Corrigé "Stat. inf" 16/10/2020

Exercice 1

1) Déf d'une loi de type exponentiel : $f_\theta(x)$ peut s'écrire sous une des 2 formes :

$$f_\theta(x) = \exp\left(C(\theta)\cdot T(x) + D(\theta) + S(x)\right) \qquad (1)$$

ou

$$f_\theta(x) = \exp\left(C(\theta)\cdot T(x) + D(\theta)\right)\cdot \tilde{S}(x) \qquad (2)$$

On montre que $f_\theta(x)$ peut s'écrire sous la forme (2) :

$$f_\theta(x) = \exp\left(\log 2 - \log\theta - \frac{x^2}{\theta}\right) x \mathbb{1}_{x>0}$$

avec :
$$\begin{cases} C(\theta) = -\frac{1}{\theta}, \quad T(x) = x^2 \\ D(\theta) = -\log\theta + \log 2 \\ \tilde{S}(x) = x\mathbb{1}_{x>0} \end{cases}$$

2) On vérifie d'abord que l'ensemble $A = \{x \,;\, f_\theta(x) > 0\}$ ne dépend pas de θ.

On a :

$$A = \{x \,;\, f_\theta(x) > 0\} =]0, \infty[$$ ne dépend pas de θ.

La vraisemblance :

$$L_n(\theta) = \prod_{i=1}^n f_\theta(x_i) = \frac{2^n}{\theta^n}\prod_{i=1}^n x_i \cdot \exp\left(-\frac{1}{\theta}\sum_{i=1}^n x_i^2\right)\mathbb{1}_{\min x_i > 0}$$

La log-vraisemblance sur A :

$$\log L_n(\theta) = n\log 2 - n\log\theta - \sum_{i=1}^n \log x_i - \frac{1}{\theta}\sum_{i=1}^n x_i^2$$

$$\frac{\partial}{\partial\theta}\log L_n(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n x_i^2 = 0 \implies \theta = \frac{1}{n}\sum_{i=1}^n x_i^2$$

Donc, l'estimateur du maximum de vraisemblance est :

$$\boxed{\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n X_i^2}$$

3) $\mathbb{E}[Y] = \int_{\mathbb{R}} y\, g(y)\, dy = \int_{0}^{\infty} y \frac{1}{\theta} e^{-y/\theta} dy$

$= -y e^{-y/\theta}\Big]_0^{\infty} + \int_0^{\infty} e^{-y/\theta} dy$

$= -\theta e^{-y/\theta}\Big]_0^{\infty} = \theta$

4) Puisque la loi de X est de type exponentiel, alors:

$\frac{1}{n}\sum_{i=1}^{n} T(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ est un estimateur sans biais et efficace pour $\mathbb{E}[T(X)] = \mathbb{E}[X^2] = \mathbb{E}[Y] = \theta$

$\sum_{i=1}^{n} T(X_i) = \sum_{i=1}^{n} X_i^2$ est exhaustif pour θ. $\Rightarrow \frac{1}{n}\sum X_i^2 = \hat{\theta}_n$ exhaustif

5) Par LFGN: $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} T(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p.s} \mathbb{E}[X^2] = \mathbb{E}[Y] = \theta$

$\Rightarrow \hat{\theta}_n$ fortement convergent pour θ.

6) On considère les moments empiriques et théoriques d'ordre 1 pour $(Y_1, \dots, Y_n)$: $\overline{Y}_n = \mathbb{E}[Y]$

$\overline{Y}_n = \theta \Rightarrow \hat{\theta} = \overline{Y}_n$

Donc, un estimateur par la méthode des moments:

$\tilde{\theta}_n = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \hat{\theta}_n$.

Exercice 2

La variable aléatoire: $X : \begin{cases} 1 & \text{longueur boulon} = 5\text{cm} \\ 0 & \neq \end{cases}$

$X \sim Ber(p)$, p = la proba que la longueur d'un boulon soit $= 5m$.

On veut tester: le taux des boulons fabriqués d'une longueur de 5cm est $> 90\%$: $p > 0.90$

③

Donc les hypothèses à tester :

$H_0 : p \leq 0,90$ $\qquad$ $H_1 : p > 0,90$

$p_0 = 0,90$

On a test sur la proportion (probabilité) d'une loi de Bernoulli.

Statistique de test : $Z = \sqrt{n} \dfrac{\bar{X}_n - 0,9}{\sqrt{0,9 \cdot 0,1}}$

$\bar{x}_n = 1 - \dfrac{5+7+8+15}{400} = 0,9125$ $\qquad$ $n = 400$

Zone de rejet : $R = \{(x_1 .., x_n) / z > u_{1-\alpha}\}$ avec $u_{1-\alpha}$ fractile d'ordre 0,95 de la loi $N(0,1)$:

$u_{1-\alpha} = u_{0,95} = 1,6449$

Sur nos données : $z = 0,83 \notin R \Rightarrow H_0$ acceptée

Donc le taux de boulons = 5cm n'est pas strictement supérieur a

3) X : durée de guérison d'un patient $\sim N(m, \sigma^2)$

On doit donner l'intervalle de confiance pour m quand σ^2 est inconnu :

$$\left[\bar{x}_n - \frac{s_n^*}{\sqrt{n}} u_{1-\frac{\alpha}{2}} , \bar{x}_n + \frac{s_n^*}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right] \qquad n = 16$$

avec $\bar{x}_n = \dfrac{1}{n} \sum_{i=1}^{n} x_i = 22,3125$

$s_n^{*2} = \dfrac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 = 22,09$

$u_{1-\frac{\alpha}{2}} = u_{0,975}$ fractile de la loi $t(15)$

$= 2,131$

Alors, l'IC est :

$[19,80 \ ; \ 24,816]$