

Extensions du modèle de Cox

Variables dépendantes du temps

Effets non linéaires

Pr Roch Giorgi

roch.giorgi@univ-amu.fr

Analyse des Durées de Survie

Introduction

- Temps écoulé jusqu'à la survenue d'un événement terminal appelé « décès »
 - Mort
 - Rechute, récurrence
 - Guérison
 - ...
- Etude du délai de la survenue de cet événement
 - Estimation de la (des) distribution(s) des temps de survie (fonction de survie)
 - Comparaison des fonctions de survie entre $k \geq 2$ groupes
 - Analyse de l'effet de variables explicatives sur les fonctions de survie

Analyse des Durées de Survie

Définitions

- Date* d'origine
 - Origine de la durée étudiée
 - Date du diagnostic, de l'entrée dans l'étude, de l'exposition à un facteur,...
 - Propre (potentiellement) à chaque individu
- Date* de point
 - Arrêt de l'étude
 - La même pour tous les individus
 - Au-delà, les informations sur les individus ne sont plus prises en compte
- Date* des dernières nouvelles

* *Calendaire*

Analyse des Durées de Survie

Définitions

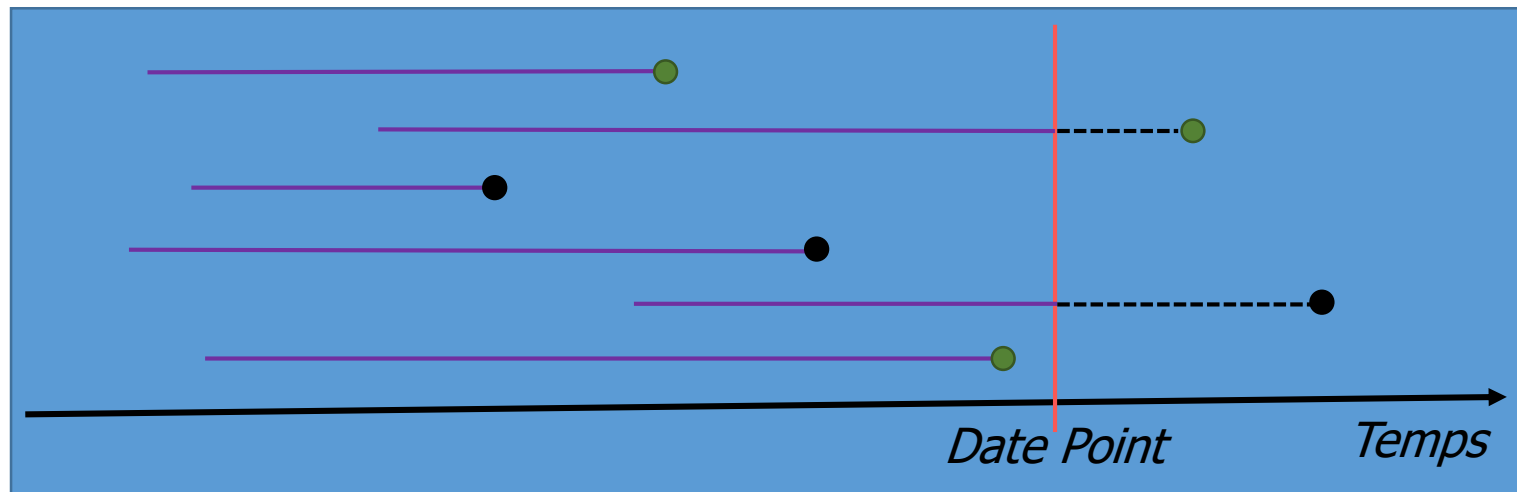
- Variables d'intérêt

- Durée

- Temps (jours, mois)
 - $\min(\text{Date dernières nouvelles, date de point}) - \text{Date origine}$

- Statut

- Binaire
 - Evènement terminal **observé** ou pas



Analyse des Durées de Survie

Définitions

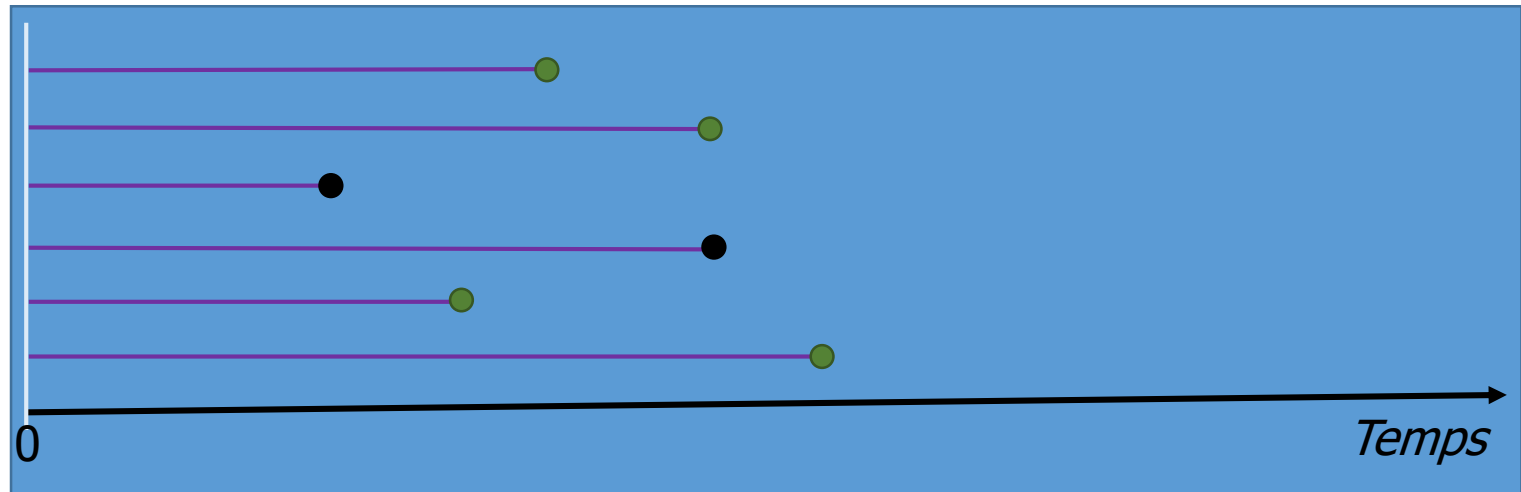
- Variables d'intérêt

- Durée

- Temps (jours, mois)
 - $\min(\text{Date dernières nouvelles, date de point}) - \text{Date origine}$

- Statut

- Binaire
 - Evènement terminal **observé** ou pas



Distributions de la Durée de Survie

- Durée de survie
 - Variable aléatoire positive ou nulle T continue
- Loi de probabilité définie par différentes fonctions
 - survie S
 - répartition F
 - densité de probabilité f
 - risque instantané (hasard) λ
 - risque cumulé (hasard) Λ
- Ces fonctions sont reliées entre elles

Distributions de la Durée de Survie

Fonction de survie S

- Probabilité de survivre jusqu'à l'instant t

$$S(t) = P(T \geq t), t \geq 0$$

Distributions de la Durée de Survie

Fonction de répartition F

- Probabilité de décéder avant l'instant t

$$F(t) = P(T \leq t) = 1 - S(t), t \geq 0$$

Distributions de la Durée de Survie

Fonction de densité de probabilité f

- Probabilité de décéder dans un petit intervalle de temps après l'instant t

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

$$\text{avec } F(t) = \int_0^t f(u) du$$

Distributions de la Durée de Survie

Fonction de risque (taux) instantané λ

- Probabilité de décéder dans un petit intervalle de temps après l'instant t , sachant que l'on a survécu jusqu'à t

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t \leq T < t+h | T \geq t)}{h} = \frac{f(t)}{S(t)} = -\ln(S(t))', t \geq 0$$

Distributions de la Durée de Survie

Fonction de risque cumulé Λ

- Probabilité de décéder sur l'intervalle de temps allant de 0 à t

$$\Lambda(t) = \int_0^t \lambda(u) du = -\ln(S(t)), t \geq 0$$

d'où

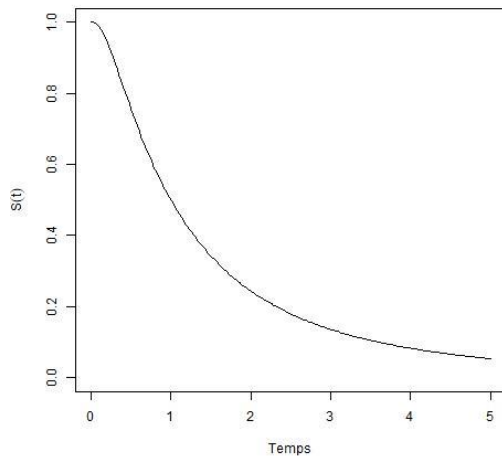
$$S(t) = \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(u) du\right)$$

$$f(t) = \lambda(t) \exp\left(-\int_0^t \lambda(u) du\right)$$

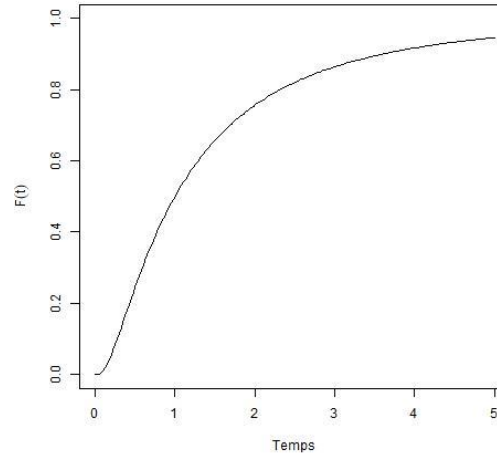
Distributions de la Durée de Survie

Fonction de

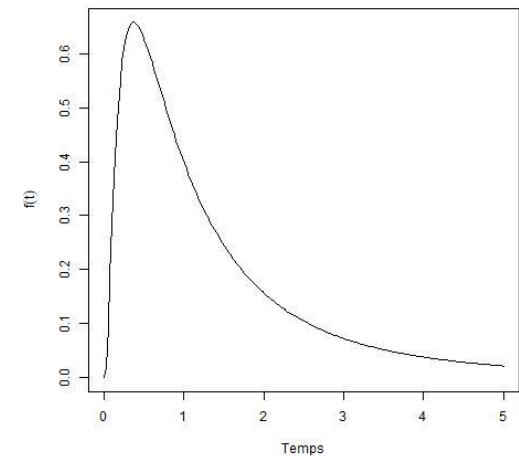
survie



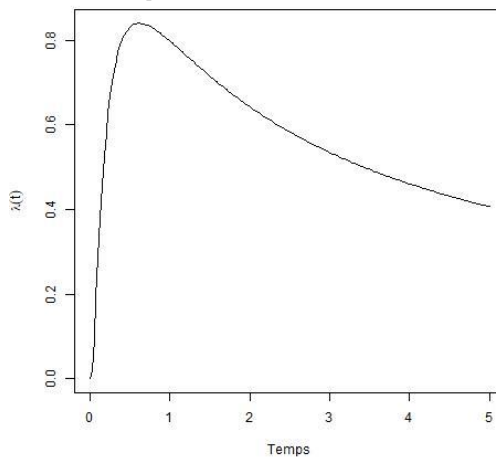
répartition



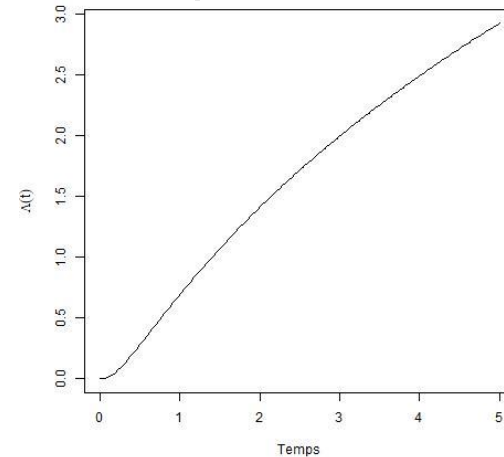
densité de probabilité



risque instantané



risque cumulé



Observations Incomplètes

Censure et Troncature

- La réalisation de la variable aléatoire T n'est pas toujours observée
 - Processus de censure
 - Processus de troncature
- Exemple essai clinique
 - Censure administrative, du fait de la fin de l'étude

Censure

- Pour chaque individus i on a
 - T_i et C_i , les variables aléatoires temps de survie et temps de censure
- Censure à droite
 - On observe $t_i = \min(T_i, C_i)$ et $\delta_i = I_{\{T_i \leq C_i\}}$
 - $\delta_i = 1$ si l'individu décède, 0 s'il est censuré
 - Hypothèse de censure aléatoire
 - Hypothèse d'indépendance entre T_i et C_i
- Censure par intervalle
 - Événement survenu entre 2 temps connus
- *Censure à gauche*
 - *Événement survenu avant que l'individu soit observé*

Troncature

- Concerne l'échantillonnage
 - Seule une partie des individus (sous-échantillon) est observable

Modèle de Cox

Modèle à Taux Proportionnels

- Repose sur la modélisation du taux auquel les individus présentent l'événement d'intérêt
- Conditionnellement à leurs covariables \mathbf{z}

$$\lambda(t|\mathbf{z}) = -\frac{d\ln(S(t|\mathbf{z}))}{d(t)} = \frac{f(t|\mathbf{z})}{S(t|\mathbf{z})}$$

Modèle de Cox

Modèle à Taux Proportionnels

$$\lambda(t|\mathbf{z}) = \lambda_0(t)\exp(\beta^t \mathbf{z})$$

- β vecteur des paramètres associés aux vecteur de covariables \mathbf{z}
- $\lambda_0(t)$ taux de mortalité de base

Modèle de Cox

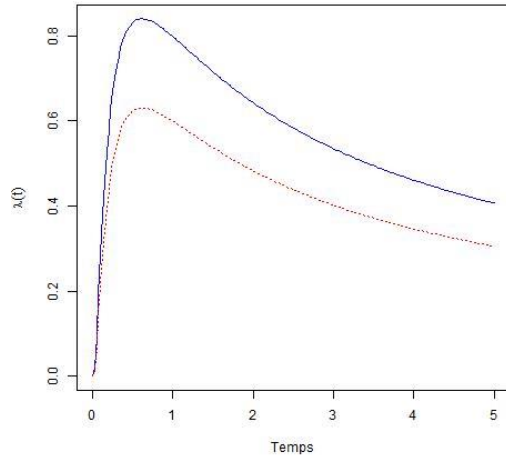
Modèle à Taux Proportionnels

- Pour 2 individus i, j ayant pour covariables z_i et z_j

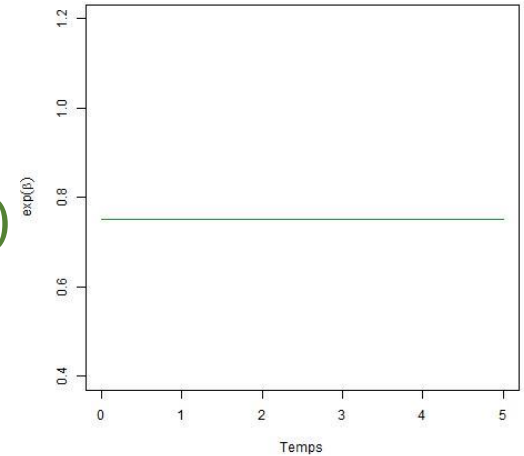
$$\frac{\lambda(t|z_i)}{\lambda(t|z_j)} = \frac{\lambda_0(t)\exp(\beta^t z_i)}{\lambda_0(t)\exp(\beta^t z_j)} = \exp(\beta^t(z_i - z_j))$$

- Par conséquence, les fonctions du taux sont proportionnelles, leur rapport est constant et ne dépend pas de t
 - Hypothèse à vérifier au vue des données
- Rapport des fonctions du taux est un risque relatif à l'instant t des sujets de caractéristique z_i par rapport à ceux de caractéristique z_j

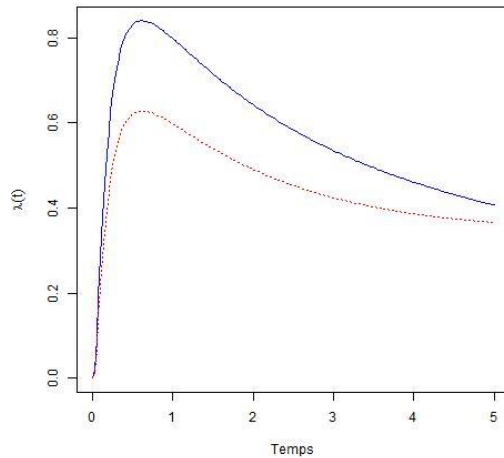
Modèle à Taux Proportionnels



$$\frac{\lambda(t|z_1)}{\lambda(t|z_2)} = \frac{\lambda_0(t)\exp(\beta z_1)}{\lambda_0(t)\exp(\beta z_2)} = \exp(\beta)$$

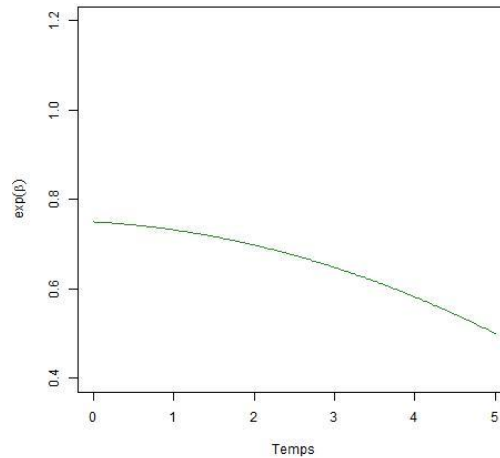


Modèle à Taux non Proportionnels



$$\lambda(t|z_1)$$

$$\lambda(t|z_2)$$



$$\frac{\lambda_0(t)\exp(\beta z_1)}{\lambda_0(t)\exp(\beta z_2)} = \exp(\beta(t))$$

Modèle de Cox

Modèle log-linéaire sur l'effet des covariables

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda(t|z_i)}{\lambda(t|z_j)} = \frac{\lambda_0(t)\exp(\beta^t z_i)}{\lambda_0(t)\exp(\beta^t z_j)} = \exp(\beta^t(z_i - z_j))$$

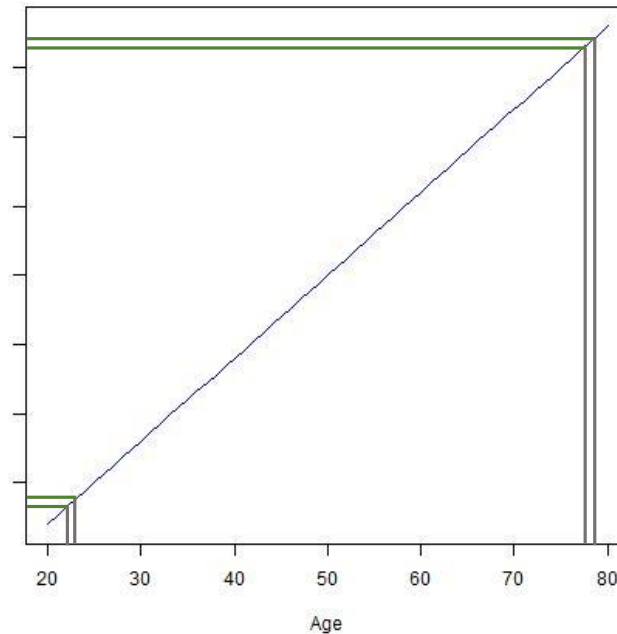
$$\ln\left(\frac{\lambda_0(t)\exp(\beta^t z_i)}{\lambda_0(t)\exp(\beta^t z_j)}\right) = \beta^t(z_i - z_j)$$

- Par conséquent, le logarithme du rapport des taux de décès est une fonction linéaire des covariables
 - Hypothèse à vérifier au vue des données

Effet Log-Linéaire sur une Covariable

$$\beta(age_{78} - age_{77})$$

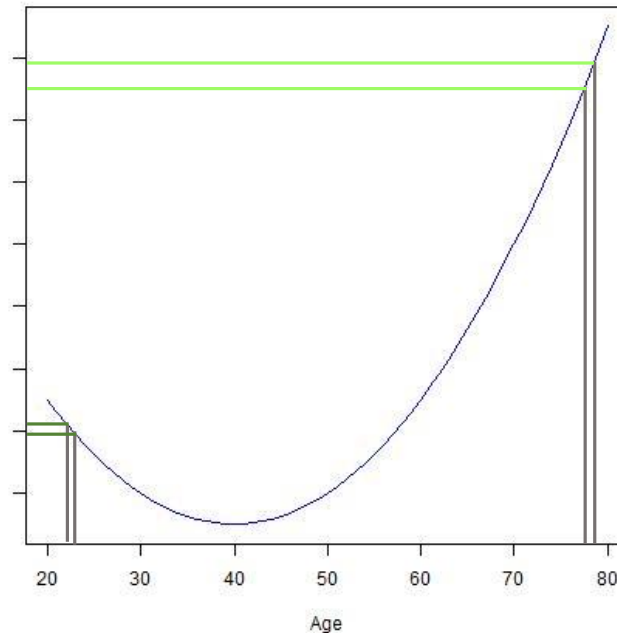
$$\beta(age_{24} - age_{23})$$



Effet Non Log-Linéaire sur une Covariable

$$\beta_{age}(age_{78} - age_{77})$$

$$\beta_{age}(age_{24} - age_{23})$$



Modèle de Cox

Modèle Linéaire Généralisé pour $S(t)$

Taux instantané : $\lambda(t|\mathbf{z}) = \lambda_0(t)\exp(\beta^t \mathbf{z})$

Taux cumulé : $\Lambda(t|\mathbf{z}) = \Lambda_0(t)\exp(\beta^t \mathbf{z})$

Survie : $S(t|\mathbf{z}) = \exp(-\Lambda(t|\mathbf{z})) = S_0(t|\mathbf{z})\exp(\beta^t \mathbf{z})$

$$\ln[S(t|\mathbf{z})] = -\Lambda_0(t|\mathbf{z})\exp(\beta^t \mathbf{z})$$

$$\ln[-\ln[S(t|\mathbf{z})]] = \ln[\Lambda_0(t|\mathbf{z})] + \beta^t \mathbf{z}$$

où $\beta^t \mathbf{z}$ est le prédicteur linéaire
(index pronostique)

Modèle de Cox

Modèle à Taux Proportionnels

$$\lambda(t|\mathbf{z}) = \lambda_0(t)\exp(\beta^t \mathbf{z})$$

- Vecteur de covariables \mathbf{z}
 - Variable « pure »
 - Variable « transformée » (log,...)
 - Interaction entre variables
- β vecteur des paramètres
 - N'inclus pas de terme constant (intégré dans $\ln(\lambda_0(t))$)
 - Si, \mathbf{z} est centrée ($\mathbf{z} - E[\mathbf{z}]$)
 - ne change pas la valeur estimée de β
 - change $\lambda_0(t)$

Vraisemblance Partielle de Cox

- A l'instant t_i ($i = 1, \dots, n$ décès) il y a \mathcal{R}_i individus encore à risque
- La probabilité de décès en t_i de chaque individu j est
$$\lambda_0(t_i) \exp(\beta^t \mathbf{z}_j) \Delta t$$

- La probabilité que ce soit l'individu i qui décède est

$$v_i(\beta) = \frac{\lambda_0(t_i) \exp(\beta^t \mathbf{z}_i) \Delta t}{\sum_{j \in \mathcal{R}_i} \lambda_0(t_i) \exp(\beta^t \mathbf{z}_j) \Delta t} = \frac{\exp(\beta^t \mathbf{z}_i)}{\sum_{j \in \mathcal{R}_i} \exp(\beta^t \mathbf{z}_j)}$$

$$V(\beta) = \prod_{i=1}^n v_i(\beta) = \prod_{i=1}^n \frac{\exp(\beta^t \mathbf{z}_i)}{\sum_{j \in \mathcal{R}_i} \exp(\beta^t \mathbf{z}_j)}$$

Vraisemblance Partielle de Cox

- Les censures ne participent pas au calcul de la vraisemblance
 - Hypothèse de censures non-informatives
 - Hypothèse d'indépendance entre les temps de censures et les temps de décès
- La vraisemblance partielle ne dépend pas de $\lambda_0(t)$
 - β estimé, sans connaissance de $\lambda_0(t)$ (pas estimé), par maximisation de la vraisemblance partielle

Estimation des Paramètres β

- Maximisation de la log-vraisemblance

$$L(\beta) = \sum_{i=1}^n \left[\beta^t \mathbf{z}_i - \ln \left(\sum_{j \in \mathcal{R}_i} \exp(\beta^t \mathbf{z}_i) \right) \right]$$

- L'estimation du maximum de vraisemblance est la valeur $\hat{\beta}$ de β qui rend maximum $L(\beta)$ (ou $V(\beta)$)
- Au maximum, la fonction score

$$U(\hat{\beta}) = \frac{\partial L(\beta)}{\partial(\beta)} = \left(\frac{\partial L(\beta)}{\partial(\beta_1)}, \dots, \frac{\partial L(\beta)}{\partial(\beta_p)} \right) = 0$$

Estimation du Taux Cumulé de Base Λ_0

- Estimateur de Breslow (extension de l'estimateur de Nelson-Aalen)

$$\hat{\Lambda}_0(t|\mathbf{z}) = \sum_{t_i \leq t} \frac{d_i}{\sum_{j \in \mathcal{R}_i} \exp(\hat{\beta}^t \mathbf{z}_i)}$$

- d_i nombre de décès en t_i

Estimation de la Fonction de Survie S

$$\hat{S}(t|\mathbf{z}) = \exp\left(-\hat{\Lambda}_0(t)\exp(\hat{\beta}^t\mathbf{z})\right)$$

- Remarque
 - $\hat{\Lambda}_0$ et \hat{S} peuvent ainsi être estimés dans R à l'aide de fonctions du package `survival`
 - Fonction `survfit` sur un objet de classe `coxph`

Interprétation des Paramètres β

- Pour 2 individus i, j ayant pour covariables z_i et z_j

$$\frac{\lambda(t|z_i)}{\lambda(t|z_j)} = \frac{\lambda_0(t)\exp(\beta^t z_i)}{\lambda_0(t)\exp(\beta^t z_j)} = \exp\left(\beta^t(z_i - z_j)\right) = RR(t) = RR$$

- z binaire (sexe, codé en 0 si féminin et 1 si masculin)
 - $RR = \exp(\beta(z = 1 - z = 0)) = \exp(\beta)$
 - **Hypothèse des risques proportionnels**
- z continu (âge, en années en prenant par exemple ces 2 valeurs)
 - $RR = \exp(\beta(z = 25 - z = 24)) = \exp(\beta)$
 - **Hypothèse de log-linéarité**
- Interprétation comme n'importe quel RR par rapport à 1

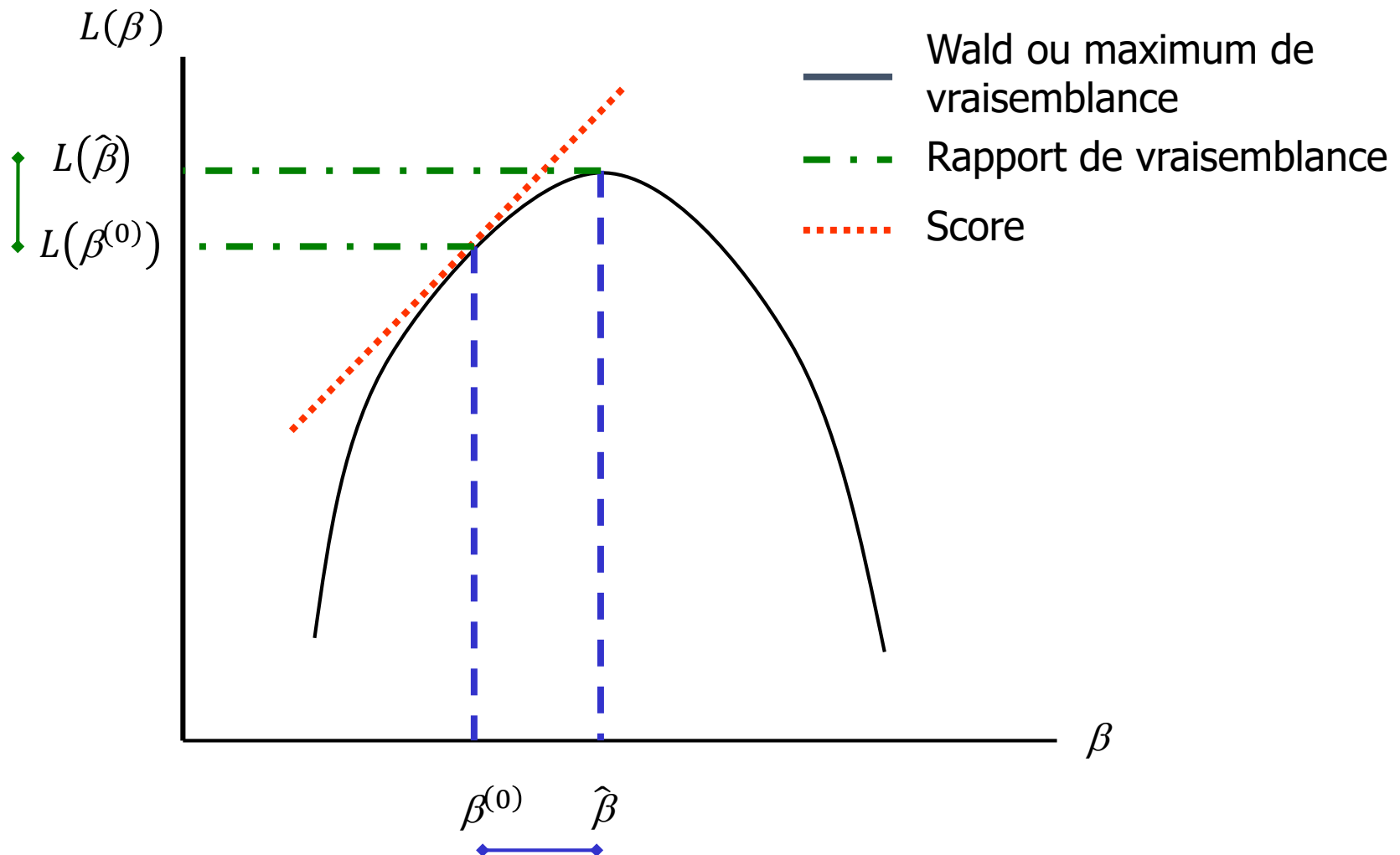
Tests des Paramètres β

Une seule variable (binaire ou continue)

- Hypothèse nulle $H_0: \beta = \beta^{(0)}$ (dimension de $\beta = 1$)
 - Test
 - Rapport de vraisemblance
 - $2[L(\hat{\beta}) - L(\beta^{(0)})]$
 - Wald
 - $\hat{\beta}^2 / \text{Var}(\hat{\beta})$
 - Score
 - $U(\beta^{(0)}) / \text{Var}(U(\beta^{(0)}))$
- } $\xrightarrow{H_0} \chi^2$ à 1 degré de liberté

Tests des Paramètres β

Une seule variable (binaire ou continue)



Tests des Paramètres β

Une seule (nominale), plusieurs variables

- Hypothèse nulle $H_0: \beta = \beta^{(0)}$ (dimension de $\beta = p$)
- Test
 - Rapport de vraisemblance
 - $2[L(\hat{\beta}) - L(\beta^{(0)})]$
 - Wald
 - $(\hat{\beta} - \beta^{(0)})^t I(\hat{\beta})^{-1} (\hat{\beta} - \beta^{(0)})$
 - Score
 - $U^t(\beta^{(0)}) I(\beta^{(0)})^{-1} U(\beta^{(0)})$

$\xrightarrow{H_0} \chi^2$ à p degrés de liberté

où I est la matrice d'information de Fisher

Tests des Paramètres β

Plusieurs variables

- Remarque

- Test du rapport de vraisemblance pour comparer 2 modèles emboîtés

- Modèle A

$$\lambda^A(t|z_1, z_2) = \lambda_0^A(t) \exp(\beta_1 z_1 + \beta_2 z_2)$$

- Modèle B

$$\lambda^B(t|z_1, z_2, z_3, z_4) = \lambda_0^B(t) \exp(\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4)$$

avec z_1, z_2, z_3, z_4 binaires

$$\Rightarrow 2[L(\hat{\beta}^B) - L(\hat{\beta}^A)] \xrightarrow{H_0} \chi^2 \text{ à } 2 \text{ degrés de liberté}$$

Exemple (Peto, 1979) : Données

Temps participation	Traitement	Fonction rénale	Temps participation	Traitement	Fonction rénale
8	A	AN	220	A	N
8	A	N	365*	A	N
13	B	AN	632	B	N
18	B	AN	700	B	N
23	B	AN	852*	A	N
52	A	AN	1296	B	N
63	A	AN	1296*	A	N
63	A	AN	1328*	A	N
70	B	N	1460*	A	N
76	B	N	1976*	A	N
180	B	N	1990*	B	N
195	B	N	2240*	B	N
210	B	N			

* Censures ; N : normale ; AN : anormale

Exemple (Peto, 1979) : Traitement

Modèle : $\lambda(t, \mathbf{z}) = \lambda_0(t) \exp(\beta_1 z_1)$ avec $z_1=1$ si traitement B

```
coxph(Surv(temps, statut) ~ trt, data=DataPeto)
```

	coef	exp(coef)	se(coef)	z	p
trt	0.561	1.75	0.51	1.1	0.27

	exp(coef)	exp(-coef)	lower .95	upper .95
trt	1.75	0.571	0.645	4.76

Log Vraisemblance initiale = -46.9

Log Vraisemblance finale = -46.3

Likelihood ratio test = 1.26 on 1 df, p=0.261

Wald test = 1.21 on 1 df, p=0.271

Score (logrank) test = 1.24 on 1 df, p=0.265

→ $2(-46.3 + 46.9)$

→ $0.561^2 / 0.51^2$

Exemple (Peto, 1979) : Fonction Rénale

Modèle : $\lambda(t, \mathbf{z}) = \lambda_0(t) \exp(\beta_2 z_2)$ avec $z_2=1$ si fonction rénale AN

```
coxph(Surv(temps, statut) ~ fr, data=DataPeto)
```

	coef	exp(coef)	se(coef)	z	p
fr	3.64	38.1	1.1	3.31	0.00094

	exp(coef)	exp(-coef)	lower .95	upper .95
fr	38.1	0.0262	4.41	330

Log Vraisemblance initiale = -46.9

Log Vraisemblance finale = -37.6

Likelihood ratio test= 18.7 on 1 df, p=0.0000152

Wald test = 10.9 on 1 df, p=0.000939

Score (logrank) test = 24.9 on 1 df, p=5.96e-007

Exemple (Peto, 1979) : Ajustement

Modèle : $\lambda(t, \mathbf{z}) = \lambda_0(t) \exp(\beta_1 z_1 + \beta_2 z_2)$

```
coxph(Surv(temps, statut) ~ trt + fr, data=DataPeto)
```

	coef	exp(coef)	se(coef)	z	p
trt	1.22	3.4	0.598	2.05	0.04100
fr	4.28	72.1	1.191	3.59	0.00033

	exp(coef)	exp(-coef)	lower .95	upper .95
TRT	3.4	0.2945	1.05	11
fr	72.1	0.0139	6.98	745

Log Vraisemblance initiale = -46.9

Log Vraisemblance finale = -35.2

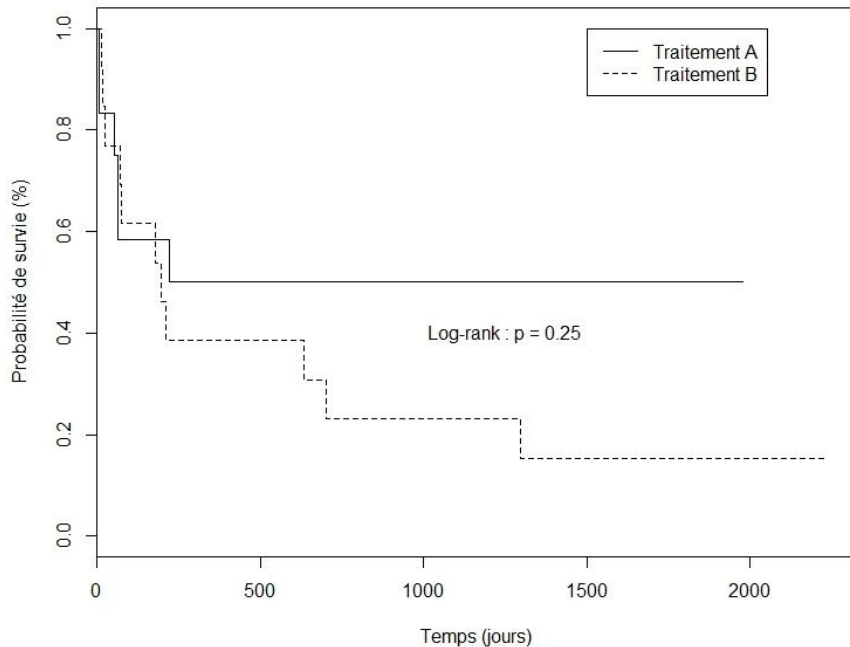
Likelihood ratio test= 23.5 on 2 df, p=7.86e-006

Wald test = 13.9 on 2 df, p=0.000982

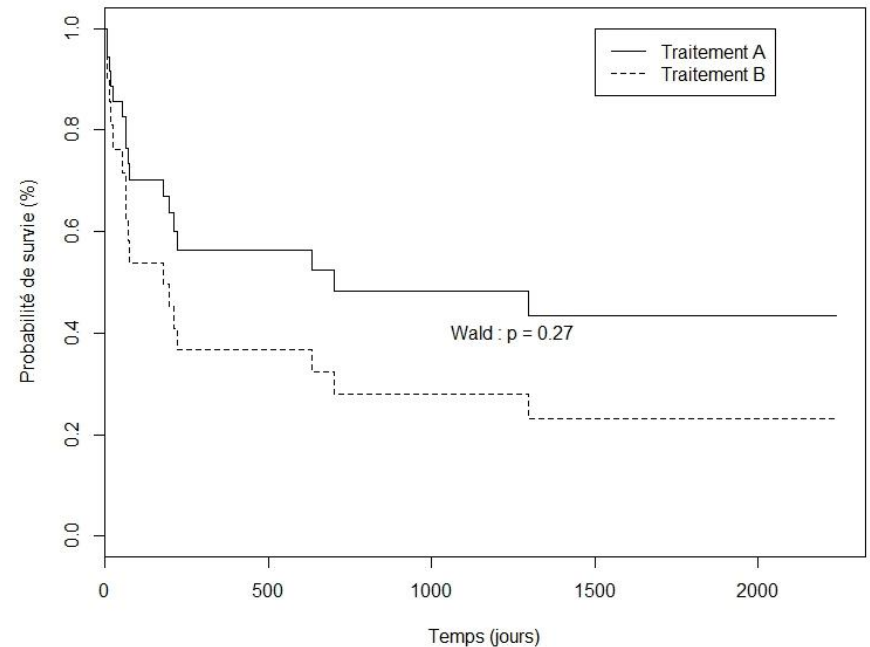
Score (logrank) test = 28.3 on 2 df, p=7.32e-007

Exemple (Peto, 1979) : Courbes de Survie

Kaplan-Meier



Cox



— Différences/Similitudes et raisons ?

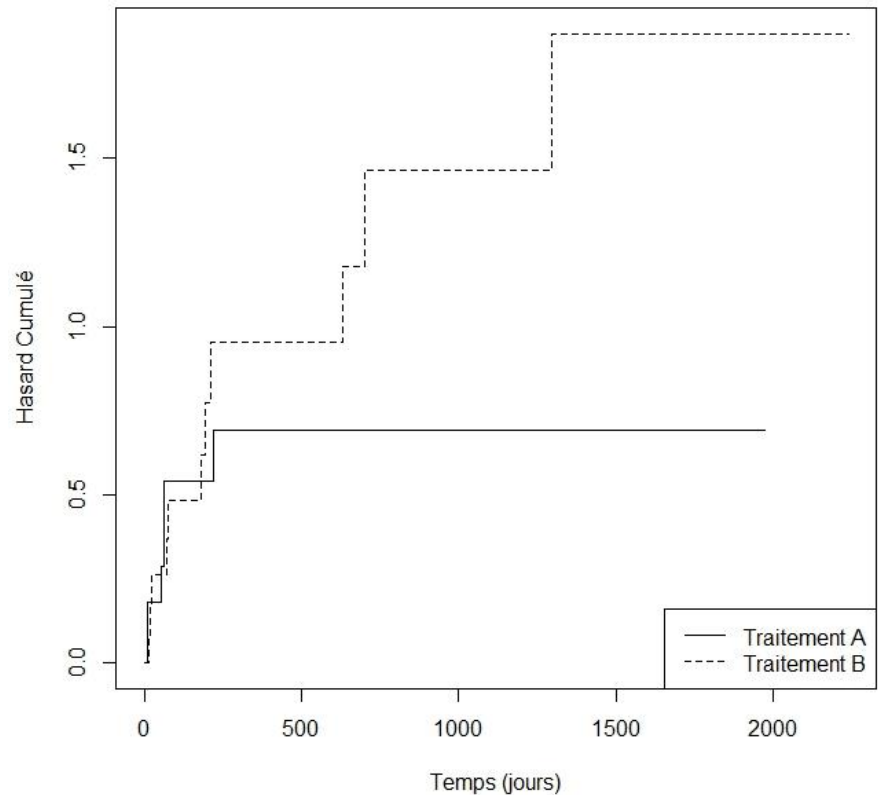
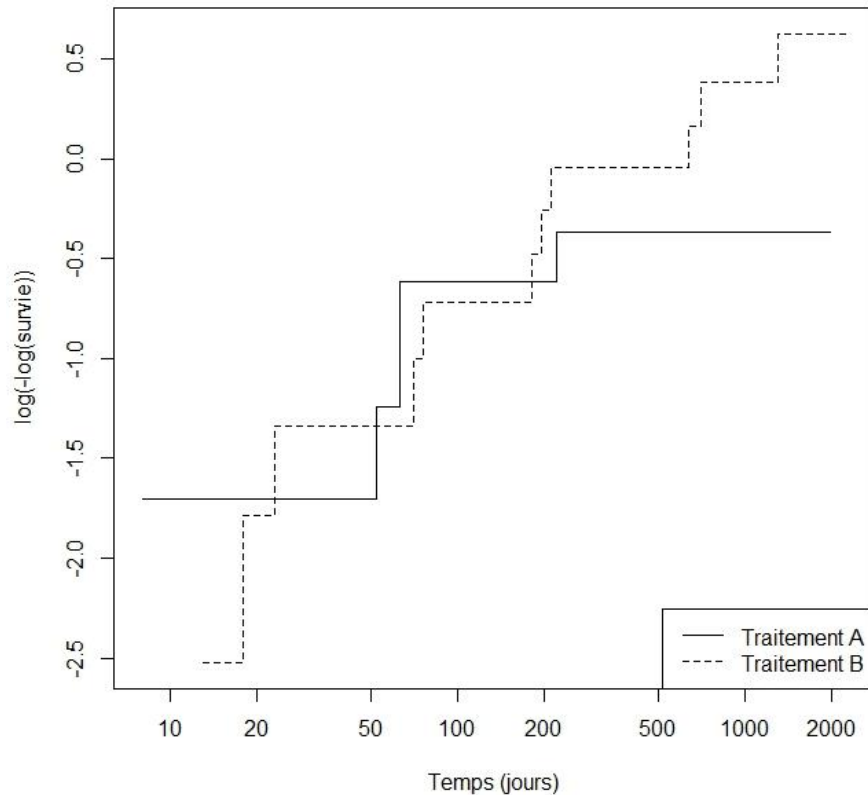
Hypothèse des Taux Proportionnels

- Modélisation à tort de l'effet d'un facteur par un modèle à taux proportionnel
 - Estimation d'un effet moyen
 - Possible conclusion à tort à l'absence d'effet
- Vérification de l'hypothèse
 - Graphique
 - $\ln \left[-\ln[\hat{S}(t|z)] \right] = \ln[\hat{\Lambda}_0(t|z)] + \hat{\beta}z$
 - Etude des résidus de Schoenfeld
 - Test des résidus
 - Graphique des résidus standardisés lissés en fonction du temps
 - Modélisation et test d'un effet dépendant du temps
 - $\lambda(t|z) = \lambda_0(t) \exp(\beta z + \gamma f(t)z)$

Hypothèse des Taux Proportionnels

Graphique

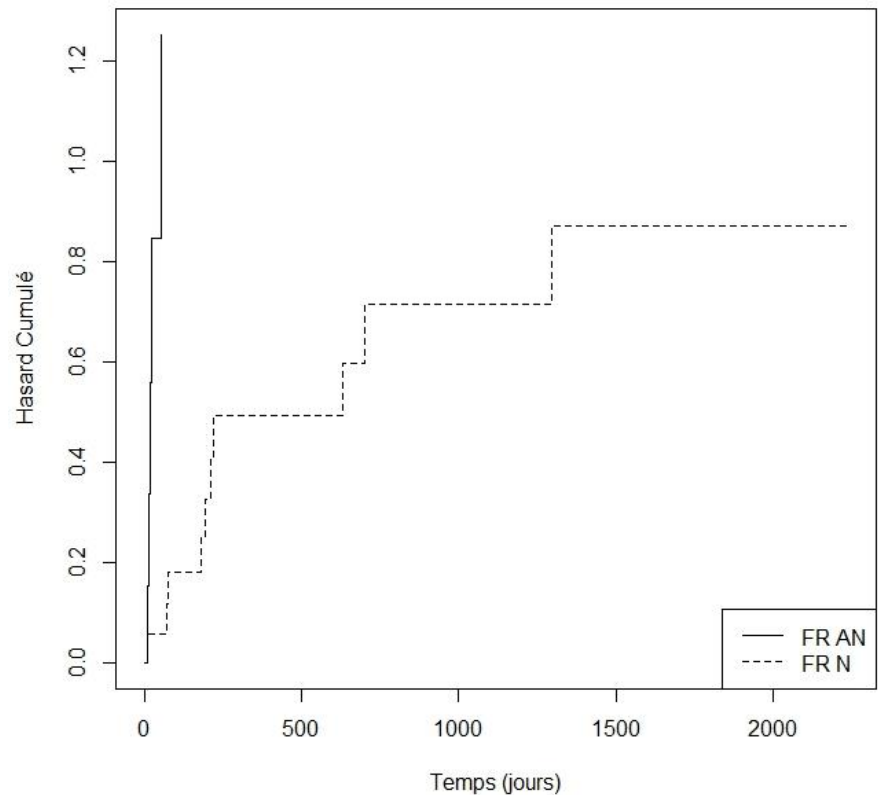
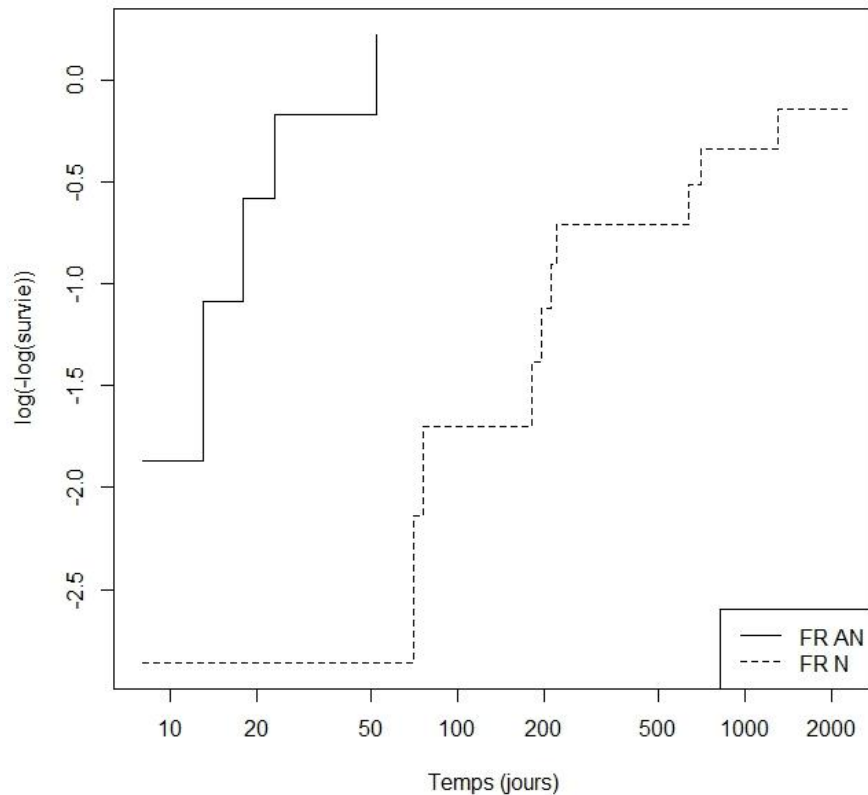
Exemple (Peto, 1979) : traitement



Hypothèse des Taux Proportionnels

Graphique

Exemple (Peto, 1979) : fonction rénale



Hypothèse des Taux Proportionnels

Résidus de Schoenfeld

- Correspondent à la contribution au score de chaque décès
- Une covariable, pas d'ex-aequo : contribution du $i^{\text{ème}}$ décès (si ex-aequo, somme des résidus)
 - $z_i - \bar{z}_i(\hat{\beta})$
- Plusieurs covariables : chaque covariable fournit son résidu
 - $z_{ki} - \bar{z}_{ki}(\hat{\beta})$

Hypothèse des Taux Proportionnels

Résidus de Schoenfeld

- Test des résidus
 - Fonction `cox.zph` du package R `survival`
 - Hypothèse nulle H_0 : taux proportionnels

Exemple (Peto, 1979) :

traitement

	rho	chisq	p
trtB	0.489	4.1	0.0428

fonction rénale

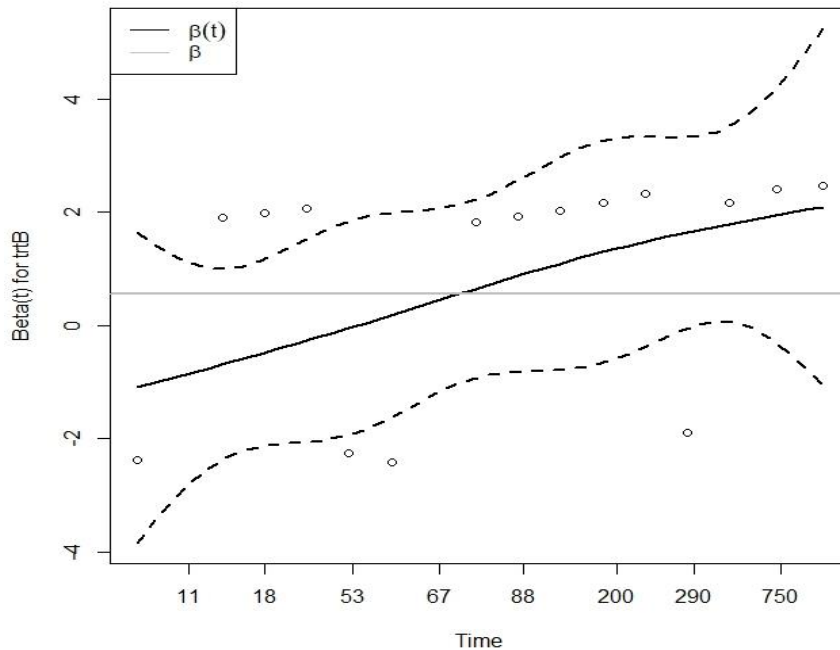
	rho	chisq	p
frN	-0.207	0.916	0.339

Hypothèse des Taux Proportionnels

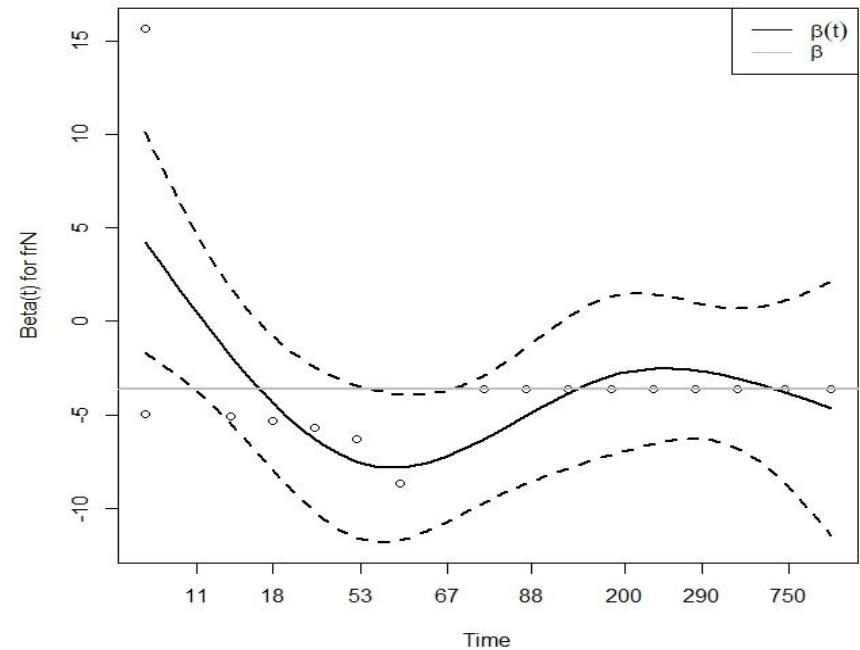
Résidus de Schoenfeld

- Graphique des résidus standardisés lissés
 - Fonction `cox.zph` du package R `survival`

Exemple (Peto, 1979) :
traitement



fonction rénale



Hypothèse de Log-Linéarité

- Modélisation à tort d'un effet linéaire d'une covariable continue
 - Estimation biaisée
 - Possible conclusion à tort à l'absence d'effet
- Vérification de l'hypothèse
 - Etude des résidus de Martingale
 - Graphique des résidus lissés en fonction des valeurs de la covariable continue

Hypothèse de Log-Linéarité

Résidus de Martingale

- Permettent d'étudier la relation fonctionnelle entre une variable et les taux de décès
- On recherche une transformation de la covariable z telle que la variable résultante – **résidus** – obéisse à la même loi pour tous les sujets
- Si z a un effet de la forme $f(z)$, alors ses résidus sont approximativement proportionnels à $f(z)$

Hypothèse de Log-Linéarité

Résidus de Martingale

- Comparent, pour un sujet i , l'observation de l'événement (δ_i) à celle qui est attendue, conditionnellement au modèle testé, et mesurée par le taux cumulé de la population

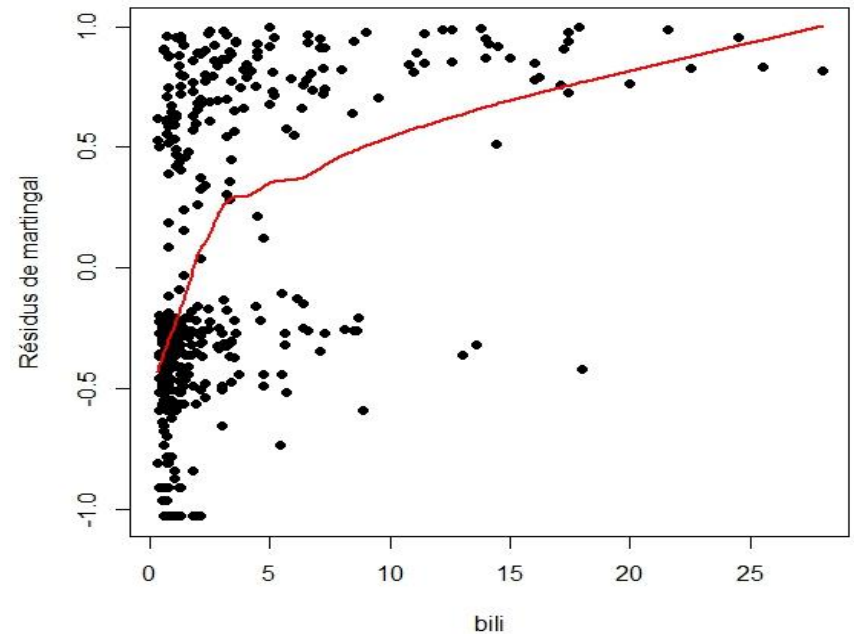
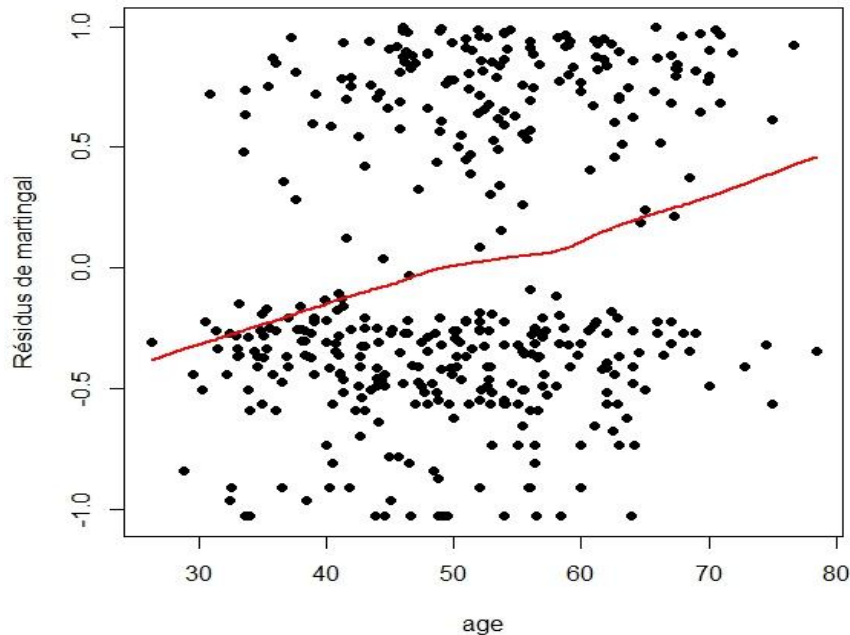
$$res_i = \delta_i - \hat{\Lambda}(t_i | \mathbf{z}_i)$$

- Évaluation de la forme
 - Représentation du nuage de points $(z_i; res_i)$ et superposition de $f(z)$ lissée
- Calcul des résidus de Martingale et lissage
 - Fonction `residuals` du package R `survival`
 - Fonction `lowess` du package R `stats`

Hypothèse de Log-Linéarité

Résidus de Martingale

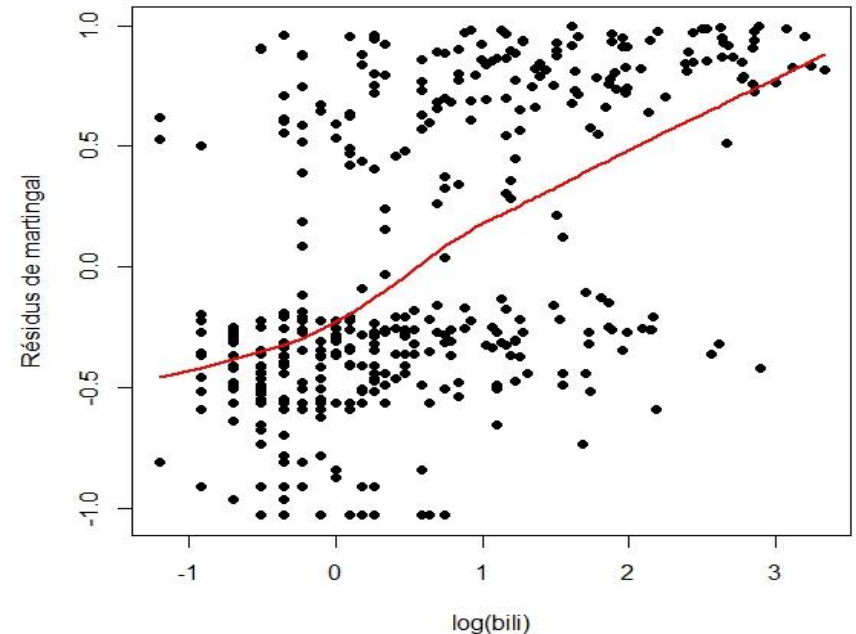
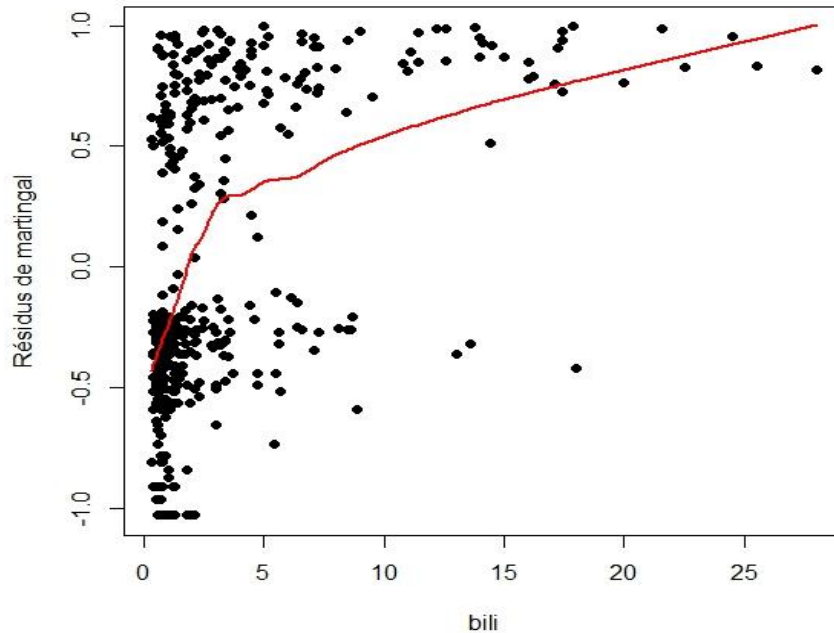
Exemple cirrhose biliaire primitive : pbc



Hypothèse de Log-Linéarité

Résidus de Martingale

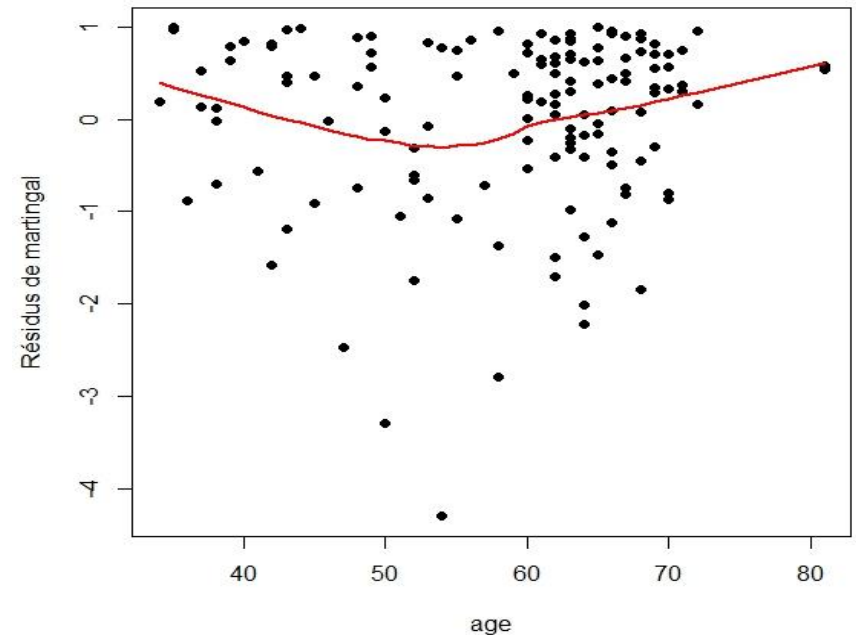
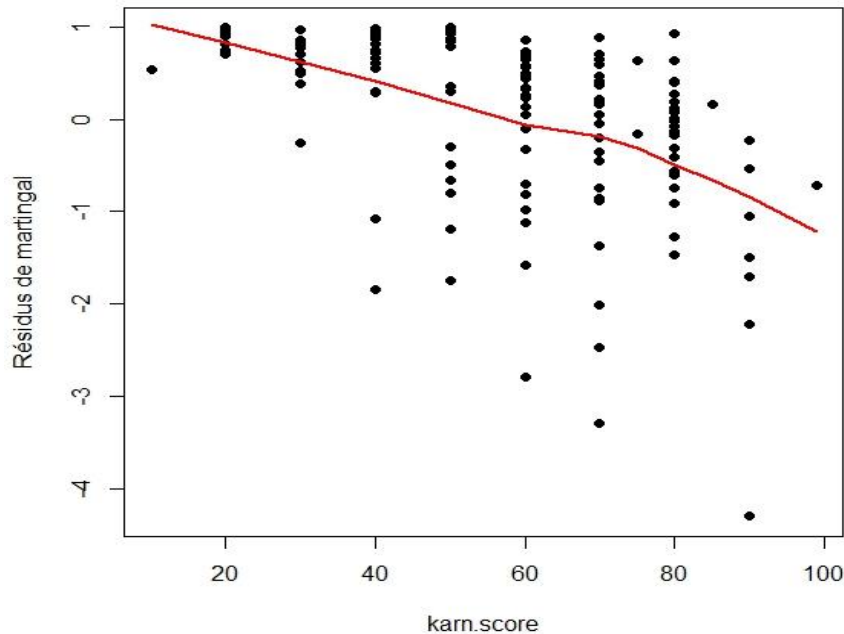
Exemple cirrhose biliaire primitive : pbc



Hypothèse de Log-Linéarité

Résidus de Martingale

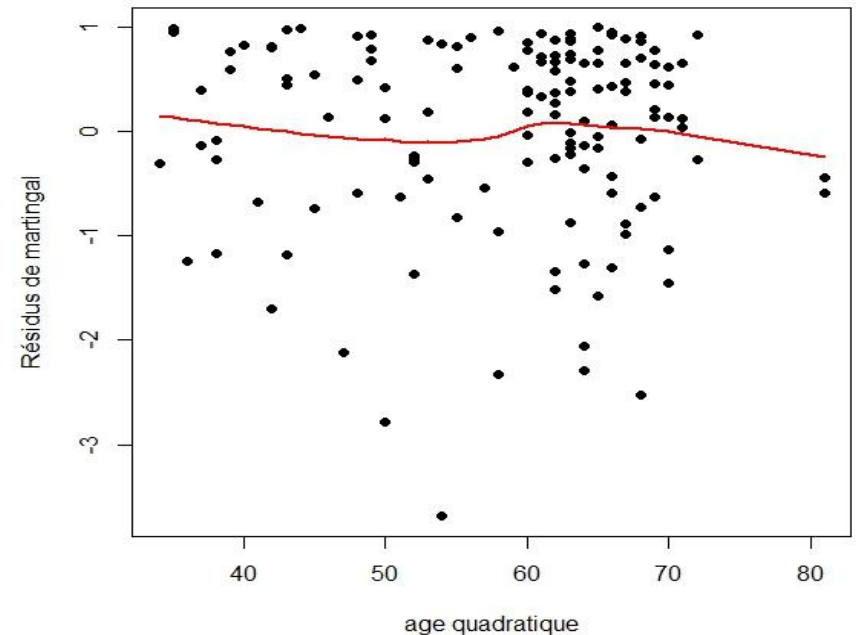
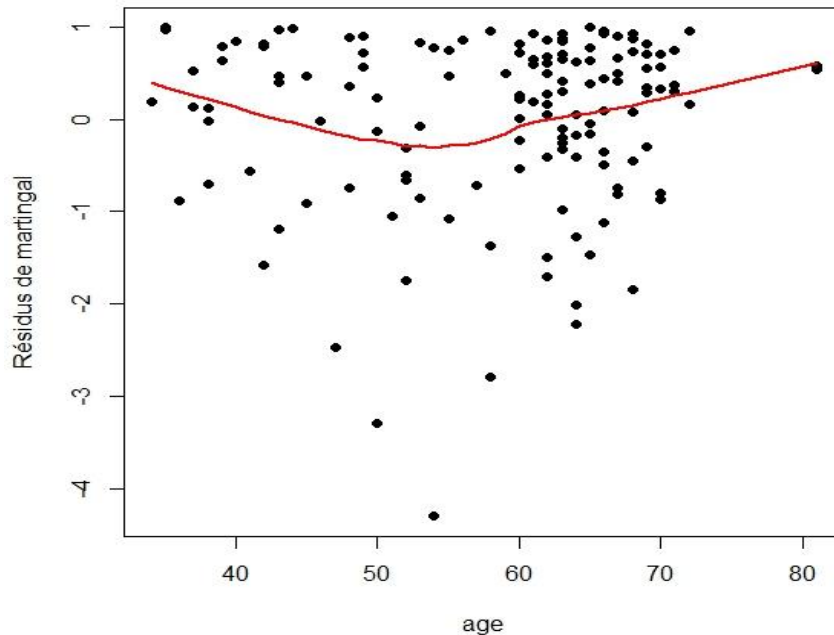
Exemple cancer du poumon



Hypothèse de Log-Linéarité

Résidus de Martingale

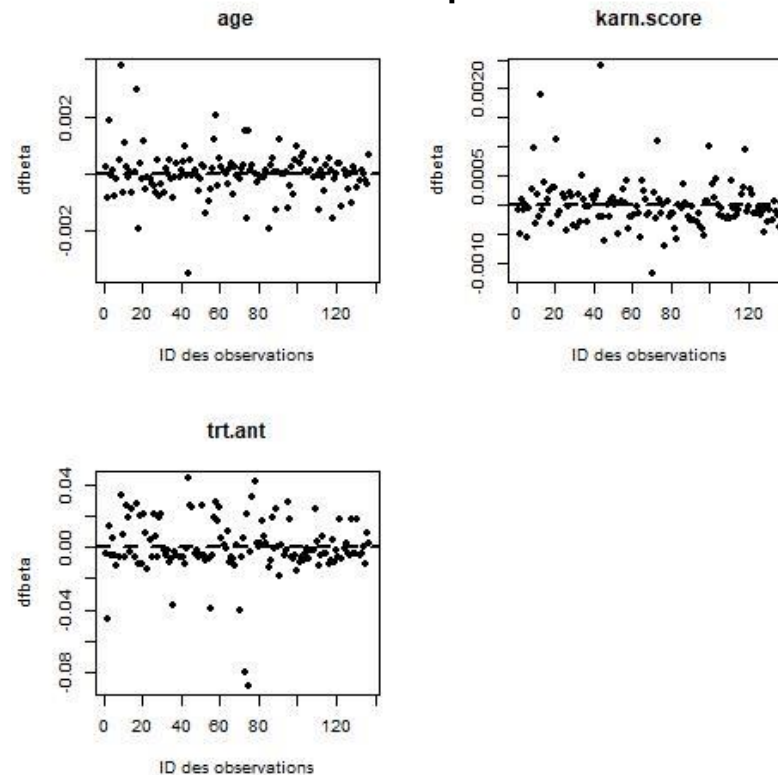
Exemple cancer du poumon



Influence des Observations

Outliers

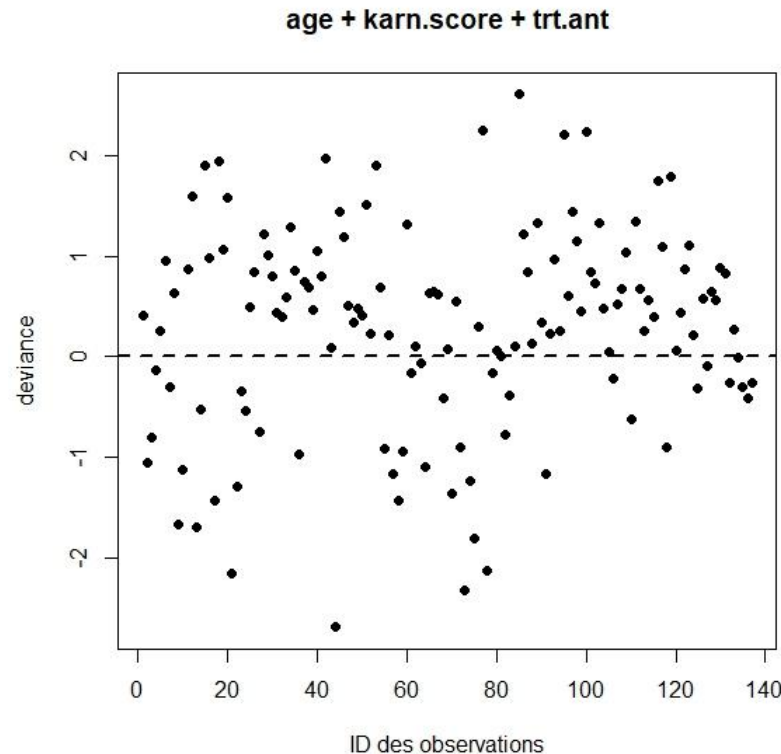
- Graphiques de résidus
 - Valeurs dfbeta
 - Changements estimés des coefficients de régression lorsque l'on supprime à tour de rôle chaque observation



Influence des Observations

Outliers

- Graphiques de résidus
 - Deviance résiduelle
 - Transformation normalisée des résidus de martingale



Extensions du Modèle de Cox

Seules seront abordées

- Modèle de Cox stratifié
- Rapport des taux dépendant du temps (cont.)
- Effet non-linéaire (cont.)
- Covariable dépendante du temps

Modèle de Cox Stratifié

Modèle

- Soit une covariable catégorielle ayant K classes
 - Sujets répartis de manière disjointe

$$\lambda(t|\mathbf{z}) = \lambda_{0k}(t) \exp(\beta^t \mathbf{z}), k = 1, \dots, K$$

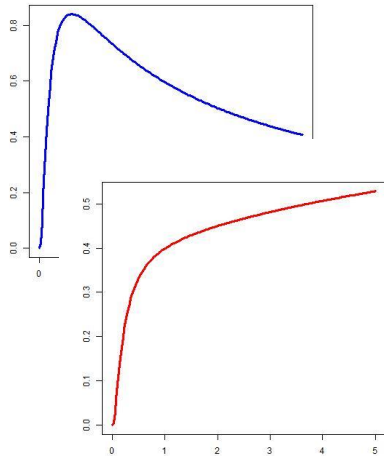
- Effet des covariables identique dans chaque strates
- Hypothèses classiques du modèle de Cox non relâchées pour \mathbf{z}
- Taux de mortalité de base différent dans chaque strates

Modèle de Cox Stratifié

Modèle

- Exemple

- Covariable de stratification : *sexe* (codé en 1 si féminin et 2 sinon)
- Covariables d'ajustement : *age*, *trt*



$$\lambda(t|\mathbf{z}, \text{féminin}) = \lambda_{01}(t) \exp(\beta_{age}age + \beta_{trt}trt)$$

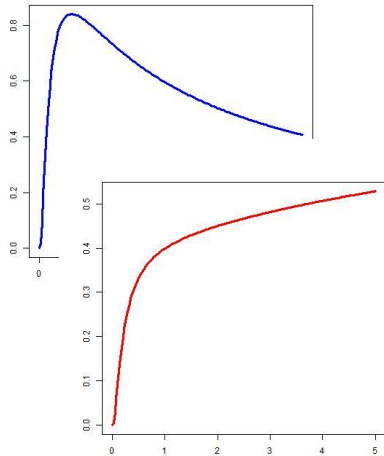
$$\lambda(t|\mathbf{z}, \text{masculin}) = \lambda_{02}(t) \exp(\beta_{age}age + \beta_{trt}trt)$$

Modèle de Cox Stratifié

Modèle

- Exemple

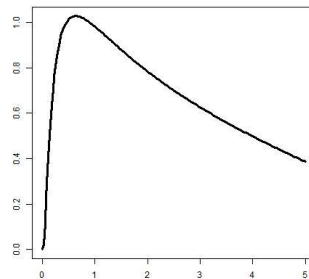
- Covariable de stratification : *sexe* (codé en 1 si féminin et 2 sinon)
- Covariables d'ajustement : *age*, *trt*



$$\lambda(t|\mathbf{z}, \text{féminin}) = \lambda_{01}(t) \exp(\beta_{age}age + \beta_{trt}trt)$$

$$\lambda(t|\mathbf{z}, \text{masculin}) = \lambda_{02}(t) \exp(\beta_{age}age + \beta_{trt}trt)$$

≠ Modèle non stratifié avec comme covariables d'ajustement : *age*, *trt*, *sexe*



$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(\beta_{age}age + \beta_{trt}trt + \beta_{sexe}sexe)$$

Modèle de Cox Stratifié

Estimation - Tests

- On considère les individus à risque dans chaque strate séparément
 - A l'instant t_i ($i = 1, \dots, n$ décès) il y a \mathcal{R}_{ik} individus encore à risque dans chaque strate
 - Vraisemblance partielle de chaque strate
- Estimation des paramètres
 - Vraisemblance totale = somme des vraisemblances de chaque strate
- Tests des paramètres
 - Rapport de vraisemblance
 - Wald
 - Score

Modèle de Cox Stratifié

Intérêts

- Prise en compte d'une hétérogénéité dans les données
- Covariable catégorielle avec écart à l'hypothèse des taux proportionnels
- Ajustement naturel sur une covariable catégorielle de confusion

Modèle de Cox Stratifié

Inconvénients

- Pas d'estimation directe de l'importance de l'effet de la covariable de stratification
- La précision dans l'estimation des coefficients et la puissance de l'analyse diminue avec le nombre de strate
- Les covariables agissent de la même manière dans chaque strate
 - Hypothèse relâchée en créant une interaction entre la covariable de stratification et celle(s) d'intérêt(s)
 - Testée simplement par un test du rapport de vraisemblance avec le nombre de degrés de liberté appropriés

Modèle de Cox Stratifié

Exemple : pbc2.id (package JM)

```
fit <- coxph(Surv(years, status2) ~ age +  
             log(serBilir) + strata(edema), data = pbc2.id)
```

	coef	se(coef)	p
age	0.05	0.01	<0.001
log(serBilir)	0.98	0.10	<0.001

D'après : <http://www.drizopoulos.com/courses/EMC/EP03.pdf>

Modèle de Cox Stratifié

Exemple : pbc2.id (package JM)

```
fit.ia <- coxph(Surv(years, status2) ~ age +  
               log(serBilir) * strata(edema), data = pbc2.id)
```

	coef	se(coef)	p
age	0.04	0.01	<0.001
log(serBilir)	1.08	0.11	<0.001
log(serBilir):edema no diuretics	-0.02	0.93	0.62
log(serBilir):edema despite diuretics	-0.75	0.26	0.04

	p
1 - pchisq(2*(fit.ia\$loglik[2] - fit\$loglik[2]), df=2))	0.0238
anova(fit, fit.ia)	0.0238

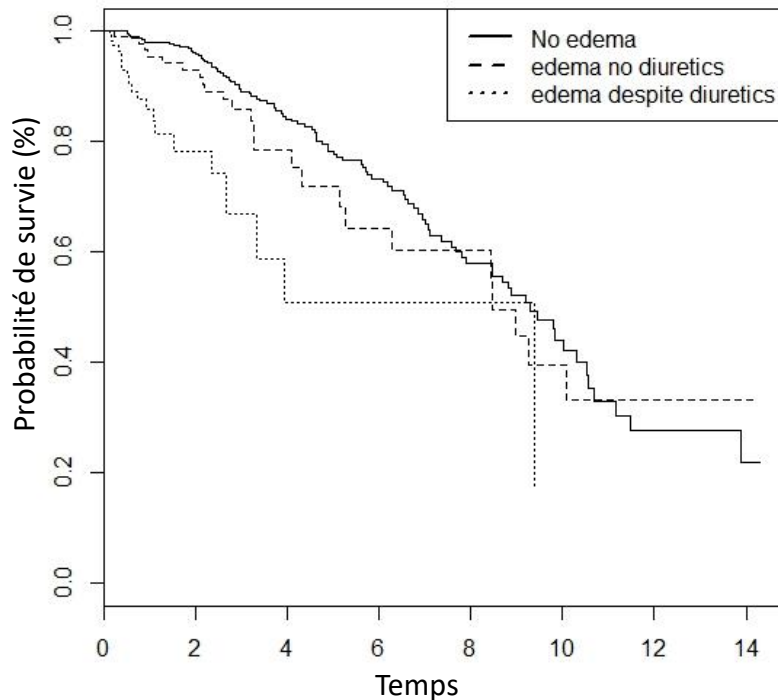
D'après : <http://www.drizopoulos.com/courses/EMC/EP03.pdf>

Modèle de Cox Stratifié

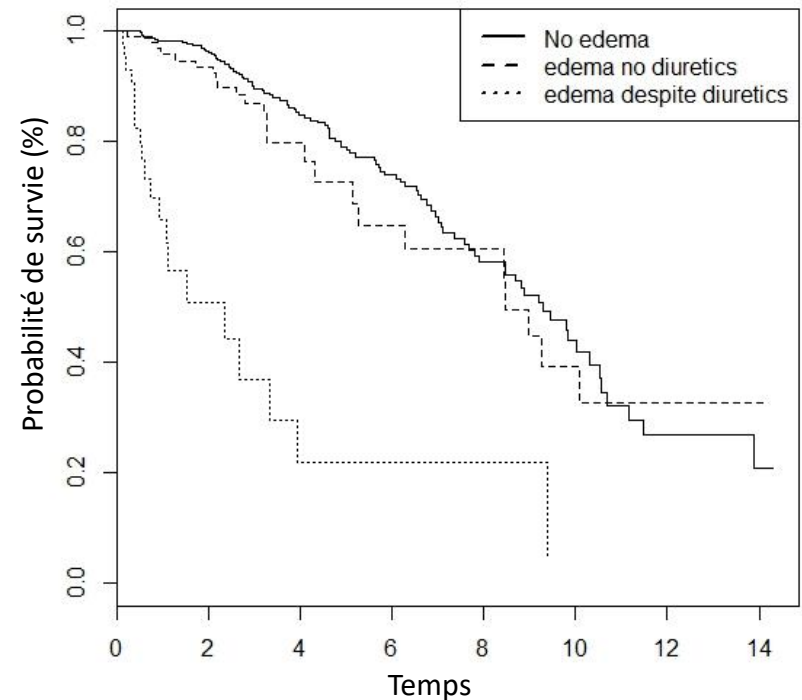
Exemple : `pbcr2.id` (package JM)

pour un profil individuel `age=60` et `serBilir=2`

Sans inter-action



Avec inter-action



D'après : <http://www.drizopoulos.com/courses/EMC/EP03.pdf>

Rapport des Taux Dépendant du Temps

- Causes
 - Omission d'une covariable importante
 - Forme fonctionnelle incorrecte d'une covariable
 - Effet de la covariable évoluant réellement au cours du temps
 - ...
- Modélisation permettant
 - fonctions du taux ne sont plus proportionnelles
 - leur rapport (RR) n'est plus constant et dépend de t

Rapport des Taux Dépendant du Temps

- Modèles de la forme

$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(\beta(t)\mathbf{z})$$

- Différentes possibilités
 - Interaction entre une covariable et une fonction de t
 - Modèle à taux proportionnels par intervalles
 - Modélisation de l'effet à l'aide d'une fonction flexible (fonction spline,...)
 - Utiliser un autre type de modèle (temps de vie accéléré, additif)
- Comparaison des modèles

Interaction Covariable * Temps

$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(\beta^t \mathbf{z} + \gamma_2 f(t) z_2)$$

- La fonction $f(t)$ est définie *a priori*
 - $f(t) = t$: interaction linéaire
 - $f(t) = t^2$: interaction quadratique
 - $f(t) = \ln(t)$
 - ...
- On teste *a posteriori* celle qui s'adapte le mieux aux données observées

Interaction Covariable * Temps

Exemple avec la fonction $\ln(t) - z_2$ binaire

$$\begin{aligned}\lambda(t|\mathbf{z}) &= \lambda_0(t) \exp(\beta_2 z_2 + \gamma_2 \ln(t) z_2) \\ &= \lambda_0(t) [\exp(\beta_2 z_2) t^{\gamma_2 z_2}]\end{aligned}$$

- Pour 2 individus i, j ayant pour covariables z_{2i} et z_{2j}

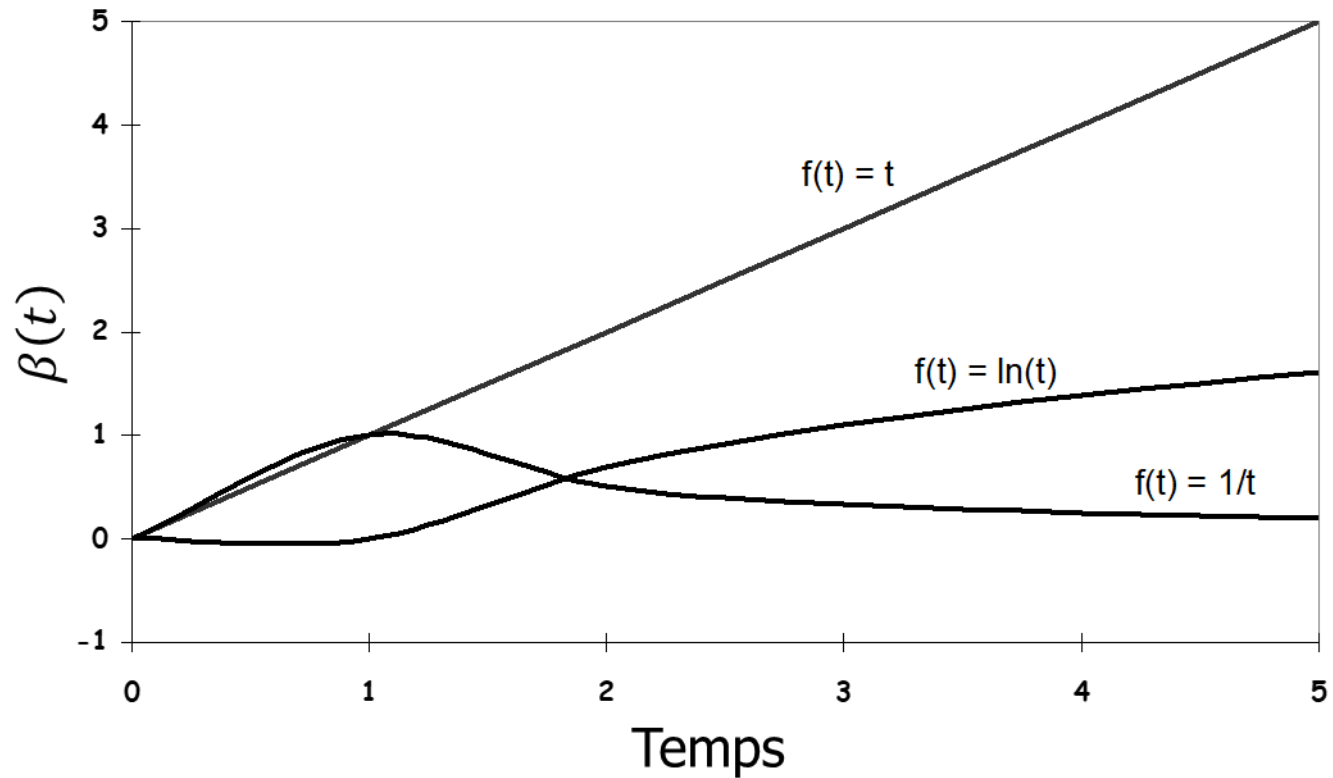
$$\frac{\lambda(t|z_{2i})}{\lambda(t|z_{2j})} = \exp\left(\beta_2(z_{2i} - z_{2j})\right) t^{\gamma_2(z_{2i} - z_{2j})}$$

$$\frac{\lambda(t|z_{2i} = 1)}{\lambda(t|z_{2j} = 0)} = \exp(\beta_2) t^{\gamma_2}$$

- Hypothèse nulle $H_0: \gamma_2 = 0$ (proportionnalité respectée)

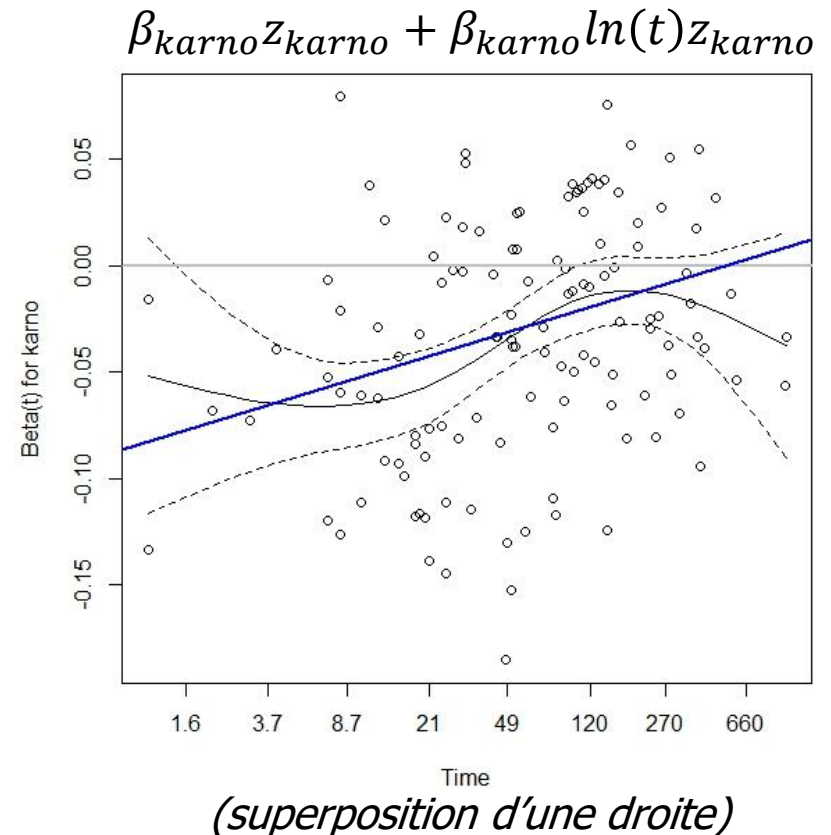
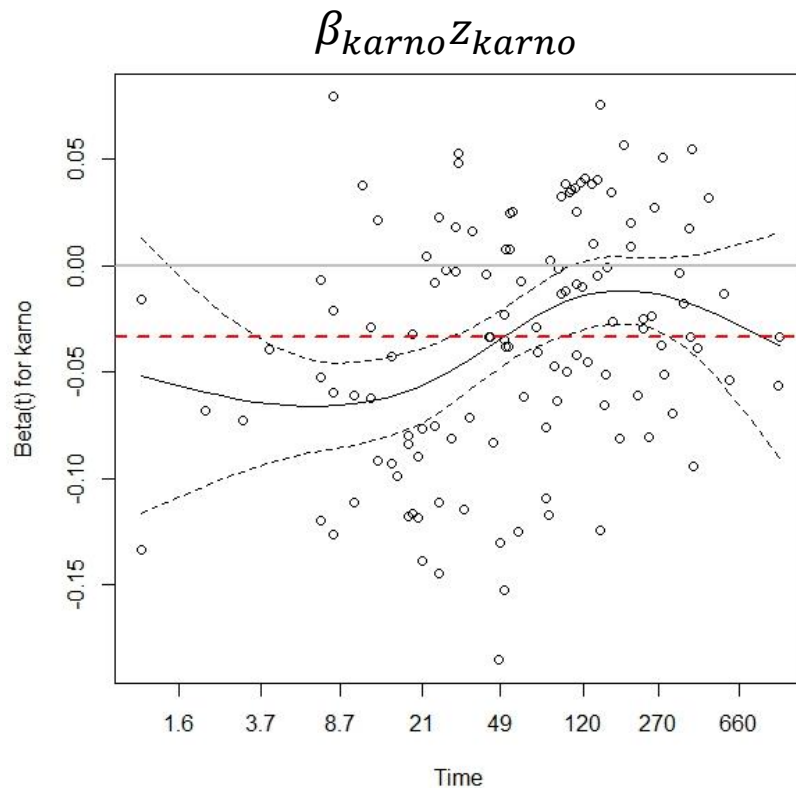
Interaction Covariable * Temps

Exemples de formes



Interaction Covariable * Temps

Exemple : veteran (package survival)



D'après : <https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf>

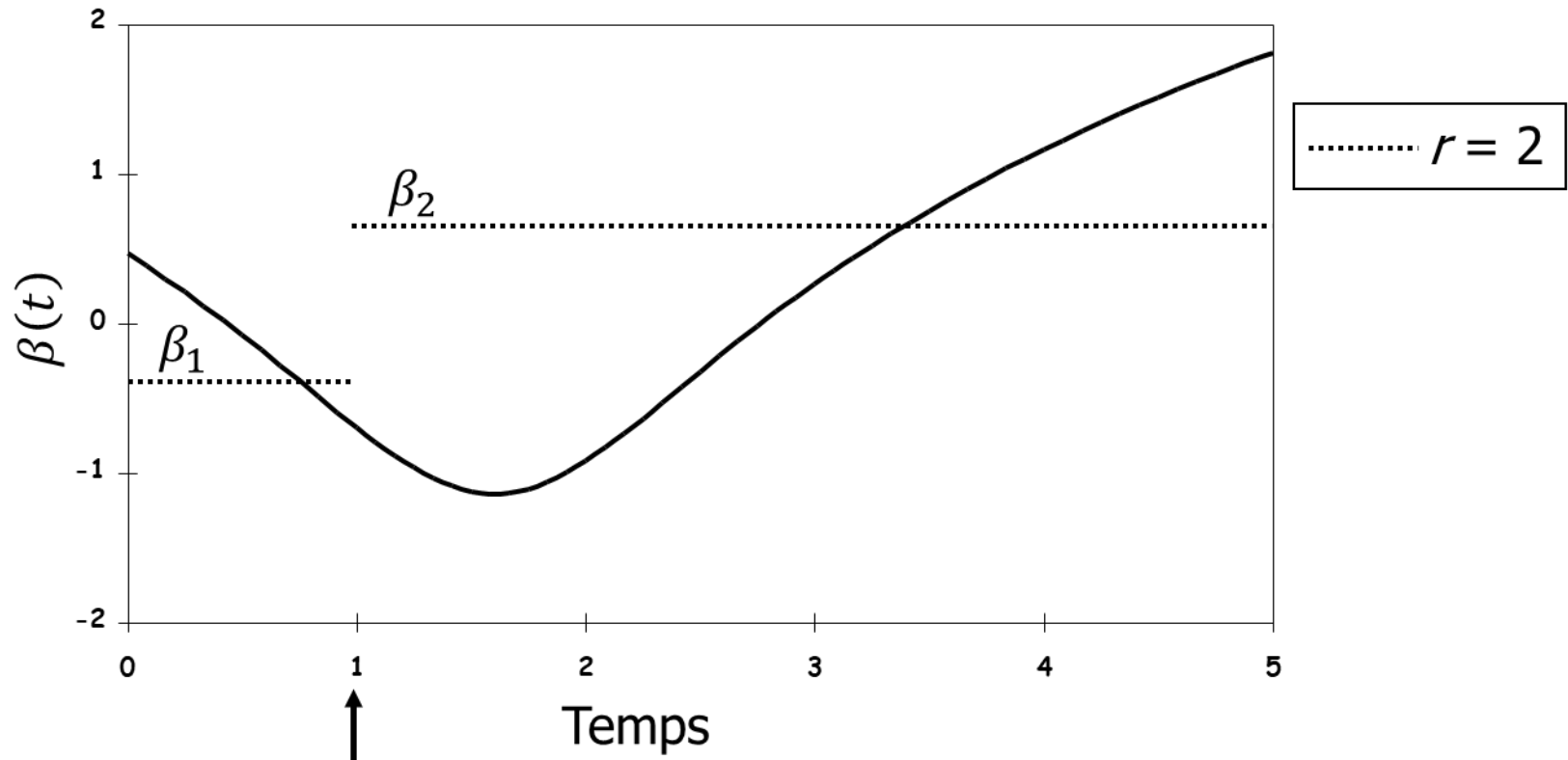
Découpage du Temps de Suivi

$$\lambda(t|\mathbf{z}) = \lambda_0(t) \sum_{k=1}^r \exp(\beta_k^t \mathbf{z}) I_k(t)$$

- Découpage *a priori* du temps de suivi en r intervalles de temps (égaux ou non)
- Hypothèses
 - fonctions du taux proportionnelles sur chacun des intervalles
 - Peuvent varier d'un intervalle à l'autre

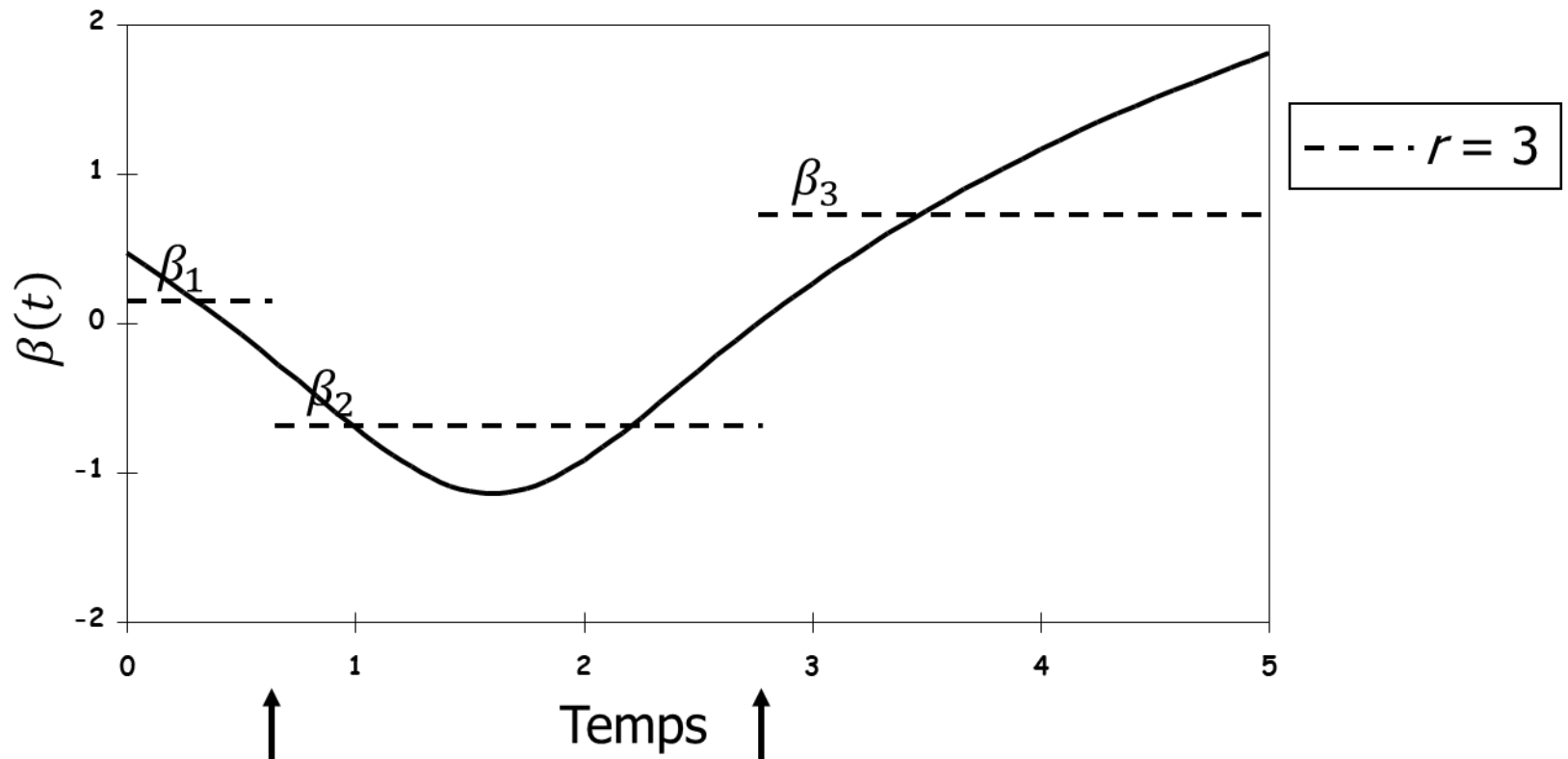
Découpage du Temps de Suivi

$r = 2$



Découpage du Temps de Suivi

$r = 3$



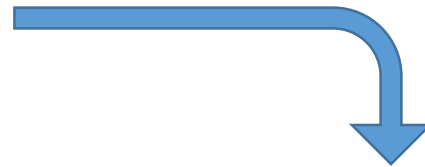
Découpage du Temps de Suivi

Exemple : veteran (package survival)

- Découpage du temps avec la fonction `survSplit`
 - 90 et 180 jours

data.frame initial

	id	time	status	age	karno
1	1	72	1	69	60
2	2	111	1	64	70
3	3	228	0	38	60



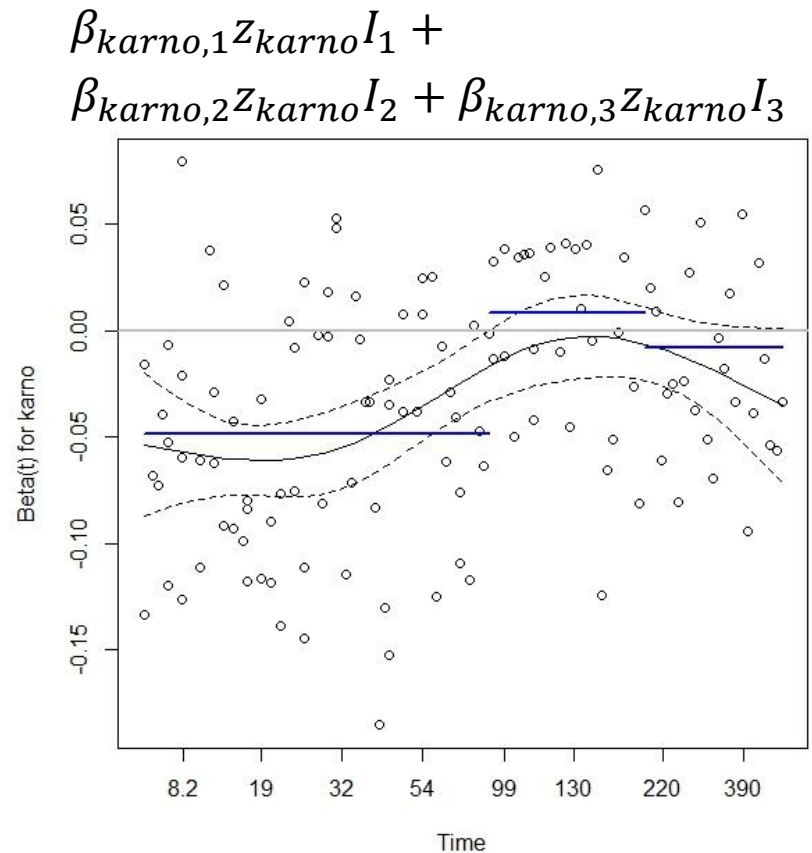
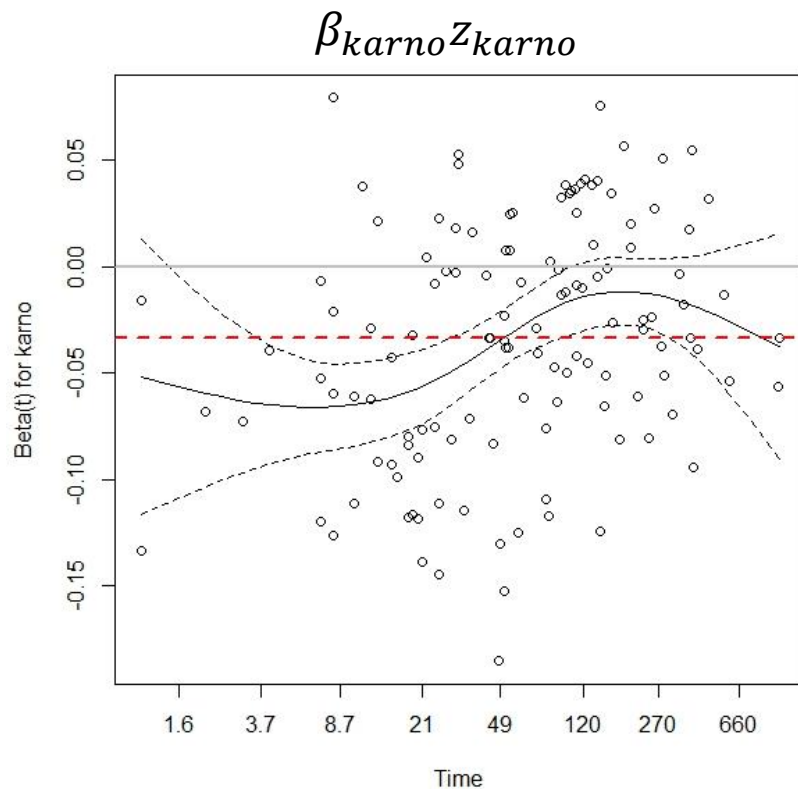
data.frame splitté

	id	tstart	time	status	tgroup	age	karno
1	1	0	72	1	1	69	60
2	2	0	90	0	1	64	70
3	2	90	111	1	2	64	70
4	3	0	90	0	1	38	60
5	3	90	180	0	2	38	60
6	3	180	228	0	3	38	60

D'après : <https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf>

Découpage du Temps de Suivi

Exemple : veteran (package survival)



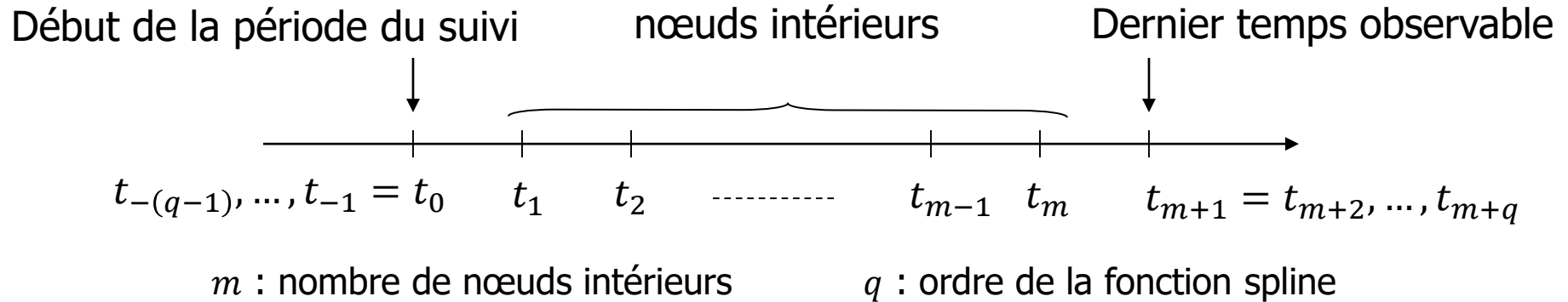
D'après : <https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf>

Fonction Flexible – Splines

- Modélisation souple et flexible des variations
- Estimation non plus basée sur des *a priori* mais sur la modélisation des données observées
- Fonctions polynomiales par morceaux d'ordre q (polynôme de degré $q - 1$)
 - Flexibilité augmente avec l'ordre
- Point de jonction de deux polynômes contiguës au niveau de points appelés nœuds
 - Nombre de nœuds contrôle la flexibilité
- Chaque morceau (polynôme) relie le morceau suivant de manière lisse et homogène

Fonction Flexible

Exemple : B-Splines



$$B_{j,q}(t) = \frac{t - t_j}{t_{j+q-1} - t_j} B_{j,q-1}(t) + \frac{t_{j+q} - t}{t_{j+q} - t_{j+1}} B_{j+1,q-1}(t)$$

$$\begin{aligned}
 & j = -(q-1), \dots, m \\
 & B_{j,1}(t) = \begin{cases} 1 & \text{si } t \in [t_j, t_{j+1}[\\ 0 & \text{sinon} \end{cases}
 \end{aligned}$$

Fonction Flexible

Exemple : B-Splines cubique, 2 nœuds int.

$$B_{j,q}(t) = \frac{t - t_j}{t_{j+q-1} - t_j} B_{j,q-1}(t) + \frac{t_{j+q} - t}{t_{j+q} - t_{j+1}} B_{j+1,q-1}(t)$$

$m = 2$: nombre de nœuds intérieurs $q = 3$: ordre de la fonction spline

Avec $j = -2$ et $q = 3$:

$$B_{-2,3}(t) = \frac{t - t_{-2}}{t_{-2+3-1} - t_{-2}} B_{-2,3-1}(t) + \frac{t_{-2+3} - t}{t_{-2+3} - t_{-2+1}} B_{-2+1,3-1}(t)$$

$$B_{-2,2}(t) = \frac{t - t_{-2}}{t_{-1} - t_{-2}} B_{-2,1}(t) + \frac{t_0 - t}{t_0 - t_{-1}} B_{-1,1}(t)$$

$$B_{-1,2}(t) = \frac{t - t_{-1}}{t_0 - t_{-1}} B_{-1,1}(t) + \frac{t_1 - t}{t_1 - t_0} B_{0,1}(t)$$

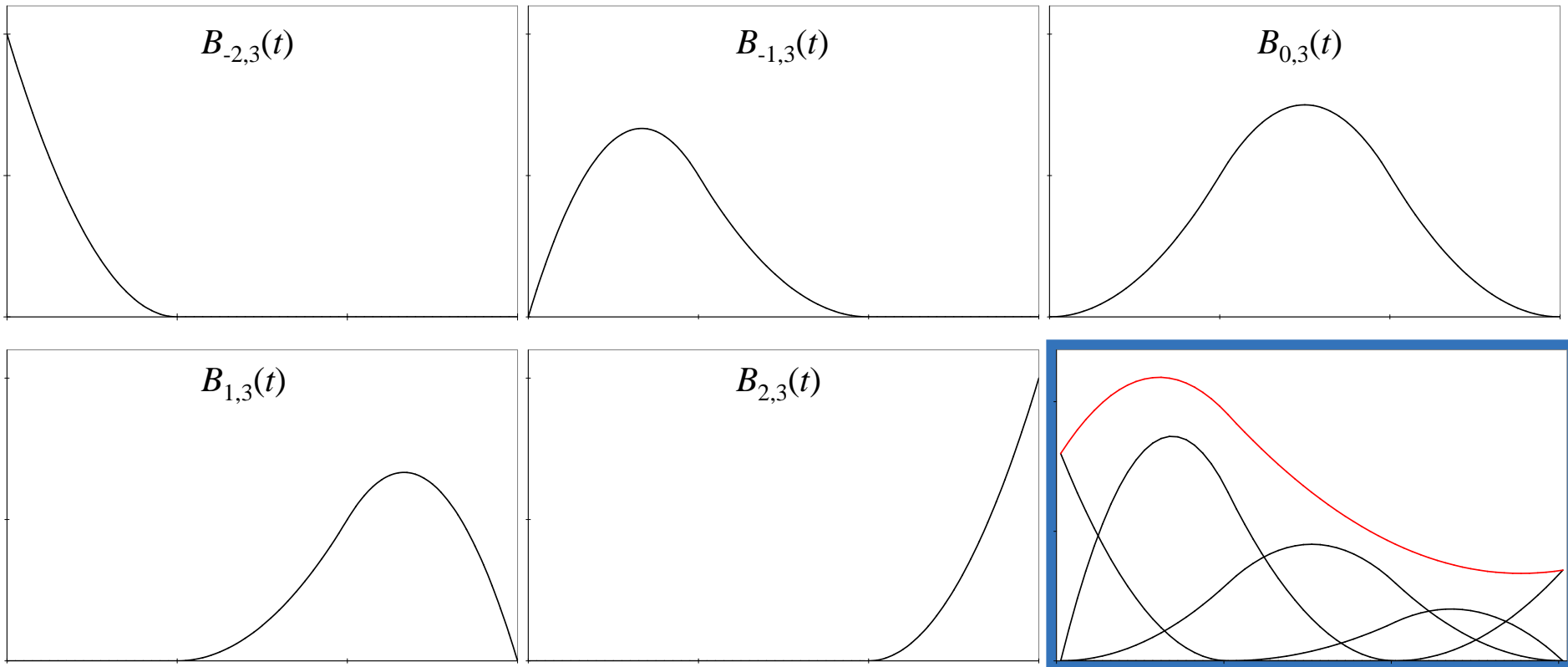
$$B_{-2,3}(t) = \frac{t_1 - t}{t_1 - t_{-1}} \times \frac{t_1 - t}{t_1 - t_0} B_{0,1}(t)$$

$$B_{-2,3}(t) = 1 - \frac{2}{t_1} t + \frac{1}{t_1^2} t^2 \text{ si } t \in [t_0, t_1[, 0 \text{ sinon}$$



Fonction Flexible

Exemple : B-Splines cubique, 2 nœuds int.



$$\beta(t) = \sum_{j=-(q-1)}^m \alpha_j B_{j,q}(t), t \in (t_0, t_{m+1})$$

Fonction Flexible

Exemple : B-Splines cubique, 2 nœuds int.

$$\beta_i(t) = \sum_{j=-(q-1)}^m \alpha_{i,j} B_{j,q}(t) \quad i = 1, \dots, p$$

- Fonctions B-splines quadratiques (ordre $q = 3$)
- Deux nœuds intérieurs ($m = 2$)
 - Localisation basée sur les quantiles de la distribution des temps de décès observés, connaissances épidémiologiques/clinique,...
- 5 fonctions de bases pour l'espace de régression des splines
 - $5p$ paramètres à estimer par covariables

Effet Non-Linéaire

- Modèles de la forme

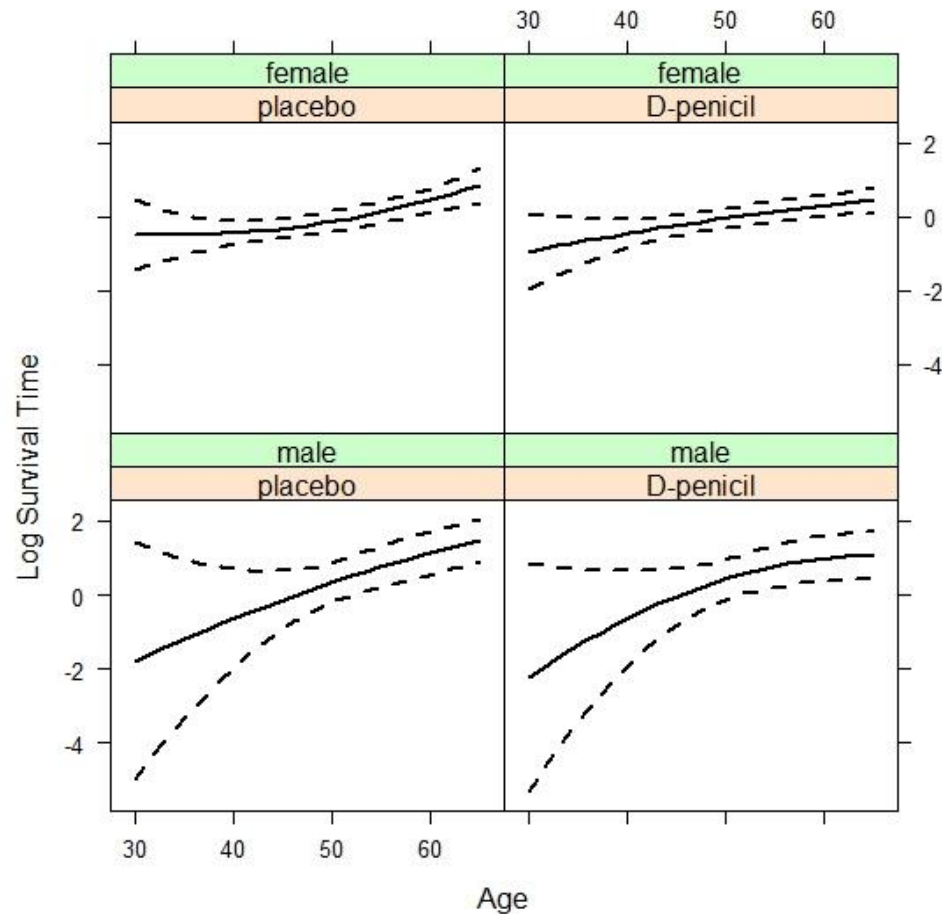
$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(f(\beta)\mathbf{z})$$

- Différentes possibilités
 - Forme paramétriques (logarithmique, quadratique,...)
 - Fonctions flexibles non-paramétriques (splines)
 - ...
- Comparaison des modèles

Effet Non-Linéaire

Quadratique

```
coxph(Surv(years, status2) ~ (drug + sex) * (age + I(age^2)))
```

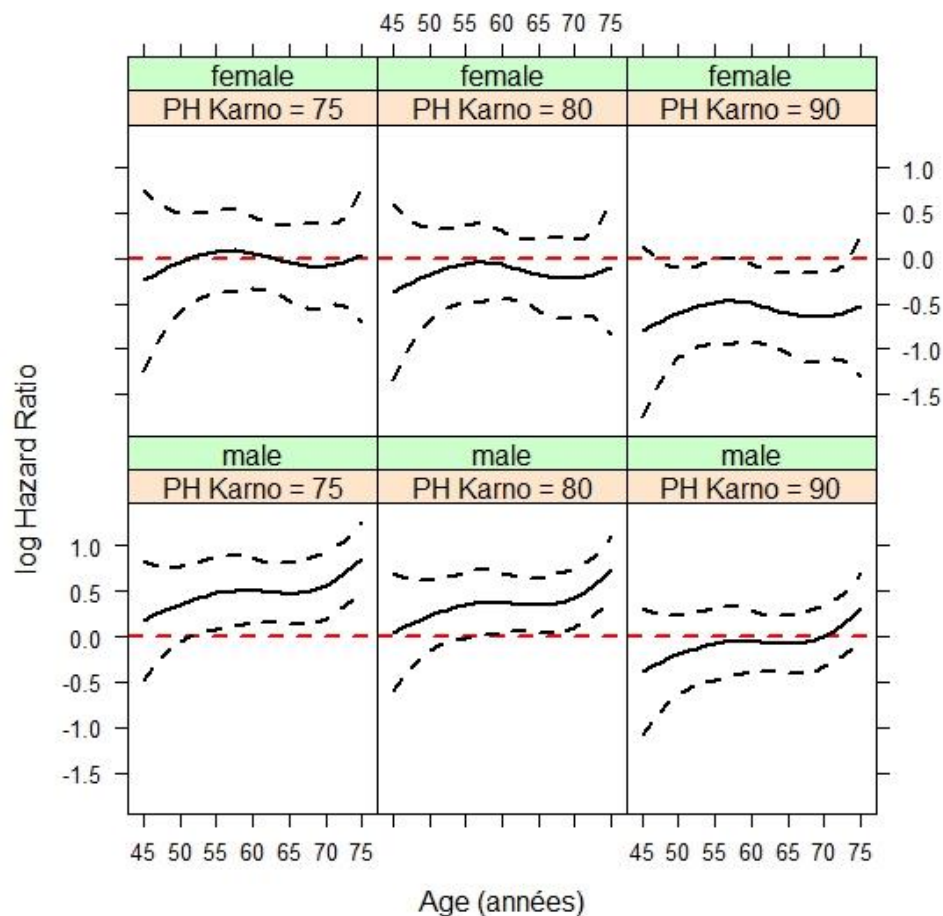


D'après : http://www.drizopoulos.com/courses/emc/ep03_%20survival%20analysis%20in%20r%20companion

Effet Non-Linéaire

Spline cubique naturelle

`coxph(Surv(time, status) ~ ns(age, 3) * sex + ns(ph.karno, 3)`



D'après : http://www.drizopoulos.com/courses/emc/ep03_%20survival%20analysis%20in%20r%20companion

Covariable Dépendante du Temps

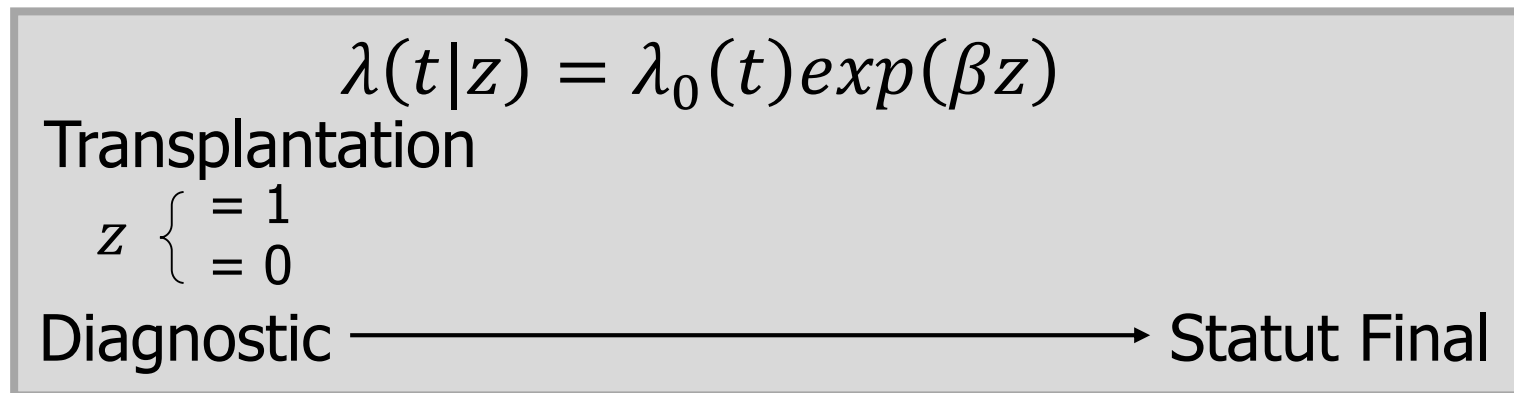
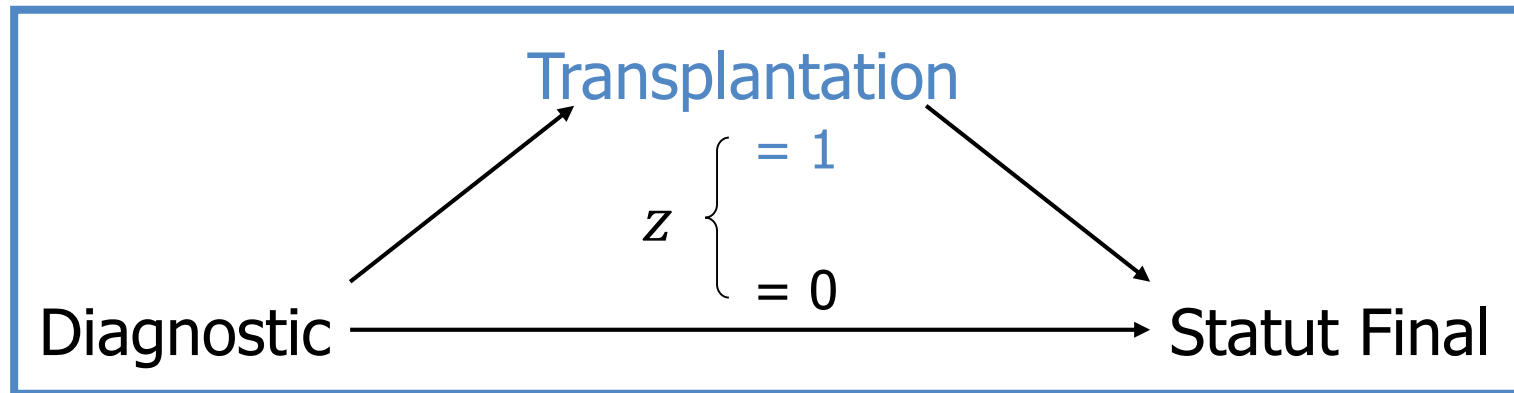
- Modèles de la forme

$$\lambda(t|\mathbf{z}) = \lambda_0(t) \exp(\beta \mathbf{z}(t))$$

Covariable Dépendante du Temps

- Modèles de la forme

$$\lambda(t|\mathbf{z}) = \lambda_0(t)\exp(\beta\mathbf{z}(t))$$



Covariable Dépendante du Temps

Processus de Comptage (Principe)

- Considère chaque individu comme étant une observation d'un lent processus de Poisson
- Les censures ne sont pas considérées comme étant des données incomplètes (le compteur est toujours à 0)
- Les covariables dépendantes du temps modifient le taux pour les événements à venir, peut dépendre des observations précédentes
- Mise en œuvre informatique simple et attractive

Covariable Dépendante du Temps

Processus de Comptage (Principe)

Remplacer

par

	Temps	Statut	$z(t)$	z_1	z_2
]Start	Stop]	Statut	$z(t)$	z_1	z_2

Exemple :

Id]Start	Stop]	Statut	Transplant	Age<45	Chir. ant.
12	0	26	1	0	1	0
16	0	2	0	0	1	1
16	2	43	0	1	1	1

Covariable Dépendante du Temps

Exemple : jasa (package survival)

- Duplication des données `tmerge`
 - Variable `trt` dépendante du temps

data.frame initial

	subject	accept.dt	tx.date	futime	fustat	age	trt
2	2	1968-01-02	<NA>	5	1	3.84	0
3	3	1968-01-06	1968-01-06	15	1	6.30	1
4	4	1968-03-28	1968-05-02	38	1	-7.74	1



data.frame dupliqué

	subject	tstart	tstop	death	age	trt
2	2	0	5	1	3.84	0
3	3	0	15	1	6.30	1
4	4	0	35	0	-7.74	0
5	4	35	38	1	-7.74	1

D'après : <https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf>

Covariable Dépendante du Temps

Exemple : jasa (package survival)

```
coxph(Surv(futime, fustat) ~ age + trt, data=jasa)
```

	coef	exp(coef)	se(coef)	z	p
age	0.0602	1.0620	0.0153	3.93	8.4e-05
trt	-1.8005	0.1652	0.2722	-6.61	3.8e-11

Likelihood ratio test=44.5 on 2 df, p=2.21e-10
n= 103, number of events= 75

```
coxph(Surv(tstart, tstop, death) ~ age + trt, data=jasa.sstop)
```

	coef	exp(coef)	se(coef)	z	p
age	0.03076	1.03124	0.01450	2.12	0.034
trt	-0.00606	0.99396	0.31175	-0.02	0.984

Likelihood ratio test=5.17 on 2 df, p=0.0754
n= 170, number of events= 75

D'après : <https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf>

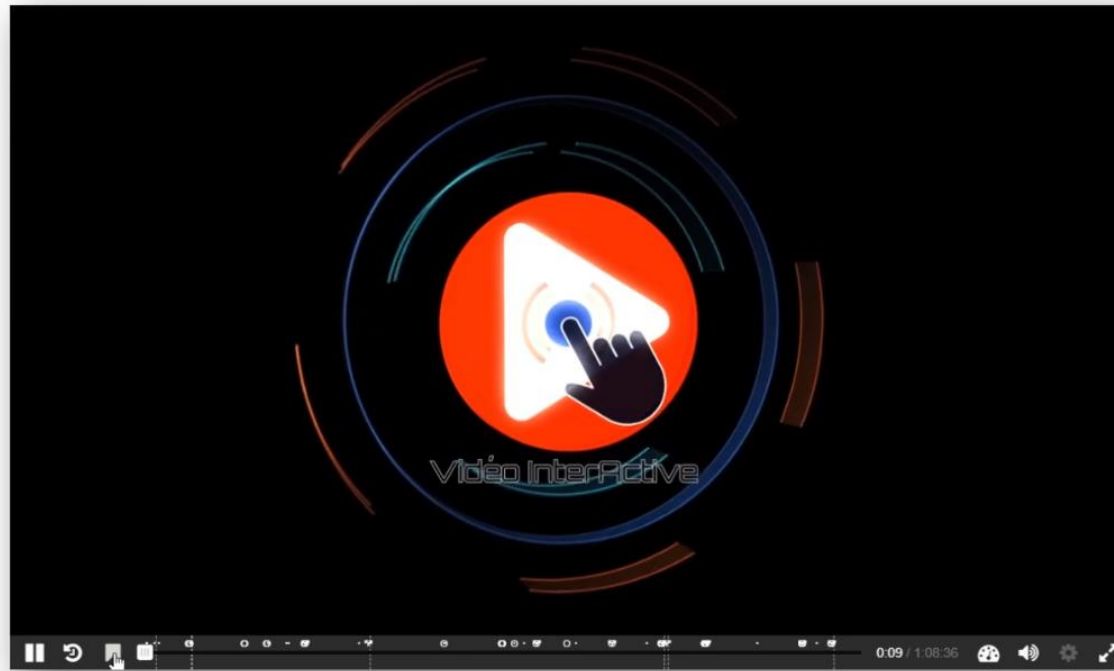
Ressource Pédagogique Complémentaire



Les vidéos interactives du SESSTIM

Avec nos vidéos interactives nous vous proposons un parcours de formation adaptatif et personnalisé, enrichi de différents éléments pédagogiques.

Durant la visualisation, vous pourrez ainsi choisir vous-même votre parcours pédagogique, en choisissant les documents à consulter et les orientations adaptées à votre niveau de connaissance.



<https://sesstim.univ-amu.fr/content/modele-regression-cox-et-extensions>

Références

- Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. Wiley, New York, 1980.
- Therneau T, Grambsch PM. Modeling survival data: Extending the Cox model. New York: Springer-Verlag 2000.
- <https://cran.r-project.org/web/packages/survival/>
- <http://www.drizopoulos.com/#teaching>