent valables (hormis bien sur l'hypothèse de proportionnalité avec le temps).

Application 2 : Modèle à risques concurrents

- On considère cette fois l'impact de la variable "seg" au sein de l'échantillon simulé dans l'application 1.
- Le package `mstate` permet d'estimer les intensités de transition et les probabilités de transition relatives à chaque sous-population.

```
> c=estimcox(data.2,c("seg"),expand=T,time_scale="clock.forward")
```

```
n= 30000, number of events= 8190
```

	coef	exp(coef)	se(coef)	z	Pr(> z)	
seg.1	0.15081	1.16278	0.04111	3.668	0.000244	***
seg.2	0.37406	1.45363	0.02876	13.008	< 2e-16	***
seg.3	-0.09807	0.90659	0.06855	-1.431	0.152538	

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> cox.zph(c)
```

	rho	chisq	p
seg.1	0.00586	0.282	0.596
seg.2	-0.01351	1.491	0.222
seg.3	-0.01167	1.109	0.292
GLOBAL	NA	2.882	0.410

Application 2 : Modèle à risques concurrents

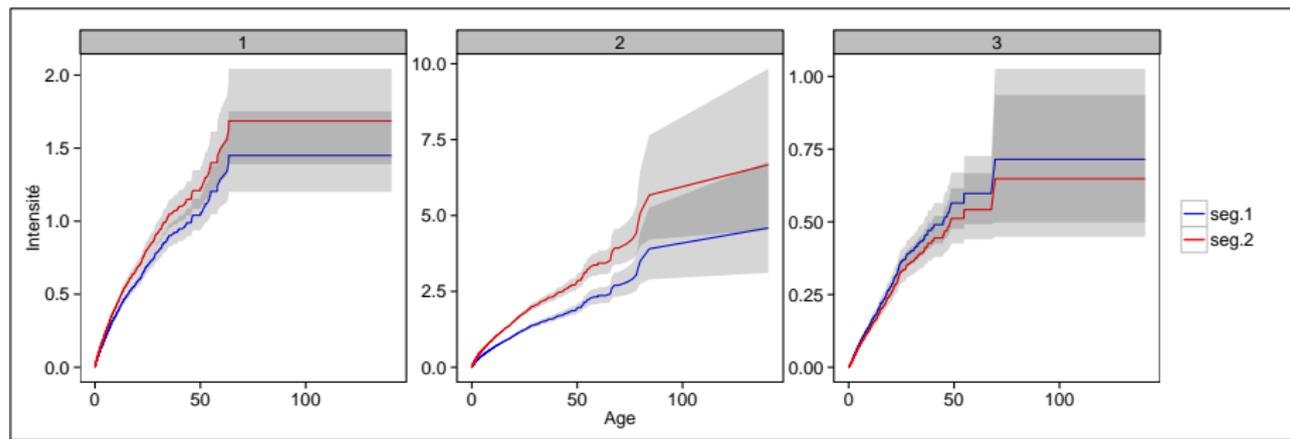


Figure: Estimation des intensités cumulées de transition et intervalles de confiance à 95% avec distinction H/F

```
> msf.1=estimintens (data.2,tmat,c("seg"),c("1"),time_scale="clock.forward",
variance=T)
> msf.2=estimintens (data.2,tmat,c("seg"),c("2"),time_scale="clock.forward",
variance=T)
```


Approches pragmatiques

- Dans la pratique, la problématique induite par le non-respect de l'hypothèse de Markov est contournée en traitant l'une des dimensions temporelles comme une source d'hétérogénéité.
- **Exemple 1** : estimation des lois de maintien en incapacité via l'estimateur de Kaplan-Meier sur la durée de séjour en distinguant les différents âges à l'entrée.
- **Exemple 2** : le modèle semi-markovien de Cox consiste à choisir une variable temporelle comme référence et considérer l'autre comme une covariable (e.g. Andersen and Perme, 2008). On a par exemple

$$\mu_{ij}(t, u) = \mu_{ij}(t | u) = \mu_{ij,0}(t) \exp(\beta f(u)),$$

avec f une fonction connue de la durée de séjour.

- Il est possible de considérer une fonction f non spécifiée (splines et B-splines) afin de prendre un compte un effet non linéaire. Ces techniques sont non spécifiques aux modèles multi-états mais on les retrouve fréquemment avec des données censurées par intervalle (Commenges *et al.*, 2007).

Estimation paramétrique

- La procédure d'estimation standard s'effectue par maximum de vraisemblance.
- La **vraisemblance individuelle** du modèle s'écrit plus simplement en considérant les dates de sauts $0 < T_1 < \dots < T_l$ sur \mathcal{T}

$$\mathcal{L}(\theta) = \prod_{m=1}^l \prod_{k=1}^n \prod_{i \neq j} \left(\mu_{ij} \left(T_m, U_{T_m}^k, \theta \right) \right)^{\Delta N_{ij}^k(T_m)} \\ \times \exp \left(-Y_i^k(T_m) \int_{T_{m-1}}^{T_m} \mu_{ij} \left(\tau, U_{\tau}^k, \theta \right) d\tau \right)$$

- Dans le **cas non-homogène**, il est délicat de donner une forme a priori à la fonction $\mu_{ij}(t, u)$ à l'inverse de cas homogène où des lois classiques (e.g. Weibull) sont retenues.
- L'utilisation de modèle à intensités **constantes par morceaux** est courante (CMIR12, 1991) dans la cas homogène mais moins fréquent dans le cas non-homogène (Monteiro *et al.*, 2006). Comme vu ci-dessus, les estimateurs obtenus correspondent au ratio "nombre des transitions" sur "exposition". Des techniques de **lissage** en deux dimensions sont alors nécessaires (Tomas and Planchet, 2013).

Estimation non-paramétrique

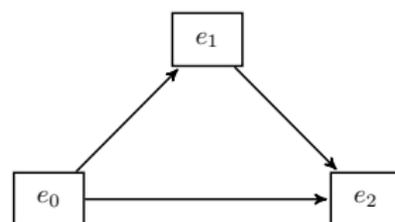
- En présence de censure, il n'existe pas de cadre général pour l'estimation non-paramétrique quand l'hypothèse de Markov n'est pas vérifiée.
- Les modèles développés en biostatistique permettent d'estimer certaines quantités particulières (Datta and Satten, 2002) qui n'intéressent pas vraiment les actuaires.
- Il existe un cadre non-paramétrique pour l'estimation des modèles pour l'incapacité avec 3 états (Meira-Machado *et al.*, 2006).
- Les approches dites à **risques semi-concurrents** sont développées en biostatistique pour estimer la durée de survie latente d'entrée dans un état "malade". Des approches ont été développées dans ce contexte (Peng and Fine, 2006) mais les quantités estimées ne sont pas utiles aux actuaires.

L'approche de non-paramétrique de Meira-Machado *et al.* (2006)

- On introduit 2 durées de vie S , la durée de vie dans l'état e_0 et T , la durée de vie globale.
- On observe les durées censurées par le processus C

$$\begin{cases} Y = \min(S, C) \text{ and } \gamma = \mathbb{1}_{\{S \leq C\}}, \\ Z = \min(T, C) \text{ and } \delta = \mathbb{1}_{\{T \leq C\}}. \end{cases}$$

- Les auteurs proposent des estimateurs des probabilités de transition utilisant des propriétés des **intégrales Kaplan-Meier**.



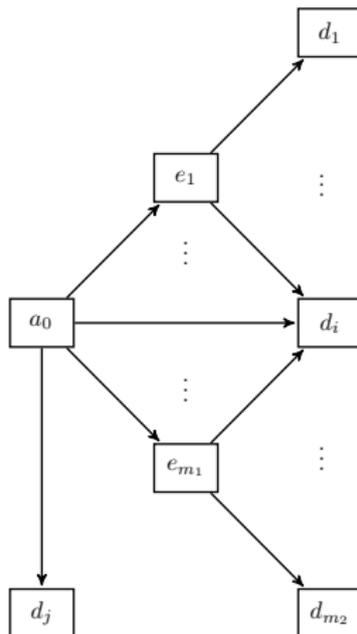
$$p_{1,1}(s, t) = \frac{\mathbb{P}(S > t)}{\mathbb{P}(S > s)},$$

$$p_{1,2}(s, t) = \frac{\mathbb{P}(s < S \leq t < T)}{\mathbb{P}(S > s)} = \frac{\mathbb{E}[\varphi_{st}^{(1)}(S, T)]}{\mathbb{P}(S > s)},$$

$$p_{2,2}(s, t) = \frac{\mathbb{P}(S \leq s, t < T)}{\mathbb{P}(S \leq s < T)} = \frac{\mathbb{E}[\varphi_{st}^{(2)}(S, T)]}{\mathbb{E}[\varphi_{ss}^{(2)}(S, T)]}.$$

Estimation à partir d'intégrales d'Aalen-Johansen

Guibert and Planchet (2014b) étendent cette approche à d'autres probabilités et une classe de modèles un peu plus large.



Estimation à partir d'intégrales d'Aalen-Johansen

- On considère intégrale de la forme

$$S^{(v)}(\varphi) = \int \varphi dF_0^{(v)},$$

avec φ une fonction quelconque, $F_0^{(v)}(t) = \mathbb{P}(S \leq s, T \leq t, V = v)$ et V le type de sortie dans un modèle à risques concurrents.

- S est considéré comme une covariable pour construire l'estimateur.

Estimateur pour un modèle à risque concurrents

$$\widehat{S}_n^{(v)}(\varphi) = \int \varphi(s, t) \widehat{F}_{0n}^{(v)}(ds, dt) = \sum_{i=1}^n \widetilde{W}_{in}^{(v)} \varphi(Y_{[i:n]}, Z_{i:n}), \quad 0 \leq s \leq t \leq \tau_Z.$$

- Les observations i.i.d. sont composées $(Z_i, \delta_i, \delta_i V_i)_{1 \leq i \leq n}$.
- Les poids d'AJ, $\widetilde{W}_{in}^{(v)} = \widetilde{W}_{in} J_{[i:n]}^{(v)}$, (Suzukawa, 2002) pour un modèle à risques concurrents avec $J_{[i:n]}^{(v)} = \mathbb{1}_{V_i=v}$ sont

$$\widetilde{W}_{in}^{(v)} = \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{[j:n]}}.$$

Estimation à partir d'intégrales d'Aalen-Johansen

Plutôt que de regarder des intensités de transition, on travaille avec des probabilités de transition utilisées par les actuaires i.e.

$$p_{0e}(s, t, \eta) = \frac{\mathbb{P}(s < S \leq \min(t, t - \eta), T > t, V_1 = e)}{\mathbb{P}(S > s)},$$

$$p_{ee}(s, t) = \frac{\mathbb{P}(S \leq s, T > t, V_1 = e)}{\mathbb{P}(S \leq s, T > s, V_1 = e)},$$

$$p_{ed}(s, \eta, \zeta) = \frac{\mathbb{P}(\eta < T - S \leq \zeta, S = s, V = (e, d))}{\mathbb{P}(S = s, T - S > \eta, V_1 = e)}.$$

Estimation à partir d'intégrales d'Aalen-Johansen

Ces probabilités sont alors estimées à partir des grandeurs suivantes

$$\hat{p}_{0e}(s, t, \eta) = \frac{\hat{S}_n^{(e, C_e)}(\varphi_{s, t, \eta}^{(1)})}{1 - \hat{H}_n(s)}, \text{ avec } \varphi_{s, t, \eta}^{(1)}(x, y) = \mathbb{1}_{\{s < x \leq \min(t, t - \eta), y > t\}},$$

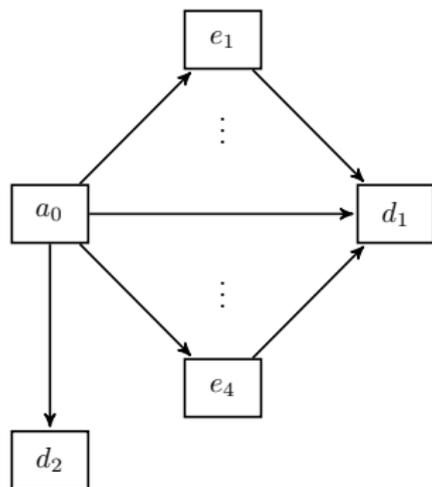
$$\hat{p}_{ee}(s, t) = \frac{\hat{S}_n^{(e, C_e)}(\varphi_{s, t}^{(2)})}{\hat{S}_n^{(e, C_e)}(\varphi_{s, s}^{(2)})}, \text{ avec } \varphi_{s, t}^{(2)}(x, y) = \mathbb{1}_{\{x \leq s, y > t\}},$$

$$\hat{p}_{ed}(s, \eta, \zeta) = \frac{\hat{S}_n^{(e, d)}(\varphi_{s, \zeta}^{(3)})}{\hat{S}_n^{(e, C_e)}(\varphi_{s, \eta}^{(4)})}, \text{ avec } \varphi_{s, \zeta}^{(3)}(x, y) = \mathbb{1}_{\{s = x, \eta < y - x \leq \zeta\}},$$

$\varphi_{s, \eta}^{(4)}(x, y) = \mathbb{1}_{\{s = x, \eta < y - x\}}$ et \hat{H}_n l'estimateur KM associé à la fonction de survie de S .

- Sous certaines hypothèses classiques, ces estimateurs sont **consistants** et **asymptotiquement normaux**. Les intervalles de confiance asymptotiques sont construits avec du **bootstrap**.

Application à l'assurance dépendance

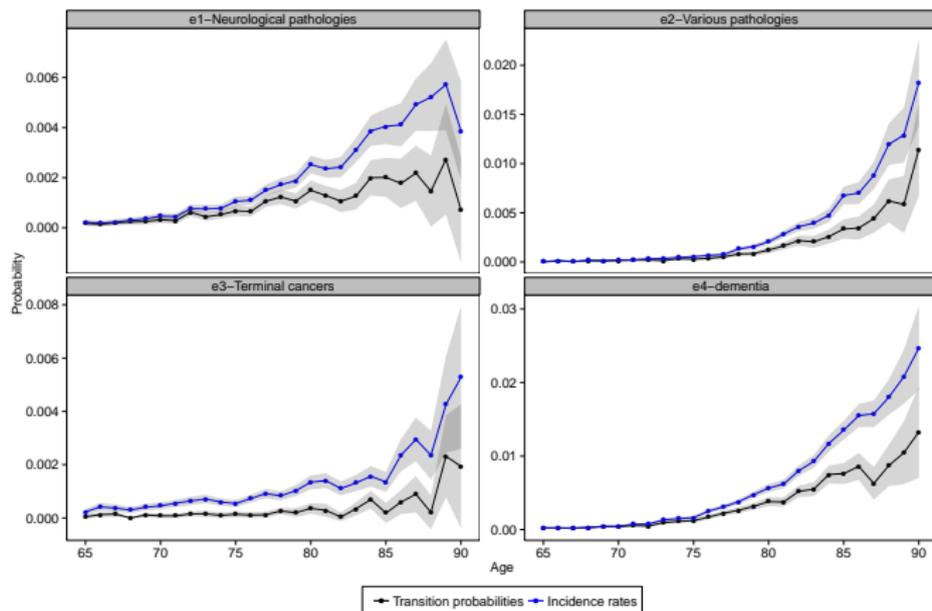


4 types de pathologies et 2 autres causes de sorties.

	Exit causes	%
e_1	Neurologic pathologies	2.5%
e_2	Various pathologies	2.7%
e_3	Terminal cancers	2.4%
e_4	Dementia	5.4%
d_1	Death	52.2%
d_2	Cancel	34.8%

Application à l'assurance dépendance

- Estimation des **probabilités** d'entrer et de rester en dépendance au moins 1 mois.
- Evaluation des intervalles de confiance ponctuels à 95% via 500 simulations de **bootstrap**.



Références I

- Andersen, P. K. and Keiding, N. (2012). Interpretability and importance of functionals in competing risks and multistate models. *Statistics in Medicine*, **31**(11-12), 1074–1088.
- Andersen, P. K. and Perme, M. P. (2008). Inference for outcome probabilities in multi-state models. *Lifetime data analysis*, **14**(4), 405–431. Cited by 0020.
- Andersen, P. K., Borgan, r., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer Series in Statistics. Springer-Verlag New York Inc.
- Beyersmann, J., Schumacher, M., and Allignol, A. (2011). *Competing Risks and Multistate Models with R*. Use R ! Springer.
- Buchardt, K. (2013). Dependent interest and transition rates in life insurance. Technical report, University of Copenhagen and PFA Pension, Copenhagen. 00002.
- Buchardt, K., Møller, T., and Schmidt, K. B. (2014). Cash flows and policyholder behaviour in the semi-markov life insurance setup. *Scandinavian Actuarial Journal*, pages 1–29. 00001.
- Carrière, J. F. (1994). Dependent decrement theory. *Transactions of the Society of Actuaries*, **46**, 45–74. Cited by 0000.
- Christiansen, M. C. (2012). Multistate models in health insurance. *Advances in Statistical Analysis*, **96**(2), 155–186.
- CMIR12 (1991). The analysis of permanent health insurance data. Technical report, Continuous Mortality Investigation Bureau, The Institute of Actuaries and the Faculty of Actuaries.

Références II

- Commenges, D. (2002). Inference for multi-state models from interval-censored data. *Statistical Methods in Medical Research*, **11**(2), 167–182.
- Commenges, D., Joly, P., Gégout-Petit, A., and Liquef, B. (2007). Choice between semi-parametric estimators of markov and non-markov multi-state models from coarsened observations. *Scandinavian Journal of Statistics*, **34**(1), 33–52. Cited by 0030.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, **34**(2), 187–220.
- Datta, S. and Satten, G. A. (2002). Estimation of integrated transition hazards and stage occupation probabilities for non-markov systems under dependent censoring. *Biometrics*, **58**(4), 792–802.
- De Wreede, L. C., Fiocco, M., and Putter, H. (2011). mstate: An r package for the analysis of competing risks and multi-state models. *Journal of Statistical Software*, **38**(7).
- Deléglise, M. P., Hess, C., and Nouet, S. (2009). Tarification, provisionnement et pilotage d'un portefeuille dépendance. *Bulletin Français d'Actuariat*, **9**(17), 70–108.
- Denuit, M. and Robert, C. (2007). *Actuariat des assurances de Personnes - Modélisation, tarification et provisionnement*. Assurance Audit Actuariat. Economica, Paris.
- Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, **94**(446), 496–509.
- Guibert, Q. and Planchet, F. (2014a). Construction de lois d'expérience en présence d'évènements concurrents – application à l'estimation des lois d'incidence d'un contrat dépendance. *Bulletin Français d'Actuariat*, **13**(27), 5–28.

Références III

- Guibert, Q. and Planchet, F. (2014b). Non-parametric inference of transition probabilities based on aalen-johansen integral estimators for semi-competing risks data: Application to LTC insurance. Les cahiers de recherche de l'ISFA 2014-14, ISFA, Lyon.
- Haberman, S. and Pitacco, E. (1998). *Actuarial Models for Disability Insurance*. Chapman and Hall/CRC, 1 edition.
- Helwich, M. (2008). *Durational effects and non-smooth semi-Markov models in life insurance*. Ph.d. thesis, University of Rostock.
- Janssen, J. and Manca, R. (2006). *Applied Semi-Markov Processes*. Springer.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**(282), 457–481. Cited by 41330.
- Levantesi, S. and Menzietti, M. (2012). Managing longevity and disability risks in life annuities with long term care. *Insurance: Mathematics and Economics*, **50**(3), 391–401.
- Martinussen, T. and Scheike, T. H. (2006). *Dynamic Regression Models for Survival Data*. Statistics for Biology and Health. Springer. Cited by 0250.
- Meira-Machado, L., de Uña-Álvarez, J., and Cadarso-Suárez, C. (2006). Nonparametric estimation of transition probabilities in a non-markov illness–death model. *Lifetime Data Analysis*, **12**(3), 325–344.
- Monteiro, A., Smirnov, G. V., and Lucas, A. (2006). Non-parametric estimation for nonhomogeneous semi-markov processes: An application to credit risk. Discussion Paper TI 2006-024/2, Tinbergen Institute, Amsterdam.

Références IV

- Peng, L. and Fine, J. P. (2006). Nonparametric estimation with left-truncated semicompeting risks data. *Biometrika*, **93**(2), 367–383.
- Prentice, R. L., Kalbfleisch, J. D., Peterson, A. V., Flournoy, N., Farewell, V. T., and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, **34**(4), 541–554.
- Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, **26**(11), 2389–2430.
- Suzukawa, A. (2002). Asymptotic properties of aalen-johansen integrals for competing risks data. *Journal of the Japan Statistical Society*, **32**, 77–93.
- Tomas, J. and Planchet, F. (2013). Multidimensional smoothing by adaptive local kernel-weighted log-likelihood: Application to long-term care insurance. *Insurance: Mathematics and Economics*, **52**(3), 573–589.