

Analyse de survie, analyse de regroupements

Analyse multidimensionnelle appliquée

Léo Belzile

HEC Montréal

automne 2022

Le **modèle à risques proportionnels de Cox** pour \mathbf{X} au temps t est

$$h(t; \mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \cdots + \beta_p X_p),$$

où $h_0(t)$ est la fonction de risque de base qui remplace l'ordonnée à l'origine.

- Postulat de risques proportionnels: le rapport de risque pour deux observations ne varie pas en fonction du temps t .

Postulat de risques proportionnels

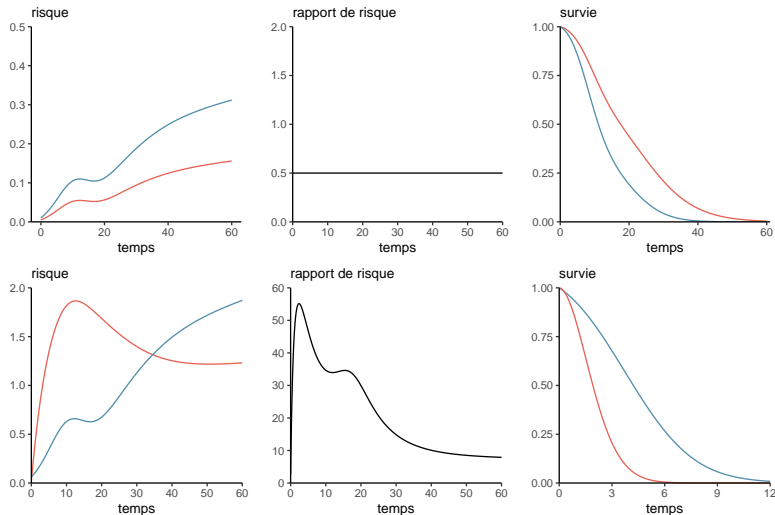


Figure 1: Courbes de risques proportionnelles (panneau supérieur) et non proportionnelles (panneau inférieur).

On peut modéliser la non-proportionnalité par la **stratification** pour une variable catégorielle $Z = 1, \dots, K$.

Supposons que l'effet de Z sur le risque varie dans le temps.

On écrit alors

$$h(t; \mathbf{X}, Z = k) = h_k(t) \exp(\beta_1 X_1 + \dots + \beta_p X_p),$$

où h_k est la fonction de risque pour $Z = k$.

Dans ce modèle

- On suppose que l'effet des variables explicatives \mathbf{X} est le même peu importe la valeur de Z .
- L'effet de $Z = k$ vs $Z = j$ pour un même ensemble de variables explicatives \mathbf{X} est $h_k(t)/h_j(t)$, qui dépend du temps.

- **Avantage:** on peut modéliser n'importe quel changement du risque en fonction de Z .
- **Désavantage:** on perd la variable explicative Z , donc on ne peut tester son effet (pas de coefficient)... on peut résumer l'information pour la variable Z en calculant par exemple les différences de survie à des temps donnés.
- **Désavantage:** la fonction de risque est estimée pour chaque sous-groupe de Z (plus faible taille d'échantillon).

Idéalement, utiliser la stratification avec des variables secondaires ou de contrôles.

Modèle de Cox avec stratification dans R

```
1 library(survival)
2 data(survie1, package = "hecmulti")
3 # Stratification par service
4 cox_strat <- coxph(
5   Surv(temps, 1-censure) ~ age + sexe + strata(service),
6   data = survie1)
7 # Décompte par service
8 with(survie1, table(service))
9 # Coefficients
10 summary(cox_strat)
```

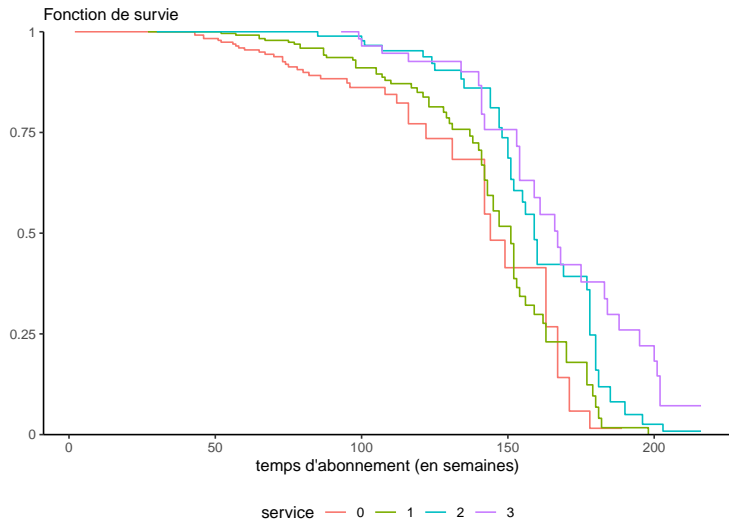
Table 1: Décompte du nombre d'observations par service.

	0	1	2	3
0	197	179	78	46

Table 2: Rapport de risques pour un modèle de Cox stratifié par service.

terme	exp(coef)	borne inf.	borne sup.
age	0.96	0.94	0.97
sexe	0.61	0.44	0.85

Courbes de survie du modèle stratifié



1. Inclusion de variables explicatives dont la valeur change dans le temps.
2. Modèle à risques compétitifs.

On considère une extension du modèle de Cox qui permet d'inclure des variables explicatives dont la valeur change dans le temps.