

Feuille de TD 2 : Estimation et intervalles de confiance

Exercice 1 (Loi géométrique)

Soit $X \sim \mathcal{G}(p)$ de paramètre p , i.e.,

$$\mathbb{P}(X = k) = (1 - p)^{k-1}p, \quad k \geq 1.$$

1. Rappeler $\mathbb{E}[X]$ et $\text{Var}(X)$.
2. Par méthode des moments, donner un estimateur $\hat{p}_n$ de p .
3. Est-il consistant ? Fortement consistant ? Asymptotiquement normal ?
4. Donner l'estimateur du maximum de vraisemblance $\hat{p}_n^{MV}$. Que pouvez-vous conclure ?

Exercice 2 (Loi de Poisson)

Soit $X \sim \mathcal{P}(\theta)$.

1. Par méthode des moments, proposer un estimateur de θ .
2. Discuter consistance, forte consistance et normalité asymptotique.
3. Donner l'estimateur du maximum de vraisemblance.
4. Discuter consistance, forte consistance et normalité asymptotique.

Exercice 3 (Loi exponentielle translatée)

Soit $Y \sim \text{Exp}(1)$, c'est-à-dire de densité $f(x) = e^{-x}\mathbf{1}_{\mathbb{R}^+}(x)$. On pose $X = Y + \theta$, pour un certain θ , X admet donc pour densité $f_\theta(x) = e^{-(x-\theta)}\mathbf{1}_{[\theta, +\infty[}(x)$.

1. Calculer $\mathbb{E}[X]$, $\text{Var}(X)$.
2. Donner un estimateur $\hat{\theta}_n$ par la méthode des moments de θ , et MV de θ . Étudier biais, (forte) consistance, normalité asymptotique.
3. Quelle est l'erreur quadratique moyenne associée à $\hat{\theta}_n$?
4. Donner l'estimateur $\hat{\theta}_n^{MV}$ par maximum de vraisemblance de θ .
5. Donner la loi de $n(\hat{\theta}_n^{MV} - \theta)$.
6. Calculer la médiane de X_1 . En déduire un nouvel estimateur de θ . Est-il consistant ? Asymptotiquement normal ?
7. Quel estimateur choisiriez-vous ?

Exercice 4 (Loi de Laplace)

1. Soient $X_1, \dots, X_n$ des v.a. i.i.d. de densité $f_\theta(x) = \frac{1}{2\theta}e^{-|x|/\theta}$ pour tout $x \in \mathbb{R}$ avec $\theta > 0$. Calculer l'estimateur MV de θ . Discuter unicité, consistance et normalité asymptotique.
2. Soient $X_1, \dots, X_n$ des v.a. i.i.d. de densité $f_\mu(x) = \frac{1}{2}e^{-|x-\mu|}$ pour tout $x \in \mathbb{R}$ avec $\mu \in \mathbb{R}$. Donner l'estimateur MV de μ . Discuter unicité, consistance et normalité asymptotique.

Exercice 5 (Propriétés des moyenne et médiane empiriques)

Soit $X_1, \dots, X_n$ un échantillon de fonction de répartition F avec $\mathbb{E}(X_1^2) < \infty$.

1. (a) Montrer que $\mathbb{E}(X_1) = \operatorname{argmin}_{t \in \mathbb{R}} \mathbb{E}[(X_1 - t)^2]$.
(b) Déterminer $\operatorname{argmin}_{t \in \mathbb{R}} \sum_{i=1}^n (X_i - t)^2/n$.
2. On note $x_{1/2} = F^{-1}(1/2)$ la médiane de la loi de X_1 .
(a) Montrer que si F est continue sur $\mathbb{R}$ et strictement croissante sur un voisinage de $x_{1/2}$, alors $x_{1/2} = \operatorname{argmin}_{t \in \mathbb{R}} \mathbb{E}[|X_1 - t|]$.
(b) Déterminer, selon la parité de n , $\operatorname{argmin}_{t \in \mathbb{R}} \sum_{i=1}^n |X_i - t|/n$.

Exercice 6 (Intervalle de confiance et loi normale)

On observe $X_1, \dots, X_n$ indépendantes et de même loi. On suppose qu'il existe $\theta > 0$ tel que cette loi admet la densité

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\theta^2} \exp\left(-\frac{x^2}{2\theta^2}\right).$$

Indication : $F_{\chi_{10}^2}^{-1}(0.025) \simeq 3.25$, $F_{\chi_{10}^2}^{-1}(0.975) \simeq 20.48$, $F_{\chi_{10}^2}^{-1}(0.95) \simeq 18.31$, $F_{\chi_{10}^2}(40/3) \simeq 0.79$ et $\Phi^{-1}(0.975) \simeq 2$.

1. On veut estimer $\tau = \theta^2$. Proposer un estimateur $\hat{\tau}$ de τ et étudier sa loi.
2. Construire un intervalle de confiance au niveau $1 - \alpha$ de la forme $[S_1, S_2]$ tel que

$$\mathbb{P}(\tau < S_1) = \mathbb{P}(\tau > S_2) = \alpha/2.$$

3. Donner la loi asymptotique de $\hat{\tau}$ et en déduire un intervalle de confiance asymptotique.
4. Lorsque $n = 10$, $\hat{\tau} = 2$ et $\alpha = 0.05$, comparer l'intervalle de confiance obtenu à la question 2 (non asymptotique) avec celui obtenu à la question 3 (asymptotique).
5. Construire un test de niveau α de

$$H_0 : \tau \leq 3 \text{ contre } H_1 : \tau > 3.$$

Pour un seuil de 5%, lorsque $n = 10$ et qu'on observe $\hat{\tau} = 4$, rejette-t-on l'hypothèse nulle ? Calculer la p -valeur du test.

Exercice Bonus (Des tanks nazis) 🚀🚀

Au cours de la seconde guerre mondiale, l'armée alliée notait les numéros de série $X_1, \dots, X_n$ de tous les tanks nazis capturés ou détruits, afin d'obtenir un estimateur du nombre total N de tanks produits.

On propose un modèle de tirage uniforme sans remise dans $\{1, \dots, N\}$. Dans ce cadre, les observations ne sont pas indépendantes, mais elles demeurent échangeables.

1. Étudier la loi de $X_{(n)} = \max X_i$ et calculer l'espérance de $\bar{X}_n$. En déduire deux estimateurs non biaisés de K .
2. Proposer deux intervalles de confiance de niveau α . On pourra utiliser le fait que l'inégalité de Hoeffding s'applique également aux tirages sans remise.

Note : Selon Ruggles et Broodie (1947, JASA), la méthode statistique a fourni comme estimation une production moyenne de 246 tanks/mois entre juin 1940 et septembre 1942. Des méthodes d'espionnage traditionnelles donnaient une estimation de 1400 tanks/mois. Les chiffres officiels du ministère nazi des Armements ont montré après la guerre que la production moyenne était de 245 tanks/mois.