

Programme de colle n° 10

à partir du lundi 16 mars

— **Chapitres pouvant faire l'objet d'exercices**

— Inégalités, convergence, estimations

— **Question de cours/exercices à préparer**

*Les exercices doivent être préparés **avant** la colle.*

(Q1) On considère une variable aléatoire X suivant une loi de Bernoulli de paramètre $p \in]0; 1[$ inconnu que l'on cherche à estimer. Pour tout entier $n \geq 1$, on dispose d'un n -échantillon

$(X_1, \dots, X_n)$ de la loi de X et on pose $\bar{Y}_n = \frac{2}{n^2} \sum_{k=1}^n kX_k$.

1. Calculer l'espérance de $\bar{Y}_n$ en fonction de p . En déduire le biais $b_p(\bar{Y}_n)$ de $\bar{Y}_n$ en tant qu'estimateur de p .

2. On donne $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$. Calculer la variance de $\bar{Y}_n$ en fonction de p .

3. Déterminer le risque quadratique $r_p(\bar{Y}_n)$ de $\bar{Y}_n$ en tant qu'estimateur de p .

4. Déterminer $\lim_{n \rightarrow +\infty} r_p(\bar{Y}_n)$.

(Q2) On considère une variable aléatoire $X \hookrightarrow \mathcal{E}(\lambda)$ où $\lambda > 0$ est inconnu. Pour tout entier $n \geq 1$, on dispose d'un n -échantillon $(X_1, \dots, X_n)$ de la loi de X et on pose $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$.

1. Déterminer, en fonction de λ , $\mathbf{E}[\bar{X}_n]$ et $\mathbf{V}(\bar{X}_n)$.

2. Soit $n \in \mathbb{N}^*$. À l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que pour tout $\varepsilon \in \mathbb{R}_+^*$:

$$\mathbf{P}\left(\left|\bar{X}_n - \frac{1}{\lambda}\right| \geq \varepsilon\right) \leq \frac{1}{n(\lambda\varepsilon)^2}.$$

3. Dans cette question, on admet que $\lambda \geq 1$.

(a) Démontrer grâce à l'hypothèse sur λ que $\frac{1}{n(\lambda\varepsilon)^2} \leq \frac{1}{n\varepsilon^2}$.

(b) Soit $\alpha \in]0, 1[$. On pose $\varepsilon = \frac{1}{\sqrt{n\alpha}}$. Montrer que $]\bar{X}_n - \varepsilon, \bar{X}_n + \varepsilon[$ est un intervalle de confiance pour $\frac{1}{\lambda}$ au niveau de confiance $1 - \alpha$.

(Q3) Une culture expérimentale de tardigrades abrite un nombre inconnu n de spécimens que l'on cherche à estimer. Pour ce faire, on capture 10 tardigrades que l'on marque radioactivement avant de les réintégrer dans la culture.

On attend quelques heures (on suppose que la population de tardigrades reste stable durant cet intervalle de temps) puis on refait des prélèvements dans la culture jusqu'à avoir obtenu 2 tardigrades radioactifs. On suppose que les prélèvements de tardigrades se font de manière indépendante.

Lors de ces prélèvements, on réintroduit les tardigrades non radioactifs prélevés tandis que l'on isole les tardigrades radioactifs prélevés.

Soit X_1 la variable aléatoire égale au nombre de prélèvements permettant d'obtenir le premier tardigrade radioactif, X_2 la variable aléatoire égale au nombre de prélèvements entre celui du premier tardigrade radioactif (exclu) et du second tardigrade radioactif (inclus).

1. Donner en justifiant soigneusement la loi de X_1 en fonction de n .

2. De même, donner la loi de X_2 en fonction de n .

3. On pose $Y = \frac{90}{19}(X_1 + X_2 + \frac{1}{9})$. Démontrer que Y est un estimateur sans biais de n .

4. Écrire un programme Python complet qui simule une réalisation de la variable aléatoire Y .

- (Q4) On considère une variable aléatoire X suivant une loi $\mathcal{U}([-2\theta, \theta])$ où $\theta > 0$ est un paramètre inconnu que l'on cherche à estimer. Pour tout entier $n \geq 1$, on dispose d'un n -échantillon $(X_1, \dots, X_n)$ de la loi de X et on pose

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

1. Donner l'expression d'une densité usuelle de X , de sa fonction de répartition, ainsi que $\mathbf{E}[X]$ et $V(X)$.
2. Montrer que $-2\bar{X}_n$ est un estimateur sans biais de θ .
3. Déterminer le risque quadratique $r_\theta(-2\bar{X}_n)$ de $-2\bar{X}_n$ en tant qu'estimateur de θ .
4. Une réalisation de $(X_1, \dots, X_{10})$ est $(-0.2, -1.9, 0.4, -1.3, 0.2, -0.5, 0.7, -1.7, -0.9, 0.8)$. Donner une estimation de θ . Peut-on améliorer cette estimation de θ au vu de l'échantillon réalisé et de la loi de X ?