Sorbonne Université — Année 2025/2026

L3 — Deuxième semestre

Feuille de TD 3 : Intervalles de confiance

Exercice 1.

Dans un centre avicole, des études antérieures ont montré que le poids d'un oeuf choisi au hasard peut être considéré comme la réalisation d'une variable aléatoire gaussienne X, d'espérance μ et de variance σ^2. On admet que les poids des oeufs sont indépendants les uns des autres. On prend un échantillon de $n = 36$ oeufs que l'on pèse. Les mesures obtenues (exprimées en g) sont données (par ordre croissant) dans le tableau suivant :

50.34	52.62	53.79	54.99	55.82	57.67	51.41	53.13	53.89
55.04	55.91	57.99	51.51	53.28	54.63	55.12	55.95	58.10
52.07	53.30	54.76	55.24	57.05	59.30	52.22	53.32	54.78
55.28	57.18	60.58	52.38	53.39	54.93	55.56	57.31	63.15

TAB. 1: Mesures des poids des œufs (en grammes).

1. La figure ci-dessous présente l'histogramme en fréquences des poids des oeufs sur lequel on a superposé une version lissée de l'histogramme. Quelles conclusions peut-on tirer de cet histogramme sur la distribution des oeufs ?

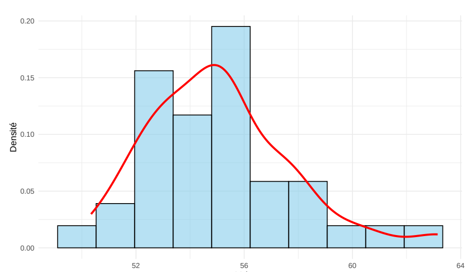

FIG. 1: Histogramme en fréquence des poids des oeufs.

2. Donner une estimation de la moyenne μ et de la variance σ^2. On les notera m et s^2.
   Pour vous aider, on fournit les informations suivantes : si $x_1, x_2, \ldots, x_{36}$ désignent les poids mesurés, alors
$$\sum_{j=1}^{36} x_j = 1982,99 \quad \text{et} \quad \sum_{j=1}^{36} x_j^2 = 109481,1$$

3. Construire un intervalle de confiance au niveau 95% pour le poids moyen des oeufs et calculer sa réalisation sachant que le quantile d'ordre 0.975 d'une loi de Student à 35 degrés de liberté vaut environ 2.03.
4. Construire un intervalle de confiance au niveau 95% pour la variance et calculer sa réalisation sachant que les quantiles d'ordre 0.025 et 0.975 valent environ 20.6 et 53.2
5. A quel niveau de confiance correspondrait un intervalle centré en m et de demi-longueur 0.76 ?

Exercice 2.

On considère une variable aléatoire X dont la loi dépend d'un paramètre inconnu $p > 0$ et telle que $\mathbb{E}\left[X^2\right] = \frac{1}{p^2}$ et $\mathbb{E}\left[X^4\right] = \frac{1}{p^6} + \frac{1}{p^4}$, et on pose $\theta = p^{-2}$.

1. Proposer un estimateur $\hat{\theta}_n$ de θ. Est-il consistant ? Fortement consistant ?
2. Est-il sans biais ?
3. Donner son erreur quadratique moyenne.
4. Donner sa normalité asymptotique.
5. En déduire un intervalle de confiance asymptotique de niveau $1 - \alpha$ pour θ.
6. En déduire un intervalle de confiance asymptotique de niveau $1 - \alpha$ pour p.
7. Proposer un estimateur $\hat{p}_n$ de p et montrer sa forte consistance.
8. Donner sa normalité asymptotique.
9. En déduire un nouvel intervalle de confiance asymptotique de niveau $1 - \alpha$ pour p.

Exercice 3.

Soit $\theta > 0$ et Y une variable aléatoire suivant une loi exponentielle de paramètre θ^{-1}. Soit $X = Y + \theta$, sa densité f_θ est définie pour tout x par

$$f_\theta(x) = C_\theta \exp\left(-\frac{x-\theta}{\theta}\right) \mathbf{1}_{[\theta,+\infty[}(x).$$

Dans ce qui suit, on considère $x_1, \ldots, x_n$ qui sont des réalisations des variables aléatoires indépendantes $X_1, \ldots, X_n$ de même loi que X.

1. Que vaut C_θ ?
2. Calculer $\mathbb{E}[X]$ et $\mathbb{V}[X]$.
3. Par la méthode des moments, en déduire un estimateur de θ.
4. Est-il consistant ? Fortement consistant ? Asymptotiquement normal ?

5. Soit $\alpha \in (0,1)$, construire un intervalle de confiance asymptotique de niveau $1-\alpha$ du paramètre θ.
6. Donner l'estimateur du maximum de vraisemblance.
7. On considère maintenant $X_{(1)} = \min_{i=1,\ldots,n}(X_i)$. Donner la fonction de répartition de X. En déduire celle de $X_{(1)}$.
8. Donner la loi de $n\left(X_{(1)} - \theta\right)$.
9. En déduire la convergence en probabilité de $X_{(1)}$.
10. Montrer que l'estimateur $X_{(1)}$ est biaisé mais asymptotiquement sans biais.
11. En déduire la convergence en moyenne quadratique de $X_{(1)}$.
12. Donner un nouvel intervalle de confiance pour le paramètre θ.
13. Commenter

Exercice 4.

Soit X une variable aléatoire de densité $f_\theta(x) = \frac{2}{\theta^2} x \mathbf{1}_{[0,\theta]}(x)$, avec $\theta > 0$. On considère dans ce qui suit des réalisations $x_1, \ldots, x_n$ des variables aléatoires indépendantes $X_1, \ldots, X_n$ de même loi que X.

1. Calculer $\mathbb{E}[X]$ et $\mathbb{V}[X]$.
2. Par la méthode des moments, en déduire un estimateur $\hat{\theta}_n$ de θ.
3. Est-il consistant ? Fortement consistant ? Asymptotiquement normal ?
4. Soit $\alpha \in (0,1)$, construire un intervalle de confiance asymptotique de niveau $1-\alpha$ du paramètre θ.
5. Donner l'estimateur du maximum de vraisemblance.
6. On considère maintenant $X_{(n)} = \max_{i=1,\ldots,n} X_i$. Donner la fonction de répartition de X.
7. En déduire la fonction de répartition de $X_{(n)}$.
8. En déduire la densité de $X_{(n)}$.
9. L'estimateur $X_{(n)}$ est-il sans biais ?
10. Calculer le risque quadratique $\mathbb{E}\left[\left(X_{(n)} - \theta\right)^2\right]$.
11. Donner la convergence en loi de $n\left(\theta - X_{(n)}\right)$.
12. Donner un nouvel intervalle de confiance au niveau $1-\alpha$ de θ.
13. Quel intervalle de confiance choisiriez vous ?

Exercice 5.

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes suivant une loi de Poisson de paramètre $\theta > 0$.

1. Donner $\mathbb{E}[X_1]$ et $\mathbb{V}[X_1]$. En déduire un estimateur $\hat{\theta}_n$ de θ.
2. Donner la normalité asymptotique de l'estimateur $\hat{\theta}_n$.
3. Par la méthode du plug-in, donner un intervalle de confiance asymptotique à 95% pour θ.
4. Trouver une fonction $g : \mathbb{R}_+ \longrightarrow \mathbb{R}$ telle que
$$\sqrt{n}\left(g\left(\hat{\theta}_n\right)\right) - g(\theta)) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$
5. En déduire un nouvel intervalle de confiance pour θ.
6. Vérifier que les deux résultats sont équivalents, i.e si on note $[a_n, b_n]$ et $[a'_n, b'_n]$ les deux intervalles obtenus, on a
$$\frac{a_n}{a'_n} \xrightarrow[n\to+\infty]{\mathbb{P}} 1 \qquad \text{et} \qquad \frac{b_n}{b'_n} \xrightarrow[n\to+\infty]{\mathbb{P}} 1.$$

Exercice 6 – Loi de Laplace.

Soit Y une variable aléatoire suivant une loi de Laplace, i.e de densité f_Y définie pour tout $x \in \mathbb{R}$ par
$$f_Y(x) = \frac{1}{2}\exp\left(-|x|\right).$$
Soit $\theta > 0$ et X une variable aléatoire suivant une loi de Laplace translatée, i.e de densité f_X définie pour tout $x \in \mathbb{R}$ par
$$f_X(x) = \frac{1}{2}\exp\left(-|\theta - x|\right).$$
Dans ce qui suit on considère des variables aléatoires indépendantes $X_1, \dots, X_n$ de même loi que X.

1. Calculer $\mathbb{E}[X]$ et $\mathbb{V}[X]$.
2. Par la méthode des moments , donner un estimateur de θ. Est-il consistant ? Asymptotiquement normal ? Que pouvez vous en conclure ?
3. Construire un intervalle de confiance asymptotique de niveau 0.90 pour θ.
4. Que vaut l'estimateur du maximum de vraisemblance ?
5. On note $\hat{m}_n$ l'estimateur du maximum de vraisemblance, on suppose qu'il existe et est consistant. Que pouvez-vous en conclure ?

Exercice 7 – Débiaiser ou ne pas débiaiser ?.

On considère une variable aléatoire $X \sim \mathcal{U}([0,\theta])$ avec $\theta > 0$ et on considère l'estimateur $\hat{\theta}_n = X_{(n)}$.

1. Est-il biaisé ? Calculer l'erreur quadratique moyenne.
2. Proposer un estimateur non biaisé et calculer son erreur quadratique moyenne.
3. On considère l'estimateur $\hat{\theta}_{\alpha,n} = \alpha X_{(n)}$. Calculer l'erreur quadratique moyenne.
4. Choisir α afin de minimiser l'erreur quadratique moyenne.
5. Conclure.

Exercice 8 – J'existe ?.

On considère une variable aléatoire $X \sim \mathcal{B}\left(p, \lambda^{-1}\right)$ avec $\lambda > 1$, et on souhaite savoir s'il existe un estimateur non biaisé de θ. Pour cela, on considère qu'il existe et on le note $\hat{\theta}(X)$.

1. Calculer $\mathbb{E}\left[\hat{\theta}(X)\right]$.
2. Sachant qu'un polynôme non nul de degré p admet au plus p racines, qu'observez-vous si $\hat{\theta}(X)$ est non biaisé ?

Exercice 9 – Estimation de deux paramètres.

Soit $Y \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$ inconnu. Soit $X = Y + \theta$ avec $\theta > 0$ inconnu. On admettra que X a pour densité $f_{\theta,\lambda}$ définie pour tout x par

$$f_{\theta,\lambda}(x) = \lambda \exp\left(-\lambda(x-\theta)\right) \mathbf{1}_{[\theta,+\infty[}(x).$$

Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d de même loi que X.

1. Calculer $\mathbb{E}[X]$ et $\mathbb{V}[X]$.
2. Déterminer l'estimateur du maximum de vraisemblance de θ.
3. Calculer la fonction de répartition de $X_{(1)}$.
4. En déduire l'erreur quadratique moyenne de $X_{(1)}$.
5. Déduire de la question 5
$$\sqrt{n}\left(X_{(1)} - \theta\right) \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$
Que pouvez-vous en déduire ?
6. En déduire un estimateur de λ.

7. Montrer que l'estimateur de λ est consistant. Pour s'aider, on admettra que si (A_n) et (B_n) sont deux suites de variables aléatoires convergeant en probabilités vers a et b, et $g : I \times J \longrightarrow \mathbb{R}$ est une fonction continue en (a, b), alors

$$g(A_n, B_n) \xrightarrow[n \to +\infty]{\mathbb{P}} g(a, b).$$

avec I, J des intervalles ouvert de $\mathbb{R}$.

8. Montrer que

$$\sqrt{n}\left(\overline{X}_n - \hat{\theta}_n - \frac{1}{\lambda}\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\lambda^2}\right).$$

9. En déduire la normalité asymptotique de l'estimateur de λ.
10. Soit $\alpha \in (0, 1)$, en déduire un intervalle de confiance asymptotique de λ.
11. Montrer que

$$n\left(X_{(1)} - \theta\right) \sim \mathcal{E}(\lambda).$$

12. Pour tout $\alpha \in (0, 1)$, donner le quantile $q_{\lambda,1-\alpha}$ d'ordre $1 - \alpha$ de la loi exponentielle de paramètre λ.
13. Donner la convergence de

$$n\left(X_{(1)} - \theta\right) - \left(q_{\hat{\lambda}_n,1-\alpha} - q_{\lambda,1-\alpha}\right).$$

avec $q_{\hat{\lambda}_n} = \frac{-\ln(\alpha)}{\hat{\lambda}_n}$.

14. Déterminer un intervalle de confiance asymptotique de niveau $1 - \alpha$ pour θ.

Exercice 10.

On considère une urne dans laquelle il y a m_1 boules rouges et m_2 boules noires. Le nombre de boules de chaque couleur est inconnu, et le nombre total de boules est bien trop conséquent pour que l'on s'amuse à les compter. L'objectif est donc de proposer différentes stratégies pour estimer la proportion $p = \frac{m_1}{m_1+m_2}$ de boules rouges.

1. On propose d'effectuer N tirages avec remise et on note $X_i = 1$ si le i-ème tirage est une boule rouge et $X_i = 0$ sinon.
   (a) Proposer un estimateur de p et donner son erreur quadratique moyenne.
   (b) Donner sa normalité asymptotique et en déduire un intervalle de confiance asymptotique de niveau $1 - \alpha$ de p.
2. On propose d'effectuer N tirages sans remise et on note Y le nombre de boules rouges tirées ($N << m_1 + m_2$).
   (a) Quelle est la loi de Y ? On admettra que $\mathbb{E}[Y] = Np$ et $\mathbb{V}[Y] = Np(1-p)\frac{m_1+m_2-N}{m_1+m_2-1}$.

(b) Proposer un nouvel estimateur de p et donner son erreur quadratique moyenne. Quel estimateur choisiriez-vous ?

(c) Donner un intervalle de confiance de niveau au moins $1 - \alpha$ de p.

3. On propose d'effectuer la même expérience n fois, mais en n'effectuant que $K = \frac{N}{n}$ (on suppose K entier) tirages sans remise et on note Y_i le nombre de boules rouges tirées à la i-ème expérience.

   (a) Proposer un nouvel estimateur de p et donner son erreur quadratique moyenne. Quel estimateur choisiriez-vous ?

   (b) Donner sa normalité asymptotique et en déduire un nouvelle intervalle de confiance.

4. On a $m_1 = 2500$, $m_2 = 7500$, $N = 1000$, $K = 10$ et on obtient des intervalles de confiance (dans l'ordre) de taille 0.054, 0.14, 0.053. Comment interpréter ces résultats ?