Analyse de Survie   M1 Proba-Stat. 2019-2020

Examen Final

Ex 1.

On dispose des données de survie de patients (guérison après un traitement en semaines) :

$$3 \quad 4 \quad 4^+ \quad 6 \quad 6 \quad 7 \quad 8^+ \quad 9 \quad 9^+ \quad 11$$

(3,5 pts) • Calculer les estimateurs de Kaplan-Meier S_{KM} et du risque $\hat{h}$ (tableau)

(1,5 pts) • Tracer les graphes.

Ex. 2.

On donne les données de survie de 2 groupes de patients (guérison en semaines) :

G1 : $5 \quad 5^+ \quad 7 \quad 7 \quad 8^+ \quad 10$ Tableau. (3,5 pts)

G2 : $3 \quad 5 \quad 6^+ \quad 8 \quad 8 \quad 9$

• Tester l'égalité des 2 survies des deux groupes G1 et G2 : $H_0 : S_1 = S_2$ à un niveau $\alpha = 5\%$ .. Conclure. (1,5 pts)

Ex. 3. Soit T une durée de survie de densité $f_T(t) = \dfrac{\theta}{(1+\theta t)^2}$, $t \geq 0$, $\theta > 0$ et C une v.a. de censure de loi uniforme $\mathcal{U}_{[0,\theta]}$. $T \perp\!\!\!\perp C$.

On pose $D = \mathbb{1}_{\{T \leq C\}}$ et $X = T \wedge C$.

(2 pts) (2 pts) 1. Donner la loi de D et la densité de X.

(1,5 pts) 2. Calculer $\mathbb{P}(X \leq x, D = 0)$

(2,5 pts) 3. Soit $(X_1, D_1) \dots, (X_n, D_n)$ un éch. de (X, D). Calculer la vraisemblance $L(X_1, \dots, X_n, \theta)$

(0,5 pts) Peut-on avoir l'emv $\hat{\theta}_n$?

(1,5 pts) 4. Pour (X_1, D_1) calculer $L(X_1, \theta)$. Peut-on avoir $\hat{\theta}_1$?

(0,5 pts)

1

Corrigé EF Analyse de Survie

28 Sept. 2020

Ex 1. (5 pts)

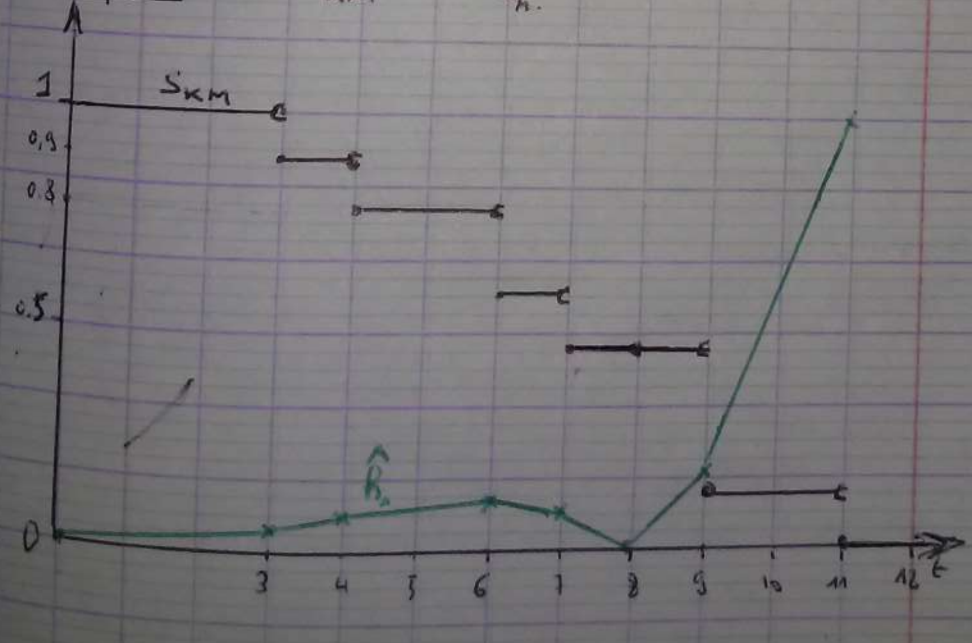
Tableau des valeurs de $\hat{h}_n(t_i)$ et $S_{KM}(t_i)$

(03,5 pts)

t_i	n_i	m_i	c_i	$\hat{h}_n(t_i)$	$S_{KM}(t_i)$
3	10	1	0	0,1	0,9
4	9	1	1	0,11	0,8
6	7	2	0	0,28	0,57
7	5	1	0	0,20	0,46
9	3	1	1	0,33	0,30
11	1	1	0	1	0

Graphes de S_{KM} et $\hat{h}_n$.

(1,5 pts)



2

Ex. 2 5 pts

Test LogRank.

t_k	$n_1(t_k)$	$n_2(t_k)$	$n(t_k)$	$d_1(t_k)$	$d_2(t_k)$	$d(t_k)$	$n_1(t_k)\frac{d}{n}$	$n_2(t_k)\frac{d}{n}$
3	0	6	6	0	1	1	0	1
5	6	5	11	1	1	2	1,09	0,91
6	0	4	4	0	0	0	0	0
7	4	0	4	2	0	2	2	0
8	0	3	3	0	2	2	0	2,0
9	0	1	1	0	1	1	0	1
10	1	0	1	1	0	1	1	0
				$O_1 = \sum_k d_1(t_k) = 4$	$O_2 = \sum d_2(t_k) = 5$		$E_1 = \sum_k n_1 \frac{d}{n} = 4,09$	$E_2 = \sum n_2 \frac{d}{n} = 4,91$

3,5 pts

On a : $T_n = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} = 3,63 \cdot 10^{-4} \approx 0,003$

Sous $H_0 : S_1 = S_2$: $T_n \rightsquigarrow \chi^2_1$. Pour $\alpha = 5\%$ on a $d_\alpha = 3,84$

La région de rejet de H_0 est : $D = \{ T_n > d_\alpha \} = \{ T_n > 3,84 \}$.

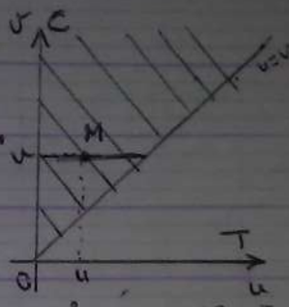
Comme le T_n observé est 0,003 ainsi on accepte $H_0 : S_1 = S_2$

1,5 pts

Final

(3)

Ex. 3 (10pts)



$\underline{v \text{ fixé}} \to u \in [0,v]$

1. $\underline{\text{Loi de } D}$: $D = \mathbb{1}_{\{T \le C\}} \sim \mathcal{B}(p)$, $p = \mathbb{P}(D=1) = \iint_{u \le v} f_T(u) f_C(v)\,du\,dv$.

$$p = \int_0^\theta \left(\int_0^v f_T(u)\,du \right) f_C(v)\,dv = \int_0^\theta \left(\int_0^v \frac{\theta}{(1+\theta u)^2}\,du \right) \frac{1}{\theta}\,dv$$

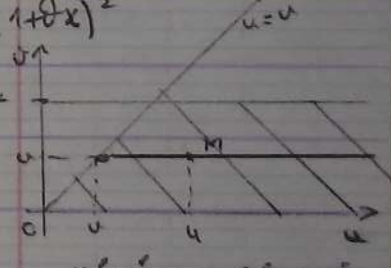
(2pts) Or $\int_0^v \frac{\theta}{(1+\theta u)^2}\,du = \left[-\frac{1}{(1+\theta u)} \right]_0^v = 1 - \frac{1}{1+\theta v}$. D'où $p = \frac{1}{\theta}\int_0^\theta \left(1 - \frac{1}{1+\theta v}\right) dv$

$$p = 1 - \frac{1}{\theta}\int_0^\theta \frac{1}{1+\theta v}\,dv = 1 - \frac{1}{\theta^2}\left[\log(1+\theta v)\right]_0^\theta = 1 - \frac{1}{\theta^2}\log(1+\theta^2)$$

$\underline{\text{Loi de } X}$: $F_X(x) = \mathbb{P}(X \le x) = 1 - \mathbb{P}(X \ge x) = 1 - S_T(x) S_C(x)$. Or $S_T(x) = \frac{1}{1+\theta x}$, $x \ge 0$.

(2pts) et $S_C(x) = \begin{cases} (1 - x/\theta), & 0 \le x \le \theta \\ 0, & x \ge \theta \end{cases}$. D'où $F_X(x) = \begin{cases} 1 - \left(\frac{1}{1+\theta x}\right)\left(1 - x/\theta\right), & 0 \le x \le \theta \\ 1, & x \ge \theta \end{cases}$.

Pour $0 \le x \le \theta$: $f_X(x) = F_X'(x) = \left(1 - \frac{1}{1+\theta x} + \frac{x}{\theta(1+\theta x)}\right)' = \frac{1+\theta^2}{\theta(1+\theta x)^2}$.



2. $\mathbb{P}(X \le x, D=0) = \mathbb{P}(C \le x, C \le T) = \iint_{v \le x,\, u \le v} f_T(u) f_C(v)\,du\,dv$

(1,5pts) $$= \int_0^x \left(\int_v^{+\infty} f_T(u)\,du \right) f_C(v)\,dv = \int_0^x \left(\int_v^{\infty} \frac{\theta}{(1+\theta u)^2}\,du \right) \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(v)\,dv$$

Or $\int_v^{+\infty} \frac{\theta}{(1+\theta u)^2}\,du = \left[-\frac{1}{1+\theta u} \right]_v^{+\infty} = \frac{1}{1+\theta v}$. D'où si $0 \le x \le \theta$ on a

$\underline{v \text{ fixé}} \to u \in [v, +\infty[$.

$\mathbb{P}(X \le x, D=0) = \int_0^x \frac{1}{1+\theta v} \cdot \frac{1}{\theta}\,dv = \frac{1}{\theta^2}\log(1+\theta x)$. Si $x > \theta$, $\mathbb{P}(X \le x, D=0) = \frac{1}{\theta^2}\log(1+\theta^2)$.

3. La vraisemblance $L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n \left(f_T(x_i,\theta) S_C(x_i,\theta)\right)^{D_i} \left(f_C(x_i,\theta) \cdot S_T(x_i,\theta)\right)^{1-D_i}$

$$= \prod_{i=1}^n \left(\frac{\theta}{(1+\theta x_i)^2} \left(1 - \frac{x_i}{\theta}\right) \mathbb{1}_{\{0 \le x_i \le \theta\}} \right)^{D_i} \left(\frac{1}{\theta} \cdot \frac{1}{1+\theta x_i} \cdot \mathbb{1}_{(0 \le x_i \le \theta)} \right)^{1-D_i} = \prod_{i=1}^n \left(\frac{\theta - x_i}{(1+\theta x_i)^2} \right)^{D_i} \left(\frac{1}{\theta(1+\theta x_i)} \right)^{1-D_i} \mathbb{1}_{0 \le x_i \le \theta}$$

(2,5pts) $\log L(x_1, \dots, x_n, \theta) = \sum_{i=1}^n D_i \left[\log(\theta - x_i) - 2\log(1+\theta x_i)\right] + (1-D_i)\left[-\log\theta - \log(1+\theta x_i)\right]$. pour $0 \le x_i \le \theta$.

(0,5pt) EMV $\hat{\theta}_n$: $0 = \frac{\partial \log L}{\partial \theta} = \sum_{i=1}^n D_i \left(\frac{1}{\theta - x_i} - \frac{2x_i}{1+\theta x_i} \right) - (1-D_i)\left(\frac{1}{\theta} + \frac{x_i}{1+\theta x_i} \right)$

Équation difficile à résoudre en θ.

4. Pour (X_1, D_1) on a $0 = \frac{\partial \log L}{\partial \theta}(X_1, \theta) = D_1 \left(\frac{1}{\theta - X_1} - \frac{2X_1}{1+\theta X_1} \right) - (1-D_1)\left(\frac{1}{\theta} + \frac{X_1}{1+\theta X_1} \right)$; $0 \le X_i \le \theta$

(1,5pt) ou $\theta^2 (D_1 - 2) X_1 + \theta\left((2D_1 - 1)(1 - 2X_1^2)\right) + (1 - D_1) X_1 = 0$

(0,5pt) Équation du 2nd degré à résoudre pour trouver $\hat{\theta}_n$.