Cours 2 : notes

Slide 16.

2 individus X_{i_1} et X_{i_2}

$$\frac{\lambda(t/X_{i_1})}{\lambda(t/X_{i_2})} = \frac{\cancel{\lambda_0^*(t)}\, e^{X_{i_1}\beta^*}}{\cancel{\lambda_0^*(t)}\, e^{X_{i_2}\beta^*}} = \exp\left((X_{i_1}-X_{i_2})\beta^*\right)$$

ça ne dépend pas du temps.

$t \mapsto \lambda(t/X_{i_1})$ et $t \mapsto \lambda(t/X_{i_2})$

elles sont proportionnelles en tant que fonctions du temps.

$\Rightarrow$ Proportional hazards model.

Slide 17.

Hazards ratio. je suppose $X_{i_1}^k = X_{i_2}^k \ \forall k \neq j$ et $X_{i_1}^j = X_{i_2}^j + 1$.

$$\frac{\lambda(t/X_{i_1})}{\lambda(t/X_{i_2})} = \exp\left((X_{i_1}-X_{i_2})\beta^*\right)$$

$$= \exp\left(\sum_{k=1}^{p} (X_{i_1}^k - X_{i_2}^k)\beta_k^*\right)$$

$$= \exp\left((X_{i_1}^j - X_{i_2}^j)\beta_j^*\right) = \underbrace{\exp(\beta_j^*)}$$

$\exp(\beta_j^*)$ = valeur par laquelle est multipliée le hazard quand la cov. j augmente, toutes choses égales par ailleurs. hazard ratio.

Slide 21.

$$\mathcal{L} \underset{\text{prop à}}{\propto} \prod_{i=1}^{n} f(T_i^c/X_i)^{\delta_i} \bar{F}(T_i^c/X_i)^{1-\delta_i}$$

Rappel
$\bar{F}(t) = \exp(-\Lambda(t))$

$$= \prod_{i=1}^{n} \left(\underbrace{\frac{f(T_i^c/X_i)}{\bar{F}(T_i^c/X_i)}}\right)^{\delta_i} \bar{F}(T_i^c/X_i)$$

$$= \prod_{i=1}^{n} \lambda(T_i^c/X_i)^{\delta_i} \exp(-\Lambda(T_i^c/X_i))$$

dans le modèle de Cox.

$$= \prod_{i=1}^{n} \lambda_0(T_i^c)^{\delta_i} \exp(X_i\beta)^{\delta_i} \exp\left(-\int_0^{T_i^c} \lambda_0(t)\exp(X_i\beta)\,dt\right)$$

$$= \prod_{i=1}^{n} \lambda_0(T_i^c)^{\delta_i} \exp(X_i\beta)^{\delta_i} \exp(-\exp(X_i\beta)\Lambda_0(T_i^c))$$

où $\Lambda_0(t) = \int_0^t \lambda_0(s)\,ds$.

$$\log\mathcal{L} = \sum_{i=1}^{n} \delta_i \log\lambda_0(T_i^c) + \delta_i (X_i\beta) - \exp(X_i\beta)\Lambda_0(T_i^c).$$

On remarque que

voir page suivante

$$\sum_{i=1}^{n} \Lambda_0(T_i^c)\exp(X_i\beta) = \sum_{i=1}^{n} \lambda_0(T_i^c) \sum_{j:T_j^c \geq T_i^c} \exp(X_j\beta).$$

$$\lambda_0 \leftrightarrow \begin{pmatrix} \lambda_0^1 \\ | \\ \lambda_0^n \end{pmatrix}. \qquad \sum_{i=1}^{n} \delta_i \log\lambda_0^i + \delta_i (X_i\beta) - \exp(X_i\beta) \sum_{j:T_j^c \leq T_i^c} \lambda_0^j$$

On dérive % à λ_0^i

$$\frac{\delta_i}{\lambda_0^i} - \sum_{j:T_j^c \geq T_i^c} \exp(X_j\beta) = 0$$

(*) . $\sum_{i=1}^{n} \Lambda_0(T_i^c) \exp(X_i \beta)$

$$= \sum_{i=1}^{n} \sum_{j : T_j^c \leq T_i^c} \lambda_0(T_j^c) \exp(X_i \beta)$$

$$= \sum_{j=1}^{n} \left(\sum_{i : T_i^c \geq T_j^c} \exp(X_i \beta) \right) \lambda_0(T_j^c)$$

en échangeant les 2 sommes.

$$\text{"}\hat{\lambda}_0^j\text{"} = \frac{\delta_i}{\sum_{j : T_j^c \geq T_i^c} \exp(X_j \beta)}$$

ce n'est pas un estimateur car il dépend de β inconnu.

Pour une valeur β fixée c'est le mieux qu'on puisse faire.

$$\sum_{i=1}^{n} \delta_i \log \frac{\delta_i}{\sum_{j : T_j^c \geq T_i^c} \exp(X_j \beta)} + \delta_i X_i \beta$$

$$- \exp(X_i \beta) \sum_{k : T_k^c \leq T_i^c} \frac{\delta_k}{\sum_{j : T_j^c \geq T_i^c} \exp(X_i \beta)}$$

on n'a plus de dépendance de nouvelle "log-vraisemblance" qui ne dépend plus du paramètre λ_0. → log-vraisemblance partielle de Cox.

On la maximise $\hat{\beta}$.

puis on injecte ce $\hat{\beta}$ dans $\text{"}\hat{\lambda}_0^j\text{"}(\hat{\beta}) = \frac{\delta_i}{\sum_{j : T_j^c \geq T_i^c} \exp(X_i \hat{\beta})}$

Slide 30.

le risque de l'individu i $\lambda_0^*(t) \exp(X_i \beta^*)$

j $\lambda_0^*(t) \exp(X_j \beta^*)$

$M_i = X_i \hat{\beta}^*$ et $M_j = X_j \hat{\beta}^*$.

Si $M_i = X_i \beta^* > X_j \beta^* = M_j$ alors le risque de i > risque de j en tout temps.

$$\lambda_0^*(t)\exp(X_i\beta^*) > \lambda_0^*(t)\exp(X_j\beta^*)$$

$\Rightarrow$ on s'attend à ce que i meurt avant j

$T^i < T_j$.

On dit qu'il y a concordance quand une paire i,j telle $T^i < T^j$ on a $X_i\beta^* > X_j\beta^*$.

on ne l'observe pas toujours à cause des censures.

$\rightarrow$ C-index = proportion d'observations concordantes dans l'échantillon.

⚠ c'est seulement une mesure de discrimination.

Prédiction dans un modèle de Cox pour un nouvel ind X_+.

coxph $\rightarrow \hat\beta$

predict(coxph, X_+) $\rightarrow X_+\hat\beta$.

survfit(coxph) $\rightarrow \hat\Lambda_0$

$$\rightarrow \exp\left(-\underbrace{\hat\Lambda_0(t)}_{\text{Breslow}}\, e^{X_+\hat\beta}\right) = \hat F(t/X_+)$$