Analyse de Survie M1 Proba - Stat. 2019-2020

Examen de Rattrapage (9Nov 2020)

Ex 1. (3,5 pts) Soit T une durée de survie de fonction de survie : $S_T(t) = \begin{cases} 1 & \text{si } t \leq 1 \\ e^{-\theta(t-1)} & \text{si } t \geq 1 \end{cases}$

a) Calculer $\mathbb{E}(T)$ et $\mathbb{E}(T-t / T>t)$ pour $t \geq 0$.

b) Calculer la densité f_T et la fonction de risque h_T.

Ex. 2. (3,5 pts) On dispose des données de survie de 2 groupes de patients (guérison en semaines) :

G1 :	5	6	6^+	7	8	9^+	10
G2 :	4	5	5^+	6	7^+	8	9

Tester l'égalité des 2 survies des 2 groupes G1 et G2 : $H_0 : S_1 = S_2$ à un niveau $\alpha = 5\%$. Conclure.

Ex 3. (13 pts) Soit T une durée de survie de densité : $f_T(t) = \dfrac{2\theta t}{(1+\theta t^2)^2}$, $t \geq 0$, $\theta > 0$ et C une v.a. de censure de loi uniforme $\mathcal{U}_{[0,\theta]}$ et $T \perp\!\!\!\perp C$.

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

(4 pts) 1. Calculer la loi de D et la densité de X.

(4,5 pts) 2. Calculer $P(X \leq x, D=0)$. X et D sont-elles indépendantes.

3. On observe un échantillon avec censure $(X_1, D_1), \ldots, (X_n, D_n)$ de (X, D).

(3 pts) Calculer la fonct. de vraisemblance $L(X_1, \ldots, X_n, \theta)$

(0,5 pt) Peut-on calculer l'EMV $\hat{\theta}_n$?

4. Pour une observation (X_1, D_1), calculer $L(X_1, D_1, \theta)$.

(1 pt) Peut-on avoir l'EMV $\hat{\theta}_1$?

Corrigé Rattrapage (9 Nov 2020)
Analyse de Survie

Ex1. $S_T(t) = \begin{cases} 1 & \text{si } t \leq 1 \\ e^{-\theta(t-1)} & \text{si } t > 1. \end{cases}$

a) $E(T) = \int_0^\infty S_T(u)\,du = \int_0^1 du + \int_1^\infty e^{-\theta(u-1)}\,du = 1 + \frac{1}{\theta}$.

$E(T-t / T>t) = \frac{1}{S(t)} \int_t^\infty S_T(u)\,du$. Or $\int_t^\infty S_T(u)\,du = \begin{cases} \int_t^1 S(u)\,du + \int_1^\infty S(u)\,du & \text{si } t \leq 1 \\ \int_t^\infty S(u)\,du & \text{si } t \geq 1 \end{cases}$

$= \begin{cases} (1-t) + \left[-\frac{e^{-\theta(t-1)}}{\theta}\right]_1^\infty = (1-t) + \frac{1}{\theta} & \text{si } t \leq 1. \\ \frac{e^{-\theta(t-1)}}{\theta}, & \text{si } t \geq 1 \end{cases}$

D'où $E(T-t / T>t) = \begin{cases} (1-t) + \frac{1}{\theta} & \text{si } t < 1 \\ 1/\theta & \text{si } t \geq 1. \end{cases}$

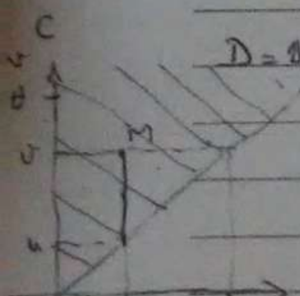
b) $f_T(t) = -S_T'(t) = \begin{cases} 0 & \text{si } t \leq 1 \\ \theta e^{-\theta(t-1)}, & \text{si } t > 1 \end{cases}$; $h_T(t) = \frac{f_T(t)}{S_T(t)} = \begin{cases} 0 & \text{si } t \leq 1 \\ \theta & \text{si } t > 1 \end{cases}$

Ex3 1) $f_T(t) = \frac{2\theta t}{(1+\theta t^2)^2}$, $t \geq 0, \theta > 0$, $C \sim \mathcal{U}_{[0,\theta]}$

On a : $F_C(t) = \begin{cases} t/\theta, & 0 \leq t \leq \theta \\ 1 & \text{si } t \geq 0 \\ 0 & \text{si } t < \theta \end{cases}$

$F_T(t) = 1 - \frac{1}{1+\theta t^2}$, $S_T = \frac{1}{1+\theta t^2}$

$S_C(t) = \begin{cases} 1, & \text{si } t \leq 0 \\ 1 - t/\theta, & 0 < t \leq \theta \end{cases}$



$D = \mathbb{1}_{(T \leq C)} \sim \mathcal{B}(p)$ où $p = P(D=1) = \int S_C(u) f_T(u)\,du$

$p = \iint_D f_{T,C}(u,v)\,du\,dv = \int_0^\infty \Big(\underbrace{\int_u^{\infty} f_C(v)\,dv}_{S_C(u)}\Big) f_T(u)\,du$ (en fixant u!)

ou bien en fixant v! on a :

$p = \int_0^\theta \left(\int_0^v f_T(u)\,du\right) f_C(v)\,dv$, $f_C(v) = \frac{1}{\theta}\mathbb{1}_{(0,\theta)}(v)$

On a : $\int_0^v f_T(u)\,du = \int_0^v \frac{2\theta u}{(1+\theta u^2)^2}\,du = \int_0^v \frac{g'}{g^2} = \left[-\frac{1}{g}\right]_0^v = 1 - \frac{1}{1+\theta v^2}$

D'où $p = \int_0^\theta \left(1 - \frac{1}{1+\theta v^2}\right)\frac{1}{\theta}\,dv = \int_0^\theta \frac{1}{\theta}\,dv - \frac{1}{\theta}\int_0^\theta \frac{1}{1+\theta v^2}\,dv = 1 - \frac{1}{\theta^{3/2}} \operatorname{Arctg} \theta^{3/2}$

$x = \sqrt{\theta}\, v$

loi de X : $F_X(x) = 1 - S_C(x) S_T(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \left(1 - \frac{x}{\theta}\right)\left(\frac{1}{1+\theta x^2}\right) & , 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}$

$f_X(x) = F_X'(x) = \frac{1}{\theta}\left(\frac{1}{1+\theta x^2}\right) - \frac{2x}{1+\theta x^2}(x-\theta)$, si $0 \leq x \leq \theta$.

2) $P(X < x, D = 0) = \int_0^{\infty} \left(\int_{v}^{\infty} f_T(u)\,du\right) f_C(v)\,dv = \frac{1}{\theta}\int_0^{x \wedge \theta} \frac{1}{1+\theta u^2}\,du = \begin{cases} \frac{1}{\theta^{3/2}} \operatorname{Arctg} \sqrt{\theta}\, x, & \text{si } 0 \leq x \leq \theta. \\ \frac{1}{\theta^{3/2}} \operatorname{Arctg} \theta^{3/2} & \text{si } x > \theta. \end{cases}$

Donc $P(X \leq x, D = 0) \neq P(X \leq x) \cdot P(D = 0)$ donc $X \not\perp\!\!\!\perp D$

3) La vraisemblance d'un échantillon $(X_1, D_1) \ldots (X_n, D_n)$:

$$L(X_1, \ldots, X_n, \theta) = \prod_{i=1}^{n} \left[f_T(X_i, \theta) S_C(X_i, \theta)\right]^{D_i} \left[f_C(X_i, \theta) \cdot S_T(X_i, \theta)\right]^{1-D_i}$$

Après remplacement on a :

$$L(X_1 \ldots X_n, \theta) = \prod_{i=1}^{n} \left(\frac{2\theta X_i}{(1+\theta X_i^2)^2} \mathbb{1}_{(X_i > 0)} \left(1 - \frac{X_i}{\theta}\right) \mathbb{1}_{(0 \leq X_i \leq \theta)}\right)^{D_i} \left(\frac{1}{\theta} \cdot \mathbb{1}_{(0 \leq X_i \leq \theta)} \cdot \frac{1}{1+\theta X_i^2} \cdot \mathbb{1}_{(X_i \geq 0)}\right)^{1-D_i}$$

$$L(X_1 \ldots X_n, \theta) = \prod_{i=1}^{n} \left(\frac{2X_i(\theta - X_i)}{(1+\theta X_i^2)} \mathbb{1}_{(0 \leq X_i \leq \theta)}\right)^{D_i} \left(\frac{1}{\theta(1+\theta X_i^2)} \cdot \mathbb{1}_{(0 \leq X_i \leq \theta)}\right)^{1-D_i}$$

Calcul de l'EMV :

Pour $0 \leq X_i \leq \theta$: $\frac{\partial \log L}{\partial \theta} = 0 \Rightarrow \sum \left(-\frac{1}{\theta} - \frac{X_i^2}{1+\theta X_i^2} + \frac{D_i}{\theta} - \frac{D_i X_i^2}{1+\theta X_i^2}\right) = 0$

$= \sum_{i=1}^{n} \frac{-2\theta X_i^2 + D_i - 1}{\theta(1+\theta X_i^2)} = 0$ équation difficile à résoudre pour trouver l'EMV $\hat{\theta}_n$!

4) Si on observe (X_1, D_1) on a :

$$\frac{\partial \log L}{\partial \theta} = \frac{-2\theta X_1 + D_1 - 1}{\theta(1+\theta X_1^2)} = 0 \Rightarrow -2\theta X_1 + D_1 - 1 = 0$$

d'où $\hat{\theta}_1 = \frac{D_1 - 1}{2X_1^2}$.

Ex2. On a : $T_n = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} = 0{,}04 + 0{,}05 = 0{,}09$.

$\alpha = 5\%$ $d_\alpha = 3{,}84$. On a : $T_n < d_\alpha \Rightarrow$ On accepte H_0 $S_1 = S_2$.

Ex3.

t_i	$n_1(t_i)$	$n_2(t_i)$	$n(t_i)$	$d_1(t_i)$	$d_2(t_i)$	$d(t_i)$	$n_1 \frac{d}{n}$	$n_2 \frac{d}{n}$
4	0	7	7	—	1	1	0	1
5	7	6	13	1	1	2	1,07	0,92
5^+	—	5	5	—	—	0	—	—
6	6	4	10	1	1	2	1,2	0,8
6^+	5	—	5	—	—	0	—	—
7	4	—	4	1	—	1	1	—
7^+	—	3	3	—	—	~~0~~	—	—
8	3	2	5	1	1	2	1,2	0,8
9	—	1	1	—	1	1	—	1
9^+	2	—	2	—	—	0	—	—
10	1	—	1	1	—	1	1	0
				$\sum d_1(t_i) = 5$ O_1	$\sum d_2(t_i) = 5$ O_2		$E_1 = 5{,}47$	$E_2 = 4{,}52$