Le modèle de Cox dépendant du temps

Yohann.Foucher@univ-nantes.fr

Master 2 - Modélisation en Pharmacologie Clinique et Epidémiologie

UNIVERSITÉ DE NANTES
CENTRE HOSPITALIER UNIVERSITAIRE DE NANTES
itun institut transplantation urologie néphrologie INSERM - UMR 643

Plan

Plan

Définitions

- La fonction de probabilité de densité $f(t)$ représente la limite de la probabilité que l'événement se produise au temps t.

$$f(t) = \lim_{\Delta t \to 0} Pr(t \leq T < t + \Delta t)/\Delta t$$

- La fonction de survie $S(t)$ est la probabilité que l'événement se produise après le temps t.

$$S(t) = Pr(T > t) = \int_t^{\infty} f(u)du$$

- La fonction de répartition $F(t)$ est la probabilité que l'événement se produise après le temps t.

$$F(t) = Pr(T \leq t) = 1 - S(t) = \int_0^t f(u)du$$

Définitions

- The fonction de risque $h(t)$ représente la limite de la probabilité que l'événement se produise au temps t, sachant qu'aucun événement ne s'est produit jusqu'au temps t.

$$\lambda(t) = \lim_{\Delta t \to 0} Pr(t \leq T < t + \Delta t | T \geq t)/\Delta t = f(t)/S(t)$$

- La fonction de risque cumulée $H(t)$.

$$\Lambda(t) = \int_0^t \lambda(u) du$$

Définitions

- Formulation générale d'un modèle à risques proportionnels (PH) :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

- Publication en 1972 par Sir David Cox *.
- $\lambda_0(t)$ est la fonction de risque de base (groupe de référence : $x_1 = x_2 = ... = 0$).
- Idée de base du modèle : $\lambda_0(t)$ n'est pas estimée.
- Modèle semi-paramétrique :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

  - Partie non-paramétrique (pas d'estimation)
  - Partie paramétrique (fonction Exponentielle avec prédicteur linéaire)

*. Cox, DR. Regression models and life-tables. Journal of the Royal Statistical Society : Series B. 34 : 187-220. 1972

Définitions

- Formulation générale d'un modèle à risques proportionnels (PH) :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t) exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

- Publication en 1972 par Sir David Cox*.
- $\lambda_0(t)$ est la fonction de risque de base (groupe de référence : $x_1 = x_2 = ... = 0$).
- Idée de base du modèle : $\lambda_0(t)$ n'est pas estimée.
- Modèle semi-paramétrique :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t) exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

  - Partie non-paramétrique (pas d'estimation)
  - Partie paramétrique (fonction Exponentielle avec prédicteur linéaire)

*. Cox, DR. Regression models and life-tables. Journal of the Royal Statistical Society : Series B. 34 : 187-220. 1972

Définitions

- Formulation générale d'un modèle à risques proportionnels (PH) :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t) exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

- Publication en 1972 par Sir David Cox *.
- $\lambda_0(t)$ est la fonction de risque de base (groupe de référence : $x_1 = x_2 = ... = 0$).
- Idée de base du modèle : $\lambda_0(t)$ n'est pas estimée.
- Modèle semi-paramétrique :

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t) exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

  - Partie non-paramétrique (pas d'estimation)
  - Partie paramétrique (fonction Exponentielle avec prédicteur linéaire)

*. Cox, DR. Regression models and life-tables. Journal of the Royal Statistical Society : Series B. 34 : 187-220. 1972

Estimation des paramètres de régression

- Supposons N individus étudiés ($j = 1, 2, ..., N$)
- Pour le $j$ème individu, posons t_j le dernier temps de suivi avec $\delta_j = 1$ si l'événement est observé et $\delta_j = 0$ sinon (censure).
- On peut démontrer que la probabilité que l'individu j subisse l'événement au temps t_j sachant qu'il est à risque est :

$$\begin{aligned}
\ell_j &= \frac{\lambda(t_j|x_{1j}, ..., x_{pj})}{\sum_{i:t_i \geq t_j} \lambda(t_i|x_{1i}, ..., x_{pi})} \\
&= \frac{\lambda_0(t_j) exp(\sum_{k=1}^{p} \beta_k x_{kj})}{\sum_{i:t_i \geq t_j} \lambda_0(t_i) exp(\sum_{k=1}^{p} \beta_k x_{ki})} \\
&= \frac{exp(\sum_{k=1}^{p} \beta_k x_{kj})}{\sum_{i:t_i \geq t_j} exp(\sum_{k=1}^{p} \beta_k x_{ki})}
\end{aligned}$$

- ℓ_j est indépendant de $\lambda_0(t)$.

Estimation des paramètres de régression

- Supposons que tous les individus sont indépendants, le produit des probabilités individuelles précédentes représente la probabilité d'observer l'échantillon. La vraisemblance partielle s'écrit donc :

$$\wp\ell = \prod_{j=1}^{N} \left\{ \frac{exp(x_{1j}, ..., x_{pj})}{\sum_{i:t_i \geq t_j} exp(x_{1i}, ..., x_{pi})} \right\}^{\delta_j}$$

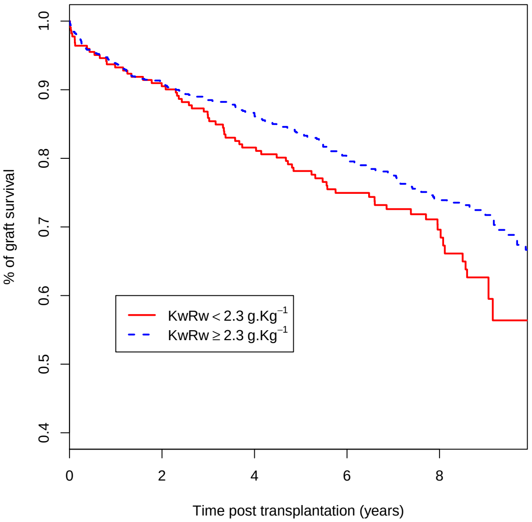
- Les estimations des paramètres de régression β_k $(k = 1, ..., p)$ correspondent aux valeurs qui maximisent la vraisemblance partielle $\wp\ell$.

La proportionnalité des risques

$$\lambda(t|X_1 = x_1, ..., X_p = x_p) = \lambda_0(t) exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

- β_k $(k = 1, ..., p)$ ne dépendent pas du temps
- $RR_k = exp(\beta_k)\ \forall t$
- Les risques relatifs sont constants au cours du temps
- La réalité clinique est souvent différente
- Exemple en transplantation rénale :
  - Variable à expliquer : survie du patient/greffon
  - Variable explicative :

$$KwRw = \frac{\text{Poids du rein}}{\text{Poids du receveur}}$$


% of graft survival
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0
2
4
6
8
Time post transplantation (years)
KwRw < 2.3 g.Kg^{-1}
KwRw ≥ 2.3 g.Kg^{-1}

La proportionnalité des risques

Tracer $log(-log(S(t)))$ en fonction de t dans chaque groupe. La proportionnalité des risques est respectée si les courbes sont parallèles.

$$\lambda(t|x_1) = \lambda_0(t)exp(\beta_1 x_1)$$

$$\Longleftrightarrow$$

$$\begin{aligned}\Lambda(t|x_1) &= \int_0^t \lambda_0(u)exp(\beta_1 x_1)du \\ &= exp(\beta_1 x_1)\int_0^t \lambda_0(u)du \\ &= \Lambda_0(t)exp(\beta_1 x_1)\end{aligned}$$

$$\Longleftrightarrow$$

$$\begin{aligned}S(t|x_1) &= exp(-\Lambda(t|x_1)) \\ &= exp(-\Lambda_0(t)exp(\beta_1 x_1)) \\ &= S_0(t)^{exp(\beta_1 x_1)}\end{aligned}$$

La proportionnalité des risques

Donc, si l'hypothèse est respectée, nous avons :

$$log(S(t|x_1)) = exp(\beta_1 x_1) log(S_0(t))$$

$$\Longleftrightarrow$$

$$-log(S(t|x_1)) = exp(\beta_1 x_1)\Lambda_0(t)$$

$$\Longleftrightarrow$$

$$log(-log(S(t|x_1))) = \beta_1 x_1 + \Lambda_0(t)$$

Exemple : si $x_1 = 1$ ou 0 :

$$log(-log(S(t|x_1 = 1))) - log(-log(S(t|x_1 = 0))) = \beta_1$$

Cirrhose hépatique

- Données issues d'un essai de la Mayo Clinic
- $N = 418$ patients inclus entre le 1974 et 1984.
- Maladie auto-immune avec le décès possible du patient.
  - $N = 161$ décès observés (status $= 2$)
  - $N = 21$ transplantations de foie (status $= 1$)
  - $N = 236$ observations censurées (status $= 0$)
- Covariables à la baseline : bilirubine, albumine, âge, oedème, temps prothrombine, etc.
- Données déjà disponibles sous **R**.

Objectif : Etudier les facteurs pronostiques de la mortalité.

Cirrhose hépatique


```
> library(survival)
> dim(pbc)

[1] 418  21

> table(pbc$status)

  0   1   2
232  25 161
```

Cirrhose hépatique

Introduction

La valeur est dépendante du temps

L'effet est dépendant du temps

```
> summary(cox.pbc <- coxph(Surv(time, status==2) ~ age + edema + log(bili) +
+ log(protime) + log(albumin), data=pbc))
Call:
coxph(formula = Surv(time, status == 2) ~ age + edema + log(bili) +
    log(protime) + log(albumin), data = pbc)

  n= 416, number of events= 160
   (2 observations deleted due to missingness)

                  coef exp(coef)  se(coef)      z Pr(>|z|)
age           0.039609  1.040404  0.007672  5.163 2.43e-07 ***
edema         0.896311  2.450547  0.271410  3.302 0.000959 ***
log(bili)     0.863551  2.371566  0.082941 10.412  < 2e-16 ***
log(protime)  2.386839 10.879054  0.768509  3.106 0.001898 **
log(albumin) -2.506923  0.081519  0.652916 -3.840 0.000123 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

             exp(coef) exp(-coef) lower .95 upper .95
age            1.04040    0.96117   1.02488    1.0562
edema          2.45055    0.40807   1.43959    4.1715
log(bili)      2.37157    0.42166   2.01575    2.7902
log(protime)  10.87905    0.09192   2.41232   49.0622
log(albumin)   0.08152   12.26713   0.02267    0.2931

Concordance= 0.835  (se = 0.025 )
Rsquare= 0.426   (max possible= 0.985 )
Likelihood ratio test= 231  on 5 df,   p=0
Wald test            = 234.2  on 5 df,   p=0
```

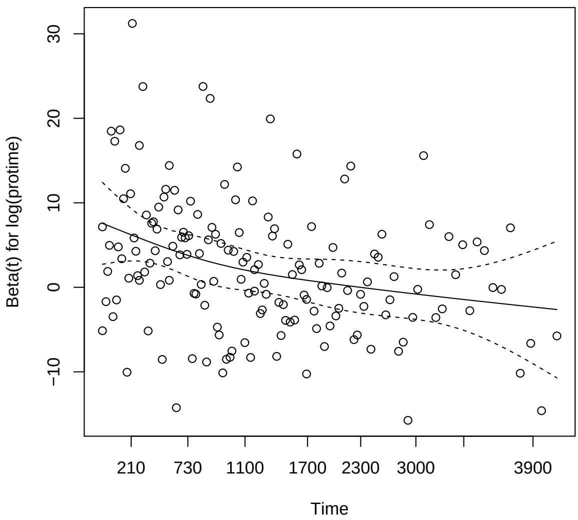
Cirrhose hépatique


```
> cox.zph(cox.pbc)

                     rho    chisq       p
age             -0.00412  0.00225 0.96220
edema           -0.06788  0.74841 0.38698
log(bili)        0.12812  2.25802 0.13292
log(protime)    -0.31827 10.76257 0.00104
log(albumin)     0.03421  0.19877 0.65572
GLOBAL                NA 14.97035 0.01049

> plot(cox.zph(cox.pbc)[4])
```

Cirrhose hépatique



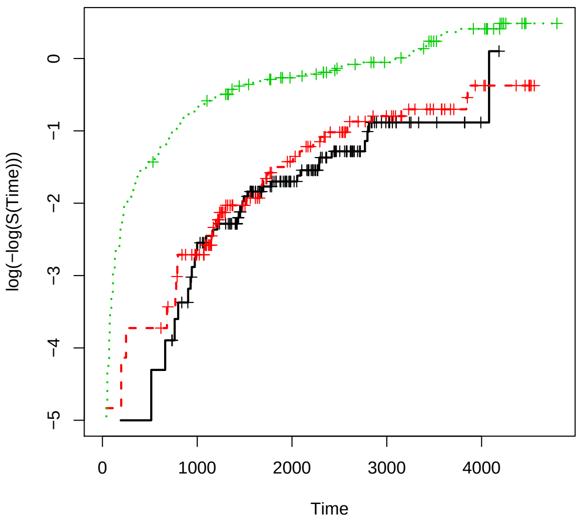
Cirrhose hépatique


```
> quantile(pbc$protime, probs=c(0, 0.33, 0.66, 1), na.rm=TRUE)

  0%  33%  66% 100%
 9.0 10.2 10.9 18.0

> pbc$protime.grp <- 1*(pbc$protime>10.2) + 1*(pbc$protime>10.9)
> f <- survfit(Surv(time, status==2) ~ protime.grp,
+  data = pbc)
> log.minus.log<-function(y)
+ {
+ log(-log(y))
+ }
> plot(f, fun=log.minus.log, ylab="log(-log(S(Time)))",
+  col=1:3, lwd=2, lty=1:3, xlab="Time")
>
```

Cirrhose hépatique



La proportionnalité des risques (5)

- Avantages de l'approche visuelle :
  - Interprétation des résultats.
  - Evaluation possible de la correction pour gérer la non-proportionalité des risques.
- Limites de l'approche visuelle :
  - Interprétations subjectives
  - Seulement une covariable
  - Covariables catégorielles uniquement

La proportionnalité des risques

- Comment modéliser la non-proportionnalité du risque lié à la covariable x_1 ?

$$\lambda(t|x_1, ..., x_p) = \lambda_0(t)exp(\beta_1(t)x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

  - La relation entre la covariable change avec le temps.
- Nous allons montrer que ce problème est proche de celui de la modélisation d'une variable dont la valeur change au cours du temps (time-dependent covariate) :

$$\lambda(t|x_1(t), ..., x_p) = \lambda_0(t)exp(\beta_1 x_1(t) + \beta_2 x_2 + ... + \beta_p x_p)$$

La proportionnalité des risques

- Comment modéliser la non-proportionnalité du risque lié à la covariable x_1 ?

$$\lambda(t|x_1, ..., x_p) = \lambda_0(t)exp(\beta_1(t)x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

  - La relation entre la covariable change avec le temps.
- Nous allons montrer que ce problème est proche de celui de la modélisation d'une variable dont la valeur change au cours du temps (time-dependent covariate) :

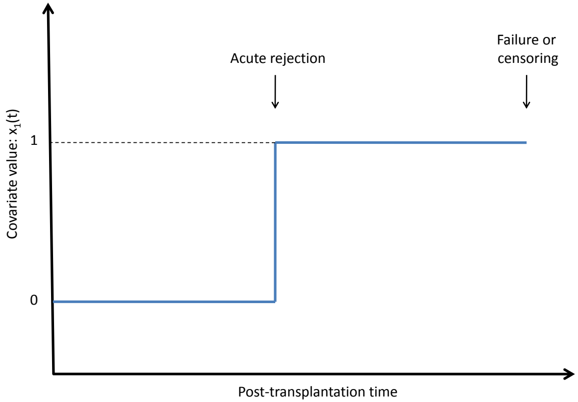
$$\lambda(t|x_1(t), ..., x_p) = \lambda_0(t)exp(\beta_1 x_1(t) + \beta_2 x_2 + ... + \beta_p x_p)$$

Plan

Le principe du modèle

- Ce type de données est fréquent lorsqu'on suit régulièrement un patient.
- Exemple en transplantation rénale :
  - L'origine du temps est la greffe.
  - Le suivi du patient est réalisé jusqu'au premier événement entre le retour en dialyse et le décès du patient (avec un greffon fonctionnel).
  - Toutes les covariables sont collectées à la baseline : sexe, âge du donneur, nombre d'incompatibilités HLA, etc.
  - Parmi ces covariables, certaines changent au cours du temps.
    - Les paramètres biologiques : créatinine, protéine, etc.
    - Evénements post-transplantation : épisodes de rejet aigu.

Le principe du modèle



Le principe du modèle

- Pour gérer ces changements, on peut noter que les contributions individuelles peuvent être divisées en deux parties.
- Si l'individu entre dans l'étude au temps $t = 0$ et subit un échec au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

$$
\begin{aligned}
\lim_{\Delta\tau \to 0} & \frac{P(\tau < T < \tau + \Delta\tau)}{\Delta\tau} \\
= & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau, T > u)}{\Delta\tau} \\
= & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u) P(T > u)}{\Delta\tau} \\
= & P(T > u) \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u)}{\Delta\tau}
\end{aligned}
$$

Le principe du modèle

- Pour gérer ces changements, on peut noter que les contributions individuelles peuvent être divisées en deux parties.
- Si l'individu entre dans l'étude au temps $t = 0$ et subit un échec au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

$$
\begin{aligned}
\lim_{\Delta\tau \to 0} & \frac{P(\tau < T < \tau + \Delta\tau)}{\Delta\tau} \\
= & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau, T > u)}{\Delta\tau} \\
= & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u) P(T > u)}{\Delta\tau} \\
= & P(T > u) \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u)}{\Delta\tau}
\end{aligned}
$$

- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = 0$ et qui est censuré en $t = u$.

Le principe du modèle

- Pour gérer ces changements, on peut noter que les contributions individuelles peuvent être divisées en deux parties.
- Si l'individu entre dans l'étude au temps $t = 0$ et subit un échec au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

$$
\begin{aligned}
\lim_{\Delta\tau \to 0} \quad & \frac{P(\tau < T < \tau + \Delta\tau)}{\Delta\tau} \\
= \quad & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau, T > u)}{\Delta\tau} \\
= \quad & \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u) P(T > u)}{\Delta\tau} \\
= \quad & P(T > u) \lim_{\Delta\tau \to 0} \frac{P(\tau < T < \tau + \Delta\tau | T > u)}{\Delta\tau}
\end{aligned}
$$

- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = 0$ et qui est censuré en $t = u$.
- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = u$ et qui subit l'échec en $t = \tau$.

Le principe du modèle

- Si l'individu entre dans l'étude en $t = 0$ et ne subit aucun échec jusqu'au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

$$\begin{aligned} P(T > \tau) &= P(T > \tau, T > u) \\ &= P(T > u)P(T > \tau | T > u) \end{aligned}$$

Le principe du modèle

- Si l'individu entre dans l'étude en $t = 0$ et ne subit aucun échec jusqu'au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

$$\begin{aligned} P(T > \tau) &= P(T > \tau, T > u) \\ &= P(T > u)P(T > \tau | T > u) \end{aligned}$$

- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = 0$ et qui est censuré en $t = u$.

Le principe du modèle

- Si l'individu entre dans l'étude en $t = 0$ et ne subit aucun échec jusqu'au temps $t = \tau$. Pour tout temps $u < \tau$, on a :

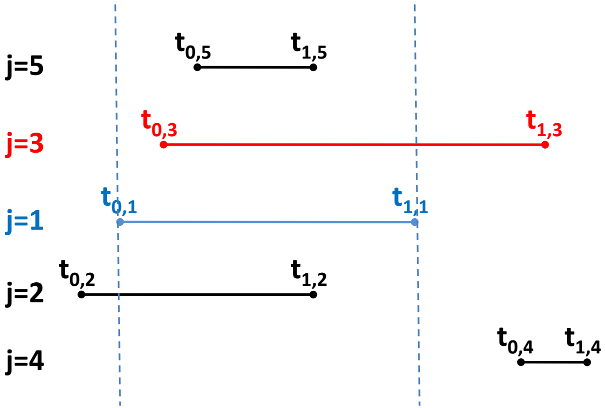
$$
\begin{aligned}
P(T > \tau) &= P(T > \tau, T > u) \\
&= P(T > u)P(T > \tau | T > u)
\end{aligned}
$$

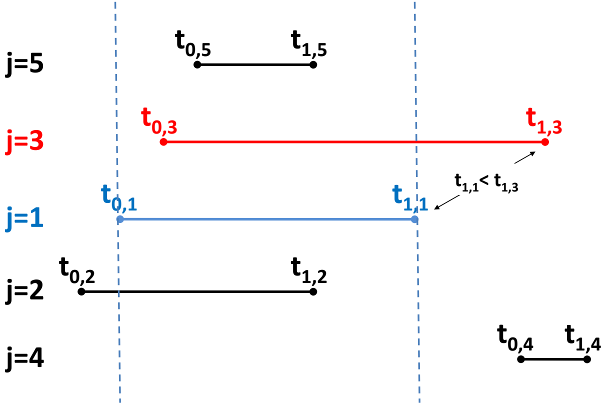
- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = 0$ et qui est censuré en $t = u$.
- Contribution individuelle à la vraisemblance d'un individu qui entre en $t = u$ et qui est censuré en $t = \tau$.

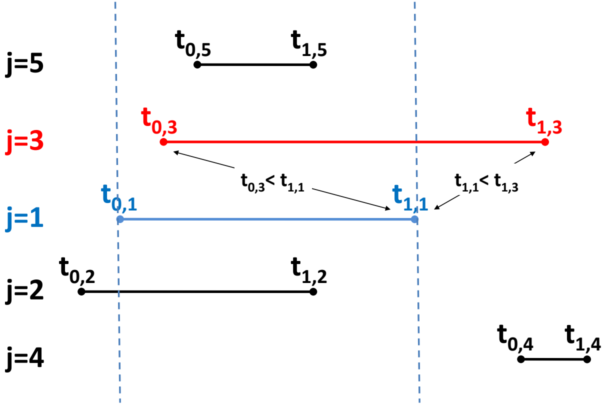
L'estimation des paramètres

$$\lambda(t|x_1(t), ..., X_p = x_p) = \lambda_0(t)exp\left(\beta_1 x_1(t) + \sum_{k=2}^{p} \beta_k x_k\right)$$

- L'approche par vraisemblance partielle peut toujours être utilisée. La seule différence est l'identification des patients considérés à risque.
- Supposons N individus inclus ($j = 1, 2, ..., N$)
- Pour le $j$ème individu *fictif*, notons $t_{1,j}$ le dernier temps de suivi avec $\delta_j = 1$ si l'événement est observé et $\delta_j = 0$ sinon (censure).
- Pour le $j$ème individu *fictif*, notons $t_{0,j}$ le temps d'entrée dans l'étude.


j=5
$t_{0,5}$
$t_{1,5}$
j=3
$t_{0,3}$
$t_{1,3}$
j=1
$t_{0,1}$
$t_{1,1}$
j=2
$t_{0,2}$
$t_{1,2}$
j=4
$t_{0,4}$
$t_{1,4}$


j=5
$t_{0,5}$
$t_{1,5}$
j=3
$t_{0,3}$
$t_{1,3}$
$t_{1,1} < t_{1,3}$
j=1
$t_{0,1}$
$t_{1,1}$
j=2
$t_{0,2}$
$t_{1,2}$
j=4
$t_{0,4}$
$t_{1,4}$


j=5
$t_{0,5}$
$t_{1,5}$
j=3
$t_{0,3}$
$t_{1,3}$
$t_{0,3} < t_{1,1}$
$t_{1,1} < t_{1,3}$
j=1
$t_{0,1}$
$t_{1,1}$
j=2
$t_{0,2}$
$t_{1,2}$
j=4
$t_{0,4}$
$t_{1,4}$

L'estimation des paramètres

- La probabilité conditionnelle que l'individu *fictif* j subit l'événement en $t_{1,j}$ sachant qu'il est à risque à ce même temps :

$$\begin{aligned} \ell_j &= \frac{\lambda(t_{1,j}|x_{1j}(t_{1,j}),...,x_{pj})}{\sum_{i:\{t_{1,i}\geq t_{1,j}\cap t_{0,i}<t_{1,j}\}} \lambda(t_{1,i}|x_{1i}(t_{1,i}),...,x_{pi})} \\ &= \frac{exp(\beta_1 x_{1j}(t_{1,j}) + \sum_{k=2}^{p} \beta_k x_{kj})}{\sum_{i:\{t_{1,i}\geq t_{1,j}\cap t_{0,i}<t_{1,j}\}} exp(\beta_1 x_{1i}(t_{1,i}) + \sum_{k=2}^{p} \beta_k x_{ki})} \end{aligned}$$

- Supposons que tous les individus sont indépendants, le produit des probabilités individuelles précédentes représente la probabilité d'observer l'échantillon. La vraisemblance partielle est :

$$\wp\ell = \prod_{j=1}^{N} \ell_j^{\delta_j}$$

- Les estimations des paramètres de régression β_k $(k = 1, ..., p)$ correspondent aux valeurs qui maximisent la vraisemblance partielle $\wp\ell$.

Les épisodes de rejet aigu

- Exemple de données pour trois individus :

Indiv	Age	Sexe	Temps.Evt	Evt	Rejet	Temps.Rejet
1	32	M	356	0	0	NA
2	56	F	102	1	1	40
3	43	M	670	0	1	180

- Transformation de la table :

Indiv	Age	Sexe	Start	Stop	Evt	Rejet
1	32	M	0	356	0	0
2	56	F	0	40	0	0
2	56	F	41	102	1	1
3	43	M	0	180	0	0
3	43	M	181	670	0	1

Les épisodes de rejet aigu


```
> library(survival)
> Divat<-read.table("Divat_rejet.csv", header = TRUE, sep = ";", dec = ",")
> Divat[1:10,]
   sexeR ageR sexeD ageD  isc nbdial.post start stop rejet evt
1      1   26     1   17  898           0     0    4     0   0
2      1   26     1   17  898           0     5 1101     1   0
3      1   72     0   68  900          NA     0    1     0   1
4      1   70     1   73 1350           5     0   24     0   0
5      1   70     1   73 1350           5    25 1922     1   0
6      0   48     1   45 1800           0     0 3318     0   0
7      1   60     1   49 1145           6     0    9     0   0
8      1   60     1   49 1145           6    10  671     1   0
9      1   30     1   23 1347           7     0  216     0   0
10     1   30     1   23 1347           7   217 1286     1   0
```

Les épisodes de rejet aigu


```
> summary(cox.divat<-coxph(Surv(start, stop, evt) ~ sexeR + I(ageD>55) +
+  I(ageR>55) + I(nbdial.post > 0) +  rejet, data = Divat))
Call:
coxph(formula = Surv(start, stop, evt) ~ sexeR + I(ageD > 55) +
    I(ageR > 55) + I(nbdial.post > 0) + rejet, data = Divat)

  n= 5039, number of events= 596
   (292 observations deleted due to missingness)

                           coef exp(coef) se(coef)      z Pr(>|z|)
sexeR                  -0.08349   0.91990  0.08451 -0.988    0.323
I(ageD > 55)TRUE        0.40645   1.50148  0.09402  4.323 1.54e-05 ***
I(ageR > 55)TRUE        0.42887   1.53553  0.09271  4.626 3.73e-06 ***
I(nbdial.post > 0)TRUE  0.58274   1.79094  0.08281  7.037 1.96e-12 ***
rejet                   0.54817   1.73009  0.09129  6.005 1.91e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                       exp(coef) exp(-coef) lower .95 upper .95
sexeR                     0.9199     1.0871    0.7795     1.086
I(ageD > 55)TRUE          1.5015     0.6660    1.2488     1.805
I(ageR > 55)TRUE          1.5355     0.6512    1.2804     1.842
I(nbdial.post > 0)TRUE    1.7909     0.5584    1.5226     2.107
rejet                     1.7301     0.5780    1.4467     2.069

Concordance= 0.653  (se = 0.013 )
Rsquare= 0.032   (max possible= 0.839 )
Likelihood ratio test= 163  on 5 df,   p=0
Wald test            = 167.6  on 5 df,   p=0
```

Les épisodes de rejet aigu

- Les patients recevant un greffon de plus de 55 ans ont environ 1.5 fois plus de risque d'échec de greffe que les plus jeunes (IC95%=[1.25, 1.80]).
- Les patients de plus de 55 ans ont environ 1.5 fois plus de risque d'échec de greffe que les plus jeunes (IC95%=[1.28 ; 1.84]).
- Les patients qui ont besoin d'au moins une dialyse au moment de la greffe ont environ 1.7 fois plus de risque d'échec de greffe que les plus jeunes (IC95%=[1.52, 2.11]).
- Les patients qui déclarent un épisode de rejet aigu voit leur risque d'échec multiplié par 1.7 (IC95%=[1.45, 2.07]).

Les épisodes de rejet aigu


```
> cox.zph(cox.divat)

                               rho  chisq     p
sexeR                      0.01510 0.1362 0.712
I(ageD > 55)TRUE          -0.00582 0.0190 0.890
I(ageR > 55)TRUE           0.00848 0.0396 0.842
I(nbdial.post > 0)TRUE -0.04797 1.3766 0.241
rejet                     -0.04382 1.1424 0.285
GLOBAL                          NA 2.7580 0.737
```

Plan

1. Introduction

2. La valeur est dépendante du temps

3. L'effet est dépendant du temps

Reformulation du problème

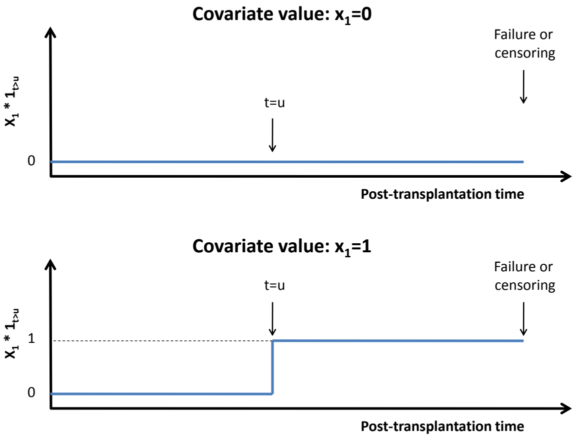
- Comment modéliser un effet non-proportionnel de la covariable x_1 ?

$$\lambda(t|x_1, ..., x_p) = \lambda_0(t) exp(\beta_1(t) x_1 + \beta_2 x_2 + ... + \beta_p x_p)$$

- Supposons que l'effet de la covariable change au temps u. Ce changement peut être modélisé avec une interaction entre la covariable et le temps.

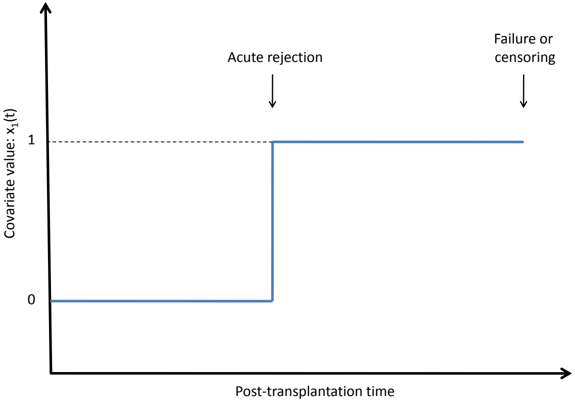
$$\lambda(t|x_1, ..., x_p) = \lambda_0(t) exp(\beta_1 x_1 + \gamma_1 x_1 \mathbb{1}_{t>u} + \beta_2 x_2 + ... + \beta_p x_p)$$

  - où $\mathbb{1}_{t>u}$ égal à 1 si $t > u$ et 0 sinon.
- Si $t \leq u$ alors $HR_{x_1=1/x_1=0} = exp(\beta_1)$.
- Si $t > u$ alors $HR_{x_1=1/x_1=0} = exp(\beta_1 + \gamma_1)$.
- Si $\gamma_1 = 0$, alors il n'y a aucun effet dépendant du temps.


Covariate value: x_1=0
$x_1 * 1_{t>u}$
Failure or censoring
t=u
0
Post-transplantation time
Covariate value: x_1=1
t=u
Failure or censoring
$x_1 * 1_{t>u}$
1
0
Post-transplantation time

Reformulation du problème

- Il s'agit de la même approche quand la valeur (et non l'effet) de la covariable est dépendant du temps.



L'anticorps anti-AT1R

- Données provenant de DIVAT ($N = 599$).
  - Données Informatisées et VAlidées en Transplantation
  - Receveurs de greffe rénale depuis 1990.
  - Les données biologiques et cliniques sont prospectivement collectées.
  - 5 centres français : Nantes, Paris Necker, Nancy, Toulouse and Montpellier.
  - Audits annuelles entre les centres qui montrent moins de 1% d'erreurs dans les données collectées.
- Les variables collectées :
  - Niveau d'anticorps anti-AT1R en pré-greffe, nombre d'incompatibilités HLA, immunisation (PRA en %).

Objectif : Etudier la relation entre le niveau d'anticorps anti-AT1R et le délai entre la greffe et l'apparition d'un épisode de rejet aigu.

L'anticorps anti-AT1R


```
> AT1R<-read.table("AT1R.csv", header = TRUE, sep = ";", dec = ",")
> AT1R[1:10,]
```

	Start	Stop	EvtRejet	AT1R.ini	AT1R0	AT1R1	IncompABDR	PRA.T
1	0	0.03011636	1	1	1	0	4	0
2	0	0.03285421	1	0	0	0	4	6
3	0	0.31211499	1	1	1	0	6	0
4	0	0.05201916	1	0	0	0	2	38
5	0	0.12046543	1	1	1	0	4	0
6	0	0.15605749	1	0	0	0	3	0
7	0	0.02190281	1	0	0	0	6	0
8	0	0.04106776	1	0	0	0	0	91
9	0	0.02737851	1	0	0	0	3	0
10	0	0.03285421	1	0	0	0	0	36

L'anticorps anti-AT1R


```
> summary(cox.at1r<-coxph(Surv(Start, Stop, EvtRejet) ~   AT1R.ini +
+  IncompABDR + I(PRA.T>0) , data = AT1R))

Call:
coxph(formula = Surv(Start, Stop, EvtRejet) ~ AT1R.ini + IncompABDR +
    I(PRA.T > 0), data = AT1R)

  n= 1144, number of events= 59
   (10 observations deleted due to missingness)

                    coef exp(coef) se(coef)     z Pr(>|z|)
AT1R.ini         0.81886   2.26791  0.30710 2.666  0.00767 **
IncompABDR       0.19076   1.21017  0.09557 1.996  0.04592 *
I(PRA.T > 0)TRUE 0.45939   1.58311  0.29131 1.577  0.11480
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                 exp(coef) exp(-coef) lower .95 upper .95
AT1R.ini             2.268     0.4409    1.2423     4.140
IncompABDR           1.210     0.8263    1.0035     1.459
I(PRA.T > 0)TRUE     1.583     0.6317    0.8944     2.802

Concordance= 0.622  (se = 0.038 )
Rsquare= 0.01   (max possible= 0.477 )
Likelihood ratio test= 11.08  on 3 df,   p=0.0113
Wald test            = 11.92  on 3 df,   p=0.007674
Score (logrank) test = 12.31  on 3 df,   p=0.006395
```

L'anticorps anti-AT1R

- Il semble que les patients avec des anticorps anti-AT1R aient 2.3 fois plus de risque que les autres de déclarer un épisode de rejet aigu ($p=0.0077$).

L'anticorps anti-AT1R

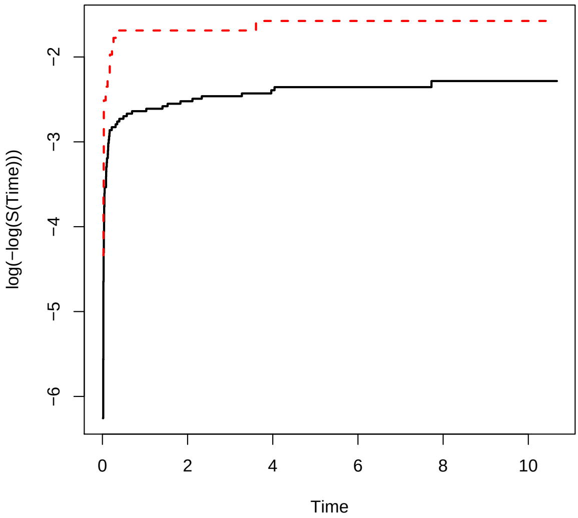

```
> cox.zph(cox.at1r)

                       rho  chisq     p
AT1R.ini           -0.0958 0.5289 0.467
IncompABDR          0.0278 0.0456 0.831
I(PRA.T > 0)TRUE   -0.1654 1.5222 0.217
GLOBAL                  NA 2.4274 0.489
```

L'anticorps anti-AT1R


```
> log.minus.log<-function(y) { log(-log(y)) }
> f <- survfit(Surv(Start, Stop, EvtRejet) ~   AT1R.ini,
+  data = AT1R)
> plot(f , fun=log.minus.log, ylab="log(-log(S(Time)))",
+ col=1:2, lwd=2, lty=1:2, xlab="Time", mark.time=FALSE)
```

L'anticorps anti-AT1R



Indiv	Age	Sexe	Temps.RA	RA	AT1R
1	32	M	276	0	1
2	56	F	96	1	0
3	43	M	873	0	0

- Transformation de la table pour $u = 120$:

Indiv	Age	Sexe	Start	Stop	RA	AT1R	AT1R.t
1	32	M	0	120	0	1	0
1	32	M	121	276	0	1	1
2	56	F	0	96	1	0	0
3	43	M	0	120	0	0	0
3	43	M	121	873	0	0	0

L'anticorps anti-AT1R


```
> AT1R[1:10,]

   Start       Stop EvtRejet AT1R.ini AT1R0 AT1R1 IncompABDR PRA.T
1      0 0.03011636        1        1     1     0          4     0
2      0 0.03285421        1        0     0     0          4     6
3      0 0.31211499        1        1     1     0          6     0
4      0 0.05201916        1        0     0     0          2    38
5      0 0.12046543        1        1     1     0          4     0
6      0 0.15605749        1        0     0     0          3     0
7      0 0.02190281        1        0     0     0          6     0
8      0 0.04106776        1        0     0     0          0    91
9      0 0.02737851        1        0     0     0          3     0
10     0 0.03285421        1        0     0     0          0    36
```

L'anticorps anti-AT1R


```
> summary(cox.at1r.time<-coxph(Surv(Start, Stop, EvtRejet) ~   AT1R.ini
+ + AT1R1 + IncompABDR + I(PRA.T>0) , data = AT1R))
Call:
coxph(formula = Surv(Start, Stop, EvtRejet) ~ AT1R.ini + AT1R1 +
    IncompABDR + I(PRA.T > 0), data = AT1R)

  n= 1144, number of events= 59
   (10 observations deleted due to missingness)

                     coef exp(coef) se(coef)      z Pr(>|z|)
AT1R.ini          1.10631   3.02318  0.33136  3.339 0.000842 ***
AT1R1            -1.72688   0.17784  1.08670 -1.589 0.112038
IncompABDR        0.18951   1.20865  0.09535  1.987 0.046876 *
I(PRA.T > 0)TRUE  0.45456   1.57548  0.29150  1.559 0.118900
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

                 exp(coef) exp(-coef) lower .95 upper .95
AT1R.ini            3.0232     0.3308   1.57911     5.788
AT1R1               0.1778     5.6231   0.02114     1.496
IncompABDR          1.2087     0.8274   1.00262     1.457
I(PRA.T > 0)TRUE    1.5755     0.6347   0.88980     2.790

Concordance= 0.637  (se = 0.038 )
Rsquare= 0.013   (max possible= 0.477 )
Likelihood ratio test= 14.76  on 4 df,   p=0.005235
Wald test            = 16.3  on 4 df,   p=0.002638
Score (logrank) test = 17.45  on 4 df,   p=0.00158
```

L'anticorps anti-AT1R

- L'effet des anticorps anti-AT1R n'a pas pu être montré comme significativement différent avant et après 120 jours ($p=0.1120$).
- Comme le manque de puissance est évident et l'écart des risques ($\exp(-1.7)$) très important avant et après 120 jours, on a conservé cet effet dépendant du temps dans le modèle.

L'anticorps anti-AT1R


```
> summary(cox.at1r.time<-coxph(Surv(Start, Stop, EvtRejet) ~   AT1R0
+ + AT1R1 + IncompABDR + I(PRA.T>0) , data = AT1R))
Call:
coxph(formula = Surv(Start, Stop, EvtRejet) ~ AT1R0 + AT1R1 +
    IncompABDR + I(PRA.T > 0), data = AT1R)

  n= 1144, number of events= 59
   (10 observations deleted due to missingness)

                 coef exp(coef) se(coef)      z Pr(>|z|)
AT1R0         1.10631   3.02318  0.33136  3.339 0.000842 ***
AT1R1        -0.62056   0.53764  1.03556 -0.599 0.549004
IncompABDR    0.18951   1.20865  0.09535  1.987 0.046876 *
I(PRA.T > 0)TRUE  0.45456   1.57548  0.29150  1.559 0.118900
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

             exp(coef) exp(-coef) lower .95 upper .95
AT1R0           3.0232     0.3308   1.57911     5.788
AT1R1           0.5376     1.8600   0.07063     4.092
IncompABDR      1.2087     0.8274   1.00262     1.457
I(PRA.T > 0)TRUE   1.5755     0.6347   0.88980     2.790

Concordance= 0.637  (se = 0.038 )
Rsquare= 0.013   (max possible= 0.477 )
Likelihood ratio test= 14.76  on 4 df,   p=0.005235
Wald test            = 16.3  on 4 df,   p=0.002638
Score (logrank) test = 17.45  on 4 df,   p=0.00158
```

L'anticorps anti-AT1R

- Il semble que les patients avec des anticorps anti-AT1R aient 3.0 fois plus de risque que les autres de déclarer un épisode de rejet aigu dans les 120 jours qui suivent la greffe (p=0.0008).
- Aucune conclusion fiable ne peut être formulée ensuite.