TD 2 : estimateurs

Les questions marquées d'un astérisque (*) *sont facultatives.*

Dans tous les exercices, on commencera par soigneusement écrire le modèle statistique considéré.

Exercice 1. (Un grand classique) Soient $X_1, \dots, X_n$ une suite de v.a. (variables aléatoires) i.i.d. (indépendantes et identiquement distribuées) suivant une loi normale de moyenne μ et variance σ^2 inconnues, qu'on cherche à estimer.

1. Calculer un estimateur du maximum de vraisemblance de (μ, σ^2). Calculer son biais et sa variance. Est-il consistant ?
2. Calculer un estimateur de (μ, σ^2) en utilisant la méthode des moments. Que constatez-vous ?

Exercice 2. Soient $X_1, \dots, X_n$ une suite de v.a. i.i.d. suivant une loi uniforme sur $[0, \theta]$. On cherche à estimer θ.

1. Calculer un estimateur du maximum de vraisemblance de θ. Calculer sa fonction de répartition. En déduire sa densité, son biais et sa variance. Est-il consistant ?
2. Calculer un estimateur de θ en utilisant la méthode des moments. Calculer son biais et sa variance. Est-il consistant ?

Exercice 3. Soient $X_1, \dots, X_n$ une suite de v.a. i.i.d. suivant une loi exponentielle de paramètre inconnu.

1. Comparer les estimateurs de θ obtenus par la méthode du maximum de vraisemblance et la méthode des moments.
2. On considère l'estimateur du maximum de vraisemblance. Est-il consistant ? Que peut-on dire de son biais ?

On rappelle l'inégalité de Jensen : soit A un intervalle de $\mathbb{R}$, $\varphi : A \longrightarrow \mathbb{R}$ une fonction convexe et X une variable aléatoire à valeurs dans A. Alors

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

De plus, si φ est strictement convexe, alors il y a égalité si et seulement si X est constante.

Exercice 4. (Unicité de l'estimateur du maximum de vraisemblance)

1. Soient $X_1, \dots, X_n$ une suite de v.a. i.i.d. uniformes sur $[\theta, \theta + 1]$. Calculer un estimateur du maximum de vraisemblance. Est-il unique ?
2. Soient $X_1, \dots, X_n$ une suite de v.a. i.i.d. de loi de Laplace de paramètres $(\mu, 1)$ (autrement dit, de densité $x \mapsto \frac{\exp(-|x-\mu|)}{2}$). Montrer que la log-vraisemblance est dérivable sur $\mathbb{R} \setminus \{X_1, \dots, X_n\}$ de dérivée
$$\mu \longmapsto \#\{i \in \{1, \dots, n\} : X_i > \mu\} - \#\{i \in \{1, \dots, n\} : X_i < \mu\}.$$
En déduire que tout estimateur du maximum de vraisemblance de μ est une médiane de l'échantillon $(X_1, \dots, X_n)$. L'estimateur du maximum de vraisemblance est-il unique ?

Problème : régression non paramétrique (*). On observe des v.a. $Y_1, \ldots, Y_n$ qu'on suppose de la forme

$$Y_j = f\left(\frac{j}{n}\right) + \varepsilon_j$$

où $\varepsilon_1, \ldots, \varepsilon_n$ sont des v.a. i.i.d. gaussiennes de variance connue σ^2 et où f est une fonction inconnue de $[0,1]$ dans $\mathbb{R}$ qu'on cherche à estimer. On fait l'hypothèse que f est (α, L)-höldérienne pour $\alpha \in (0,1]$ et $L > 0$, autrement dit : pour tout x, y, $|f(x) - f(y)| \leq L|x-y|^\alpha$.

1. Écrire le modèle statistique correspondant.

Comme ce modèle se prête mal à la méthode du maximum de vraisemblance, on procède plutôt comme suit. Pour tout $i \in \{1, \ldots, D\}$, soit $I_i =]\frac{i-1}{D}, \frac{i}{D}]$. On cherche un estimateur parmi les fonctions constantes sur chaque I_i (attention, la vraie fonction f reste (α, L)-höldérienne). Pour simplifier, on suppose $D|n$ et on note $K = n/D$ le nombre de points de la forme j/n qui tombent dans I_i (pour un i fixé).

2. Écrire ce nouveau modèle statistique sous la forme d'un modèle paramétrique.
3. Montrer que l'estimateur du maximum de vraisemblance de f s'écrit
$$\hat{f}(x) = \sum_{i=1}^{D} \left(\frac{1}{K} \sum_{j=K(i-1)+1}^{Ki} Y_j \right) \mathbf{1}_{I_i}(x).$$
Autrement dit, $\hat{f}$ est constant sur chaque I_i égal à la moyenne des Y observés sur ce I_i.
4. Calculer l'espérance de $\hat{f}$. Cet estimateur est-il sans biais ?
5. Montrer que l'erreur $\mathbf{L}^2$ de cet estimateur se décompose suivant le compromis biais-variance
$$\mathbb{E}[\|\hat{f} - f\|^2] = \mathbb{E}[\|\hat{f} - \mathbb{E}[\hat{f}]\|^2] + \|\mathbb{E}[\hat{f}] - f\|^2.$$
En déduire que si D reste constant quand $n \to \infty$, cet estimateur n'est pas consistant.
6. Montrer qu'il existe une constante $c \leq 1$ telle que le biais vérifie $\|\mathbb{E}[\hat{f}] - f\|^2 \leq cL^2 D^{-2\alpha}$.
7. Montrer que la variance vérifie $\mathbb{E}[\|\hat{f} - \mathbb{E}[\hat{f}]\|^2] \leq \sigma^2 D/n$.
8. Quel choix de D proposez-vous ? En admettant que les réponses aux questions 5 à 7 restent vraies même si D ne divise pas n, quelle est la vitesse de convergence du risque de l'estimateur ?