Corrigé TD n° 3

II) $\hat{h}(t_i) = \frac{m_i}{n_i}$ m_i = nbre de "événements" à t_i , n_i = nbre de "individus à risque" à t_i^- (avant t_i) (i.e. à t_{i-1} + données censurées)

($n_i = n_{i-1} - m_{i-1} - c_{i-1}$

nbre de censures à t_{i-1} t_{i-1} t_i

t_i	n_i	m_i	$\hat{h}(t_i)$	$\hat{H}_{NA}(t_i)$	$\hat{S}_{KM}(t_i)$
4.5	30	1	0,033	0,033	0,967
11,5	28	2	0,071	0,104	0,89
16	25	1	0,04	0,144	0,862
18,5	23	4	0,174	0,318	0,712
20,3	18	2	0,111	0,429	0,633
20,7	15	2	0,133	0,56	0,54
22	12	1	0,083	0,645	0,50
30,0	11	4	0,364	1,009	0,320
32,0	6	1	0,167	1,176	0,267
34,5	5	1	0,2	1,37	0,217
37,5	4	2	0,5	1,87	0,107
41,5	2	1	0,5	2,37	0,053

$$\hat{H}_{NA}(t_i) = \sum_{j: t_j \le t_i} \frac{m_j}{n_j} \quad , \quad \hat{S}_{KM}(t_i) = \prod_{j: t_j \le t_i} \left(1 - \frac{m_j}{n_j}\right)$$

$\hat{S}_A(a_i) = \prod\left(1 - \frac{m_{i-1}}{e_i}\right)$

Méthode actuarielle

$e_i = n_{i-1} - \frac{c_{i-1}}{2}$

$]a_{i-1}, a_i]$	n_{i-1}	m_{i-1}	c_{i-1}	e_i	$\hat{S}_A(a_i)$	$\hat{h}$
]0,4]	30	0	0	30	1	[illegible]
]4,8]	30	1	1	29,5	0,966	0,01[illegible]
]8,12]	29	2	0	29	0,899	0,017
]12,16]	27	1	1	26,5	0,865	0,010
]16,20]	26	4	2	25	0,723	0,04
]20,24]	22	5	2	21	0,55	0,0[illegible]
]24,28]	17	0	0	17	0,554	0,0[illegible]
]28,32]	17	5	1	16,5	0,38	[illegible]
]32,36]	12	1	0	12	0,55	0,[illegible]
]36,40]	11	2	0	11	0,28	[illegible]
]40,44]	9	1	1	8,5	0,5	[illegible]

$$\hat{h}_A\left(\frac{a_i + a_{i-1}}{2}\right) = \frac{2 \cdot \frac{m_{i-1}}{e_i}}{(a_i - a_{i-1})\left(2 - \frac{m_{i-1}}{e_i}\right)}$$

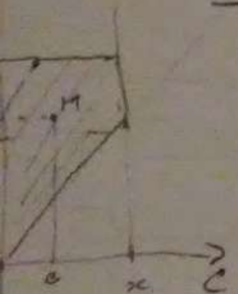
$, t \geq a \qquad \Big| = \frac{f_T(t)}{S_T(t)} \Big(\lambda \qquad , t \geq a$

Ex 2. 1)

$T \sim \mathcal{U}_{[0,\theta]}$, $C \sim \mathcal{E}(\theta)$, $X = T \wedge C$, $D = \mathbb{1}_{(T \leq C)}$, $f_T(t) = \frac{1}{\theta}\mathbb{1}_{(0,\theta)}(t)$, $f_C(t) = \theta e^{-\theta t}$, $t \geq 0$

On a : $D \sim \mathcal{B}(p)$; $p = \mathbb{P}(D=1) = \mathbb{P}(T \leq C) = \iint\limits_{u \leq v} f_{(T,C)}(u,v)\,du\,dv = \iint\limits_{u \leq v} f_T(u) f_C(v)\,du\,dv = \iint\limits_{u \leq v} \frac{1}{\theta}\mathbb{1}_{(0,\theta)}(u) \cdot \theta e^{-\theta v}\mathbb{1}_{(0,+\infty)}(v)\,du\,dv$

$$= \int_0^\theta \Big[\int_u^{+\infty} e^{-\theta v}\,dv\Big]\,du = \int_0^\theta \Big[-\frac{1}{\theta}e^{-\theta v}\Big]_u^{+\infty} du = \int_0^\theta \frac{1}{\theta}e^{-\theta u}\,du = \frac{1}{\theta^2}\Big[-e^{-\theta u}\Big]_0^\theta = \boxed{\frac{1}{\theta^2}\big(1 - e^{-\theta^2}\big) = p}$$



Loi de X

$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(T \wedge C \leq x) = 1 - \mathbb{P}(T \wedge C \geq x) = 1 - S_T(x)\,S_C(x)$

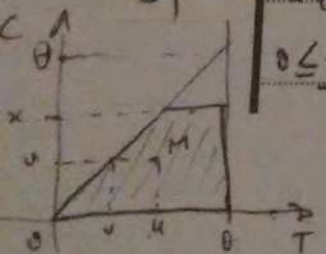
$F_T(x) = \begin{cases} 0, & x \leq 0 \\ x/\theta, & 0 \leq x \leq \theta \\ 1 & \text{si } x \geq \theta \end{cases}$, $F_C(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\theta x}, & \text{si } x \geq 0 \end{cases}$

$S_T(x) = \begin{cases} 1, & x \leq 0 \\ 1 - \frac{x}{\theta}, & 0 \leq x \leq \theta \\ 0 & \text{si } x \geq \theta \end{cases}$, $S_C(x) = \begin{cases} 1, & x \leq 0 \\ e^{-\theta x}, & x \geq 0 \end{cases}$

D'où $F_X(x) = \begin{cases} 0, & \text{si } x \leq 0 \\ 1 - \big(1 - \frac{x}{\theta}\big)e^{-\theta x}, & 0 \leq x \leq \theta \\ 1, & x \geq \theta \end{cases}$

et $f_X(x) = \begin{cases} 0, & x \notin [0,\theta] \\ \big(\theta - x + \frac{1}{\theta}\big)e^{-\theta x} & \text{si } x \geq 0 \end{cases}$

2)



$\mathbb{P}(X \leq x, D = 0) = \mathbb{P}(T \wedge C \leq x, T > C) = \mathbb{P}(C \leq x, T > C) = \iint f_T(u) f_C(v)\,du\,dv$

$\underline{0 \leq x \leq \theta}$

$$= \int_0^x \Big[\int_v^\theta f_T(u)\,du\Big] f_C(v)\,dv = \int_0^x \frac{1}{\cancel{\theta}}(\theta - v)\,\cancel{\theta} e^{-\theta v}\,dv = \int_0^x \theta e^{-\theta v}\,dv - \int_0^x v e^{-\theta v}\,dv = \cancel{\theta}\big(1 - e^{-\theta x}\big) - \int_0^x v e^{-\theta v}\,dv$$

Or $\int_0^x v e^{-\theta v}\,dv = \Big[-\frac{v}{\theta}e^{-\theta v}\Big]_0^x + \frac{1}{\theta}\int_0^x e^{-\theta v}\,dv = -\frac{x}{\theta}e^{-\theta x} + \frac{1}{\theta^2}\big(1 - e^{-\theta x}\big)$.

Donc $\mathbb{P}(X \leq x, D = 0) = \cancel{\theta}\big(1 - e^{-\theta x}\big) + \frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^2}\big(1 - e^{-\theta x}\big)$
$= \cancel{\theta}\big(1 - \frac{1}{\theta^2}\big)\big(1 - e^{-\theta x}\big) + \frac{x}{\theta}e^{-\theta x}$

$\mathbb{P}(X \leq x, D=0) = \frac{x}{\theta}\left(1-\frac{1}{\theta^2}\right)\left(1-e^{-\theta x}\right) + \frac{x}{\theta} e^{-\theta x}$; si $0 \leq x \leq \theta$.

Pour $0 \leq x \leq \theta$: $\mathbb{P}(X \leq x, D=1) = \mathbb{P}(X \leq x) - \mathbb{P}(X \leq x, D=0)$

$$= \left[1-\left(1-\frac{x}{\theta}\right)e^{-\theta x}\right] - \left[\frac{x}{\theta}\left(1-\frac{1}{\theta^2}\right)\left(1-e^{-\theta x}\right) + \frac{x}{\theta} e^{-\theta x}\right]$$

$$= 1 - e^{-\theta x} - \frac{x}{\theta}\left(1-\frac{1}{\theta^2}\right)\left(1-e^{-\theta x}\right)$$

A-t-on $X \perp\!\!\!\perp D$?

si $\mathbb{P}(X \leq x, D=0) \overset{?}{=} \mathbb{P}(X \leq x) \cdot \mathbb{P}(D=0)$

$0 \leq x \leq \theta$: $\frac{x}{\theta}\left(1-\frac{1}{\theta^2}\right)\left(1-e^{-\theta x}\right) + \frac{x}{\theta} e^{-\theta x} \overset{??}{=} \left(1-\left(1-\frac{x}{\theta}\right)e^{-\theta x}\right)\left(1-\frac{1}{\theta^2}\left(1-e^{-\theta^2}\right)\right)$

Or le 2ème membre $= \left(\left(1-e^{-\theta x}\right) + \frac{x}{\theta} e^{-\theta x}\right)\left(1-\frac{1}{\theta^2}\left(1-e^{-\theta^2}\right)\right)$

$$= \underbrace{\left(1-\frac{1}{\theta^2}\left(1-e^{-\theta^2}\right)\right)}_{(*) \neq \left(1-\frac{1}{\theta^2}\right)}\left(1-e^{-\theta x}\right) + \underbrace{\left(1-\frac{1}{\theta^2}\left(1-e^{-\theta^2}\right)\right)}_{(**) \neq 1}\left(\frac{x}{\theta} e^{-\theta x}\right)$$

Car 2 facteurs $(*)$ et $(**)$ sont différents des facteurs de le 1er membre

Donc $X \not\perp\!\!\!\perp D$.

). On dispose des données de survie de 2 groupes de patients (guérison en semaines pour 2 traitements) :

G1 : $5 \quad 5^+ \quad 6 \quad 7 \quad 9 \quad 10 \quad 10^+ \quad 11^+ \quad 12 \quad 14^+$

G2 : $3 \quad 3 \quad 5^+ \quad 6 \quad 8 \quad 9 \quad 10 \quad 11^+ \quad 13 \quad 14$

Tester l'égalité des survies : $H_0 : S_1 = S_2$ (des 2 groupes) par le test du LogRank au seuil de $\alpha = 5\%$. Conclure.

Test logRank :

t_k	$n_1(t_k)$	$n_2(t_k)$	$n(t_k)$	$d_1(t_k)$	$d_2(t_k)$	$d(t_k)$	$n_1 \frac{d(t_k)}{n(t_k)}$	$n_2 \frac{d(t_k)}{n(t_k)}$
3	-	10	10	-	2	2	0	2
5	10	-	10	1	-	1	1	0
5^+	9	8	17	0	0	0	0	0
6	8	7	15	1	1	2	1,06	0,93
7	7	-	7	1	-	1	1	0
8	-	6	6	-	1	1	0	1
9	6	5	11	1	1	2	1,09	0,90
10	5	4	9	1	1	2	1,11	0,88
10^+	4	-	4	0	-	0	0	0
11^+	3	3	6	0	0	0	0	0
12	2	-	2	1	-	1	1	0
13	-	2	2	-	1	1	0	1
14	-	1	1	-	1	1	0	1
14^+	1	-	1	0	-	0	0	0
				$O_1 = \sum d_1(t_k) = 6$	$O_2 = \sum d_2(t_k) = 8$		$E_1 = 6,26$	$E_2 = 7,71$

$n_i(t_k)$ = nbr ind à risque à t_k^- du Groupe i

$d_i(t_k)$ = nbr d'évt observé à t_k du Groupe i

Test de $H_0 : S_1 = S_2$.

$$T_n = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

$$= 0,010 + 0,010$$

$$\boxed{T_n = 0,02}$$

Sous H_0 $T_n \sim \chi^2_1$

$\alpha = 5\%$ → table $d_\alpha = 3,84$

Donc on accepte H_0 égalité de 2 survies.