Analyse de Survie

Professeur Abdellatif El Afia

Rappel : Estimateur de Kaplan Meier (KM)

Hypothèses :

- Censure non-informative.
- Temps de survie indépendants.

▪l'hypothèse nulle: H_0:"$Il\ n'\ y\ a\ pas\ de\ difference\ statistique\ entre\ les\ deux\ groupes$"

Estime : la fonction de survie ***S(t)*** en fonction en éscaliers**.**

Utilisé souvent pour mesurer la fraction d'individus en vie pour une certaine durée, et comparer la survie de deux ou plusieurs groupes.

Formule : $\hat{S}(t_k) = \prod_{t_k<t} S(t_{k-1})\left(1 - \frac{d_k}{n_k}\right) \qquad 1 < k < j$

Log-Rank test : : $\chi^2 = \sum_g \frac{(O^g - E^g)^2}{E^g}$

Modèles paramétriques

Risque instantané constant :

- loi exponentielle

$$f(t/\theta) = \theta e^{-\theta t},$$
$$\lambda(t/\theta) = \theta,$$
$$S(t/\theta) = e^{-\theta t}$$

Risque instantané monotone :

- Loi de Weibull

$$f(t \mid \theta, \nu) = \nu \left(\frac{1}{\theta}\right)^{\nu} t^{\nu-1} \exp\left(-\left(\frac{t}{\theta}\right)^{\nu}\right), \qquad t \geqslant 0 \text{ et } \theta, \nu > 0$$
$$\lambda(t \mid \theta, \nu) = \nu \left(\frac{1}{\theta}\right)^{\nu} t^{\nu-1},$$
$$S(t \mid \theta, \nu) = \exp\left(-\left(\frac{t}{\theta}\right)^{\nu}\right).$$

- Loi Gamma

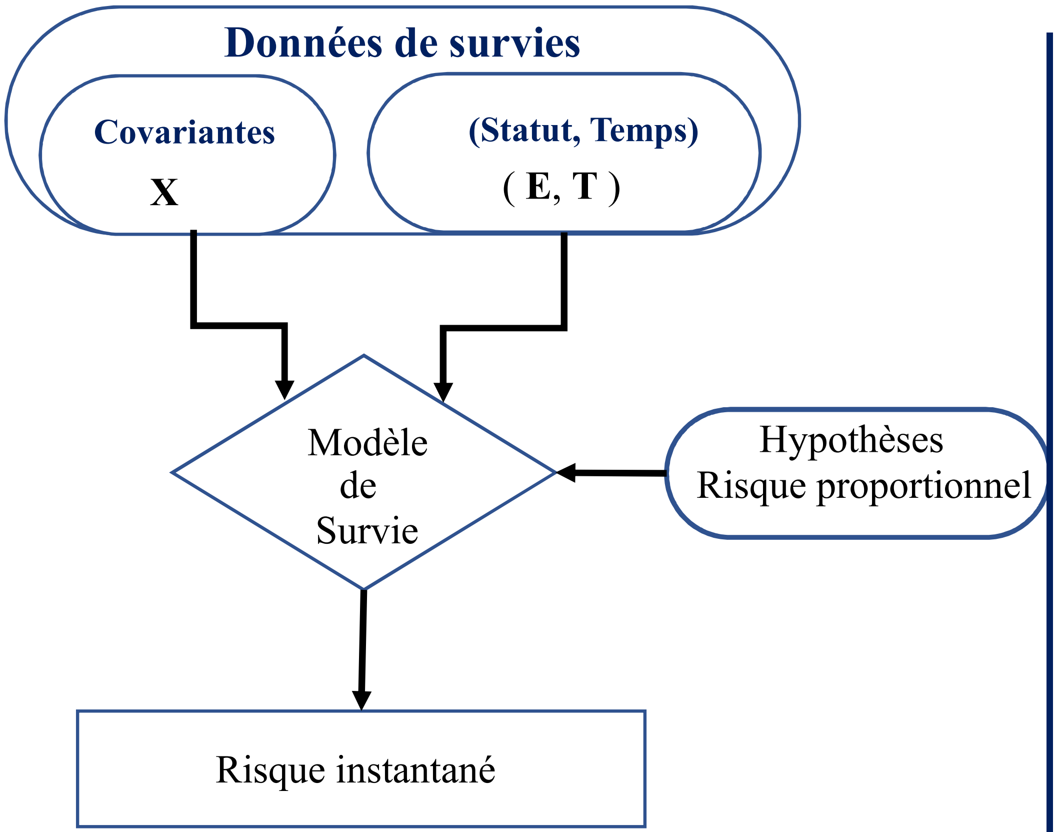
$$f(t \mid \nu, \theta) = \frac{\theta^{\nu}}{\Gamma(\nu)} t^{\nu-1} e^{-\theta t}, \qquad t \geqslant 0 \text{ et } \theta, \nu > 0$$
$$F(t \mid \nu, \theta) = \frac{1}{\Gamma(\nu)} \int_0^{\theta t} u^{\nu-1} e^{-u} du,$$
$$\lambda(t \mid \nu, \theta) = \frac{f(t \mid \theta, \nu)}{1 - F(t \mid \theta, \nu)},$$

Risque instantané en ∩ et ∪ :

- Lois de Weibull généralisées
- La loi Log-logistique de paramètre (α,β)
- La Log-normales de paramètre (μ,σ)
- La loi gaussienne inverse (μ,λ)

$$f(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left(-\frac{\lambda}{2\mu^2 t}(t-\mu)^2\right),$$
$$\lambda > 0, \qquad t \geq 0$$

Semi Paramétriques COX PH



Fonction de risque instantané :

$$\lambda(t, X_1, \ldots, X_n) = \lambda_0(t).\exp(\sum_i \beta_i X_i)$$

Risque de base $\boldsymbol{\lambda_0(t)}$, risque instantané quand tout les variables sont nulles.

Hypothèse des risques proportionnels : l'effet de chaque covariable est indépendant du temps.

Modèles semi-paramétriques : Cox PH

Hypothèses :
- Censure non-informative.
- Temps de survie indépendants.
- Risques proportionnels (HR est constant).
- $\ln\big(h(t, X)\big)$ est une fonction linéaire de **X.**
- **Covariante** ne changent pas par rapport au temps.

Utilisé pour :
- Comparer l’éventuel effet d’une covariable sur la survie (coefficient).
- Hazard ration HR
- Courbes de survie ajustées.

Modèles semi-paramétriques : Cox PH

Leukemia Remission Data

Group 1($n = 21$)		Group 2($n = 21$)	
t (weeks)	log WBC	t (weeks)	log WBC
6	2.31	1	2.80
6	4.06	1	5.00
6	3.28	2	4.91
7	4.43	2	4.48
10	2.96	3	4.01
13	2.88	4	4.36
16	3.60	4	2.42
22	2.32	5	3.49
23	2.57	5	3.97
6+	3.20	8	3.52
9+	2.80	8	3.05
10+	2.70	8	2.32
11+	2.60	8	3.26
17+	2.16	11	3.49
19+	2.05	11	2.12
20+	2.01	12	1.50
25+	1.78	12	3.06
32+	2.20	15	2.30
32+	2.53	17	2.95
34+	1.47	22	2.73
35+	1.45	23	1.97

+ denotes censored observation

Modèles semi-paramétriques : Cox PH

Model 1:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.509	0.410	3.68	0.000	4.523	2.027	10.094
No. of subjects = 42		Log likelihood = −86.380			Prob > chi2 = 0.0001		

Model 2:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.294	0.422	3.07	0.002	3.648	1.595	8.343
log WBC	1.604	0.329	4.87	0.000	4.975	2.609	9.486
No. of subjects = 42		Log likelihood = −72.280			Prob > chi2 = 0.0000		

Model 3:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	2.355	1.681	1.40	0.161	10.537	0.391	284.201
log WBC	1.803	0.447	4.04	0.000	6.067	2.528	14.561
Rx x log WBC	−0.342	0.520	−0.66	0.510	0.710	0.256	1.967
No. of subjects = 42		Log likelihood = −72.066			Prob > chi2 = 0.0000		

Modèles semi-paramétriques : Cox PH

Model 1:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.509	0.410	3.68	0.000	4.523	2.027	10.094

No. of subjects = 42 | Log likelihood = −86.380 | Prob > chi2 = 0.0001

Model 2:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.294	0.422	3.07	0.002	3.648	1.595	8.343
log WBC	1.604	0.329	4.87	0.000	4.975	2.609	9.486

No. of subjects = 42 | Log likelihood = −72.280 | Prob > chi2 = 0.0000

Model 3:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	2.355	1.681	1.40	0.161	10.537	0.391	284.201
log WBC	1.803	0.447	4.04	0.000	6.067	2.528	14.561
Rx x log WBC	−0.342	0.520	−0.66	0.510	0.710	0.256	1.967

No. of subjects = 42 | Log likelihood = −72.066 | Prob > chi2 = 0.0000

Modèles semi-paramétriques : Cox PH

Model 1:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.509	0.410	3.68	0.000	4.523	2.027	10.094
No. of subjects = 42		Log likelihood = −86.380			Prob > chi2 = 0.0001		

Model 2:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.294	0.422	3.07	0.002	3.648	1.595	8.343
log WBC	1.604	0.329	4.87	0.000	4.975	2.609	9.486
No. of subjects = 42		Log likelihood = −72.280			Prob > chi2 = 0.0000		

Model 3:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	2.355	1.681	1.40	0.161	10.537	0.391	284.201
log WBC	1.803	0.447	4.04	0.000	6.067	2.528	14.561
Rx x log WBC	−0.342	0.520	−0.66	0.510	0.710	0.256	1.967
No. of subjects = 42		Log likelihood = −72.066			Prob > chi2 = 0.0000		

Modèles semi-paramétriques : Cox PH

Model 1:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.509	0.410	3.68	0.000	4.523	2.027	10.094
No. of subjects = 42		Log likelihood = −86.380			Prob > chi2 = 0.0001		

8.067

Model 2:

	Coef.	Std. Err.	z	p > |z|	Haz. Ratio	[95% Conf. Interval]	
Rx	1.294	0.422	3.07	0.002	3.648	1.595	8.343
log WBC	1.604	0.329	4.87	0.000	4.975	2.609	9.486
No. of subjects = 42		Log likelihood = −72.280			Prob > chi2 = 0.0000		

6.748

Modèles semi-paramétriques : Cox PH

Formule :

$$h(t, X_1, .., X_n) = h_0(t). \exp(\sum_i \beta_i X_i)$$

$h_0(t)$ the baseline hazard (risque de base). **Non spécifié** **(semi-paramétrique).**
X_i covariables (indépendantes de temps).
$\exp(\sum_i \beta_i X_i)$ est positif, ce qui arrange le risque.

Modèles semi-paramétriques : Cox PH

Robuste :
Si on connait e modèle paramétrique, tant mieux, sinon, Cox donne une approximation robuste

Pas besoin de spécifier le risque de base $h_0(t)$ pour calculer **HR**.

Maximum Likelyhood

Vraisemblance partielle :

$$L_i(\beta) = \frac{h_i(t_i)}{\sum_{l \in R_i} h_l(t_i)} = \frac{e^{\beta^T x_i}}{\sum_{l \in R_i} e^{\beta^T x_l}}$$

Vraisemblance totale :

$$L(\beta) = \prod_{i=1}^{k} L_i(\beta) = \prod_{i=1}^{k} \frac{e^{\beta^T x_i}}{\sum_{l \in R_i} e^{\beta^T x_l}}$$

La fonction de Score U(β) est définie comme la dérivée première du log de vraisemblance :

$$U(\beta) = \frac{\partial}{\partial \beta}\big(l(\beta)\big) = x_i - \frac{\sum_{l \in R_i} x_i e^{\beta^T x_l}}{\sum_{l \in R_i} e^{\beta x_l}}$$

On tire l'estimation $\hat{\beta}$ du paramètre β, par :

$$U(\hat{\beta}) = 0$$

Courbes de survie ajustées

$$\hat{h}(t,\mathbf{X}) = \hat{h}_0(t)e^{1.294\ Rx\ +\ 1.604\ \log\ \text{WBC}}$$
$$\hat{S}(t,\mathbf{X}) = [\hat{S}_0(t)]^{\exp(1.294\ Rx\ +\ 1.604\ \log\ \text{WBC})}$$

Specify values for $\mathbf{X} = (Rx, \log \text{WBC})$

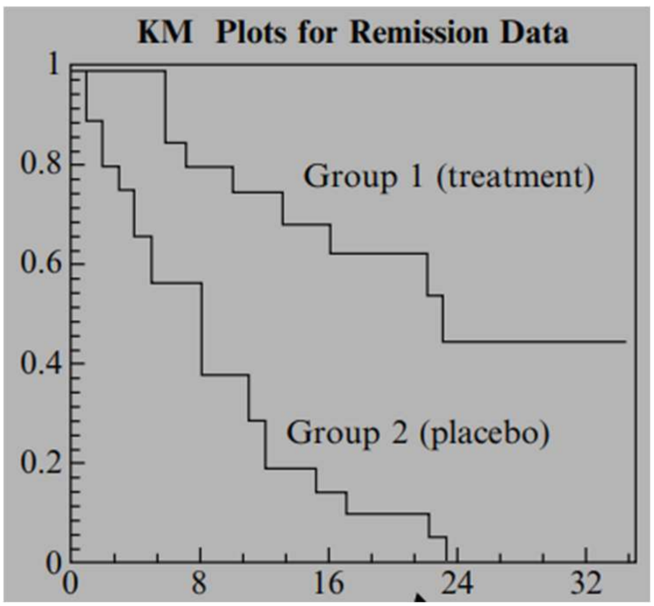
$Rx = 1, \log \text{WBC} = 2.93$:

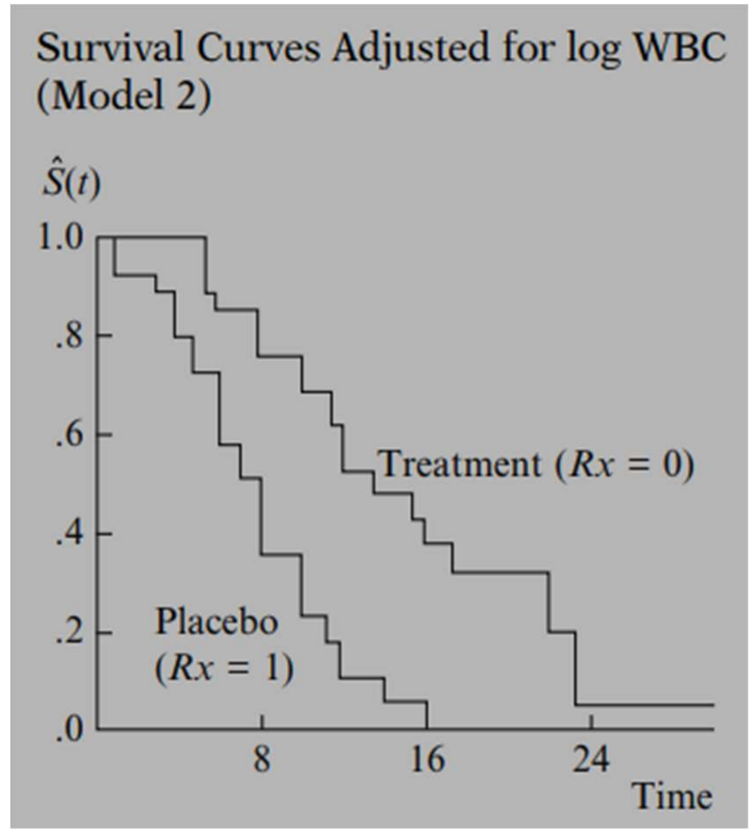
$$\begin{aligned}\hat{S}(t,\mathbf{X}) &= [\hat{S}_0(t)]^{\exp(\hat{\beta}_1\overline{Rx}+\hat{\beta}_2\overline{\log\ \text{WBC}})}\\ &= [\hat{S}_0(t)]^{\exp(1.294(0.5)+1.604(2.93))}\\ &= [\hat{S}_0(t)]^{\exp(5.35)} = \boxed{[\hat{S}_0(t)]^{210.6}}\end{aligned}$$

$Rx = 0, \log \text{WBC} = 2.93$:

$$\begin{aligned}\hat{S}(t,\mathbf{X}) &= [\hat{S}_0(t)]^{\exp(1.294(0)+1.604(2.93))}\\ &= [\hat{S}_0(t)]^{\exp(4.70)} = \boxed{[\hat{S}_0(t)]^{109.9}}\end{aligned}$$

Courbes de survie ajustées





Courbes de survie ajustées

Adjusted survival curves	KM curves
Adjusted for covariates	No covariates
Use fitted Cox model	No Cox model fitted

TP 3:

- Reprendre votre data de survie et appliquer le modèle Cox PH
- Vérifier si l'hypothèse de Proportional Hazard tient.