

EXERCICES DU CHAPITRE 16

ESTIMATIONS PONCTUELLE ET PAR INTERVALLE DE CONFIANCE

EXERCICE 1 – ●●● – Maximum de vraisemblance

1. On cherche l'estimateur du maximum de vraisemblance pour le paramètre λ d'une variable aléatoire X suivant la loi $\mathcal{P}(\lambda)$. Soient $n \in \mathbb{N}$ et $(X_1, \dots, X_n)$ un n -échantillon de X .

Soit $(x_1, \dots, x_n)$ une réalisation de $(X_1, \dots, X_n)$. On pose $L_n : \lambda \mapsto \prod_{k=1}^n \mathbb{P}(X_k = x_k)$.

- 1.a. Rappeler, pour tout $i \in \mathbb{N}$, la probabilité $\mathbb{P}(X = i)$.
 - 1.b. Démontrer que la fonction $\lambda \mapsto \ln(L_n(\lambda))$ possède un maximum sur $\mathbb{R}^{++}$ atteint en un unique réel que l'on déterminera en fonction de $x_1, \dots, x_n$ et n .
 - 1.c. Conclure sur l'estimateur du maximum de vraisemblance recherché.
2. On cherche l'estimateur du maximum de vraisemblance pour le paramètre λ d'une variable aléatoire X suivant la loi $\mathcal{E}(\lambda)$. Soient $n \in \mathbb{N}$ et $(X_1, \dots, X_n)$ un n -échantillon de X .
- Soit $(x_1, \dots, x_n)$ une réalisation de $(X_1, \dots, X_n)$. On pose $L_n : \lambda \mapsto \prod_{k=1}^n f_\lambda(x_k)$, où f_λ désigne la densité de X continue sur $\mathbb{R}^+$.
- 2.a. Démontrer que la fonction $\lambda \mapsto \ln(L_n(\lambda))$ possède un maximum sur $\mathbb{R}^{++}$ atteint en un unique réel que l'on déterminera en fonction de $x_1, \dots, x_n$ et n .
 - 2.b. Conclure sur l'estimateur du maximum de vraisemblance recherché.
3. On considère une variable aléatoire X suivant la loi uniforme sur $[a; 0]$, avec $a \leq 0$. On cherche l'estimateur du maximum de vraisemblance pour a . Soient $n \in \mathbb{N}$ et $(X_1, \dots, X_n)$ un n -échantillon de X .
- Soit $(x_1, \dots, x_n)$ une réalisation de $(X_1, \dots, X_n)$. On pose $L_n : a \mapsto \prod_{k=1}^n f_a(x_k)$, où f_a désigne la densité de X continue sur $[a; 0]$.
- 3.a. Démontrer que la fonction $a \mapsto \ln(L_n(a))$ possède un maximum sur $\mathbb{R}^-$ atteint en un unique réel que l'on déterminera en fonction de $x_1, \dots, x_n$ et n .
 - 3.b. Conclure sur l'estimateur du maximum de vraisemblance recherché.

EXERCICE 2 – ●●● – Intervalles de confiance pour l'espérance d'une $\mathcal{N}(m, 1)$

On considère une variable aléatoire X suivant une loi $\mathcal{N}(m, 1)$, où m est inconnu. On dispose d'un n -échantillon $(X_1, \dots, X_n)$ de X .

1. Donner la loi de $\bar{X}_n$ et celle de $\bar{X}_n^* = \sqrt{n}(\bar{X}_n - m)$.
2. Soit $\alpha \in]0; 1[$. Justifier l'existence d'un unique réel t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$, où Φ désigne la fonction de répartition d'une variable aléatoire suivant la loi $\mathcal{N}(0; 1)$. Montrer que $t_\alpha > 0$.
3. En déduire que $\left[\bar{X}_n - t_\alpha \frac{1}{\sqrt{n}}, \bar{X}_n + t_\alpha \frac{1}{\sqrt{n}} \right]$ est un intervalle de confiance de m au niveau de confiance $1 - \alpha$.
4. A l'aide de l'inégalité de Bienaymé-Tchebychev, démontrer que $\left[\bar{X}_n - \frac{1}{\sqrt{n\alpha}}, \bar{X}_n + \frac{1}{\sqrt{n\alpha}} \right]$ est un intervalle de confiance de m au niveau de confiance $1 - \alpha$.
5. **Comparaison des deux intervalles.** On considère $n = 100$ et $\alpha = 0,05$. A l'aide d'une table de la loi $\mathcal{N}(0; 1)$, déterminer une valeur approchée de t_α . Comparer alors l'amplitude des deux intervalles de confiance obtenus dans les questions précédentes.
6. Écrire une fonction **Python** prenant en arguments un réel m et un entier naturel non nul n , qui simule 1000 réalisations d'un échantillon de taille n d'une loi $\mathcal{N}(m, 1)$, et qui compte le nombre de réalisations fournissant un intervalle de confiance contenant la valeur m choisie.

EXERCICE 3 – ●●● – Prise de décision

Une entreprise souhaite acquérir une machine qui fabrique un certain type d'objets et qui, en fonctionnement normal, produit une proportion p ($0 < p < 1$) d'objets défectueux. Le directeur veut connaître la valeur de p . Pour cela, il teste la machine et prélève un échantillon de n objets qu'il analyse, avec $n \geq 1$.

Pour tout $i \in \llbracket 1, n \rrbracket$, soit X_i la variable aléatoire de Bernoulli définie par

$$X_i = \begin{cases} 1 & \text{si le } i\text{-ème objet prélevé est défectueux} \\ 0 & \text{sinon} \end{cases}$$

On suppose que dans les conditions de prélèvement, les variables aléatoires $X_1, \dots, X_n$ sont indépendantes.

On pose $S_n = \sum_{k=1}^n X_k$.

1. 1.a. Montrer que $F_n = \frac{S_n}{n}$ est un estimateur de p d'espérance p .
- 1.b. Démontrer : $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(|F_n - p| \geq \varepsilon) = 0$.

2. Soit α un réel de $]0, 1[$. On souhaite déterminer dans cette question un intervalle de confiance du paramètre p inconnu, au niveau de confiance $1 - \alpha$, à partir de l'échantillon $(X_1, \dots, X_n)$.

2.a. Quelle est la limite en loi de la suite $\left(\sqrt{n} \frac{F_n - p}{\sqrt{p(1-p)}} \right)_{n \in \mathbb{N}^*}$?

- 2.b. Justifier l'existence d'un unique réel t_α tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$, où Φ désigne la fonction de répartition d'une variable aléatoire suivant la loi normale centrée réduite. Vérifier que $t_\alpha > 0$.

Montrer qu'un intervalle de confiance de p au niveau $1 - \alpha$ est donné par $\left[F_n - \frac{t_\alpha}{2\sqrt{n}}, F_n + \frac{t_\alpha}{2\sqrt{n}} \right]$.

- 2.c. On suppose dans cette question qu'en fonctionnement normal la machine produit une proportion $p = 0,05$ d'objets défectueux. Le directeur analyse 10000 objets et compte 700 objets défectueux sur cet échantillon. Décide-t-il d'acheter la machine, avec un risque égal à 0,05 ?
On donne $\Phi(2) \simeq 0,975$.

EXERCICE 4 - ●●●

Soit $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi $\mathcal{U}([0, \theta])$, où θ est inconnu. On pose, pour tout $n \in \mathbb{N}^*$, $M_n = \max(X_1, \dots, X_n)$.

1. Soit $(x_1, \dots, x_n)$ un n -uplet de réalisations de $(X_1, \dots, X_n)$. On pose $L : \theta \mapsto \prod_{k=1}^n f(x_k)$, où f désigne une densité de X continue sur $[0; \theta]$.

Montrer que la fonction L possède un maximum atteint en un unique réel à déterminer.

2. Démontrer que M_n est une variable aléatoire à densité et en déterminer une densité.

3. Étudier la convergence en loi de M_n .

4. Démontrer :

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(|M_n - \theta| \geq \varepsilon) = 0$$

5. 5.a. Démontrer que pour $\theta > 0$ fixé, il existe deux réels x_1 et x_2 , que l'on exprimera en fonction de θ , tels que

$$F_{M_n}(x_1) = 0,025 \quad ; \quad F_{M_n}(x_2) = 0,975$$

- 5.b. En déduire un intervalle de confiance de θ au niveau de confiance 0,95.

6. 6.a. Démontrer que

$$n \left(1 - \frac{M_n}{\theta} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X$$

où X est une variable aléatoire suivant la loi exponentielle $\mathcal{E}(1)$.

- 6.b. Déterminer un intervalle de confiance asymptotique de θ au niveau de risque α de la forme $[M_n, U_n]$, où U_n est un estimateur de θ .

EXERCICE 5 - ●●● - EDHEC 2016 S

Soit X une variable aléatoire suivant la loi normale de paramètres m et σ^2 (avec $\sigma > 0$). On suppose que l'on ne connaît pas les paramètres $\theta_1 = m$ et $\theta_2 = \sigma^2$ et on souhaite les estimer par une méthode appelée méthode du maximum de vraisemblance.

Pour ce faire, on considère un n -échantillon $(X_1, \dots, X_n)$ de la loi de X , avec $n \geq 2$.

On appelle vraisemblance du couple (θ_1, θ_2) , la fonction notée L définie par :

$$L(\theta_1, \theta_2) = \prod_{i=1}^n \varphi_{\theta_1, \theta_2}(x_i), \text{ où } x_1, \dots, x_n \text{ sont des nombres réels donnés}$$

1. Donner l'expression de $L(\theta_1, \theta_2)$, puis celle de $\ln(L(\theta_1, \theta_2))$ en fonction de θ_1, θ_2 et $x_1, \dots, x_n$.

2. 2.a. Justifier que la fonction $f : (\theta_1, \theta_2) \mapsto \ln(L(\theta_1, \theta_2))$, définie sur l'ouvert $U = \mathbb{R} \times \mathbb{R}^{++}$, est de classe $\mathcal{C}^2$ sur U .

- 2.b. Montrer que f admet un seul point critique $A = (\hat{\theta}_1, \hat{\theta}_2)$ sur U tel que :

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^n x_i \text{ et } \hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \hat{\theta}_1^2$$

- 2.c. Déterminer les valeurs des dérivées partielles d'ordre 2 de f en A .

On vérifiera en particulier que : $\partial_{\hat{\theta}_2, \hat{\theta}_2}^2(f) \left(\hat{\theta}_1, \hat{\theta}_2 \right) = \frac{-n}{2\hat{\theta}_2^2}$.

- 2.d. En déduire que f admet un maximum local en $(\hat{\theta}_1, \hat{\theta}_2)$.

- 2.e. Expliquer pourquoi la fonction L admet aussi un maximum local en $(\hat{\theta}_1, \hat{\theta}_2)$.