Université de Pau et des Pays de l'Adour | **Semestre printemps 2023–2024**

L3 Math | Introduction à la Statistique

CC 1 - Estimation statistique

Question de cours (3pts)

Soit $(\mathcal{H}^n, (Q_\theta)_{\theta\in\Theta})$ un modèle statistique. On note $\hat{\theta}_n$ un estimateur du véritable paramètre inconnu θ^*.

1. (1pt) Qu'est-ce qu'un estimateur sans biais ?
2. (2pts) Pour le modèle $(\{0,1\}^n, (\mathcal{B}(p)^{\otimes n})_{p\in(0,1)})$, montrez que l'estimateur suivant est sans biais

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i.$$

Exercice 1 Méthode des moments (3pts)

Soit $(\mathbb{R}_+^n, (\mathbb{P}^{\otimes n}_{a\in\mathbb{R},b\in\mathbb{R},c>0}))$ un modèle statistique vérifiant pour $X \sim \mathbb{P}_{a,b,c}$

$$\mathbb{E}[X] = a + b - c^2;\ \mathrm{Var}(X) = a - b + c^2;\ \mathbb{E}[X^4] = c^2.$$

Déterminer un estimateur de (a^*, b^*, c^*) à l'aide de la méthode des moments.

Exercice 2 Estimation du paramètre de la loi de Weibull (15pts)

Soit X une variable aléatoire réelle suivant la loi de Weibull de densité : pour tout $x > 0$

$$f(x) = \frac{e^{-\frac{\sqrt{x}}{\beta}}}{2\beta\sqrt{x}}$$

où β est un paramètre strictement positif inconnu.

Partie 1 : Estimation classique du paramètre

1. (1pt) Écrire le modèle statistique dans le cas où les observations sont indépendantes.
2. (2pts) Montrer que pour $\beta > 0$, on a

$$\mathbb{E}[X] = 2\beta^2.$$

En déduire un estimateur $\hat{\beta}_{MM}$ par la méthode des moments.

3. (1.5pt) Soit $x_1, \ldots, x_n \in \mathbb{R}_+^*$, déterminer la vraisemblance du modèle et montrer que pour tout $\beta \in \mathbb{R}_*^+$:

$$\frac{\partial \log L_n}{\partial \beta}(x_1, \ldots, x_n; \beta) = -\frac{n}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^n \sqrt{x}_i.$$

On rappelle qu'on utilise la notation log pour le logarithme népérien.

4. (1.5pts) Montrer que l'estimateur du maximum de vraisemblance du modèle est

$$\hat{\beta}_{MV} = \frac{1}{n}\sum_{i=1}^n \sqrt{X_i}.$$

On justifiera bien le fait qu'il s'agit d'un maximum.

5. (3pts) Déterminer le biais de l'EMV, puis sans calculer explicitement le biais, montrer que l'estimateur $\hat{\beta}_{MM}$ est biaisé.
6. (1pt) Montrer que les estimateurs $\hat{\beta}_{MM}$ et $\hat{\beta}_{MV}$ sont fortement consistants.
7. (1pt) Déterminer le risque de l'EMV.

Partie 2 : Estimation à l'aide des statistiques d'ordre

8. (1.5pts) Déterminer la fonction de répartition F_X de la variable X. En déduire que

$$F_X^{-1}(0.5) = (\beta \log 2)^2.$$

Interpréter le terme $F_X^{-1}(0.5)$.
9. (1.5pts) On range de manière croissante notre échantillon, et notons

$$X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)},$$

le nouveau réarrangement. Montrer que presque sûrement, ces inégalités sont strictes.
10. (1pt) À l'aide du réarrangement, déterminer un nouvel estimateur de β.

Rappel du cours de probabilité : On pourra utiliser sans preuve les résultats suivants :

1. Si $X \sim \mathcal{B}(p)$ pour $p \in [0,1]$ alors $\mathbb{E}[X] = p$ et $\mathrm{Var}(X) = p(1-p)$.
2. **La Loi des Grands Nombres** : Soit X une variable aléatoire possèdant un moment d'ordre 1 et soit $(X_n)_{n\in\mathbb{N}}$ une suite i.i.d. de variables aléatoires de même loi X, on a

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p.s.} \mathbb{E}[X].$$

3. **Le Théorème Central Limite** : Si de plus, X possède un moment d'ordre 2 alors

$$\sqrt{n}\frac{\bar{X}_n - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}} \xrightarrow{loi} Z \sim \mathcal{N}(0,1).$$

4. **Inégalité de Jensen** : Soit X une variable aléatoire non constante ayant un moment d'ordre 1, pour φ une fonction strictement convexe, si $\varphi(X)$ possède un moment d'ordre 1 alors

$$\varphi(\mathbb{E}[X]) < \mathbb{E}[\varphi(X)].$$