Apprentissage statistique - TD3

Lucie Le Briquer

15 février 2018

Exercice 5.1

1. Proposition 5.2 :

   (a)

$$\begin{aligned}(\text{supp}(\mu))^C &= \{x \in \mathbb{R}^d \mid \exists \text{ voisinage } U_x \ \mu(U_x) = 0\} \\ &= \bigcup_{O \text{ ouvert tq } \mu(O)=0} O\end{aligned}$$

   (b) $A \in \mathcal{B}(\mathbb{R}^d)$, $A \subset (\text{supp}(\mu))^C$ alors $\mu(A) = 0$

   Il suffit de montrer que $\mu(\text{supp}(\mu)^C) = 0$. μ mesure finie sur $\mathbb{R}^d$ donc de Radon. Alors $\forall A \in \mathcal{B}(\mathbb{R}^d)$:

$$\mu(A) = \sup_{K \subset A \text{ cpct}} \mu(K)$$

   Donc :

$$\mu((\text{supp}(\mu))^C) = \sup_{K \subset (\text{supp}(\mu))^C \text{ cpct}} \mu(K)$$

   On montre que $\forall K \subset (\text{supp}(\mu))^C$ compact $\mu(K) = 0$. Soit un tel K. $\forall x \in K$ $\exists U_x$ tel que $\mu(U_x) = 0$.

$$K = \bigcup_{x \in K} U_x \underset{\text{compacité}}{=} \bigcup_{i=1}^{n} U_{x_i}$$

   Donc $\mu(K) \leqslant \sum_{i=1}^{n} \mu(U_{x_i}) = 0$.

2. On veut montrer que :

$$\mathbb{P}\left(\lim_{n \to +\infty} \|X_{(k_n)}(x) - x\| = 0\right) = 1$$

   Cela revient à :

$$\mathbb{P}\left(\bigcap_{\varepsilon \in \mathbb{Q}} \bigcup_{N \in \mathbb{N}} \bigcap_{n \geqslant N} \{\|X_{(k_n)}(x) - x\| \geqslant \varepsilon\}\right) = 1$$

   Il suffit de montrer que $\forall \varepsilon \in \mathbb{Q}$:

$$\mathbb{P}\left(\bigcup_{N \in \mathbb{N}} \bigcap_{n \geqslant N} \{\|X_{(k_n)}(x) - x\| \geqslant \varepsilon\}\right) = 0$$

Soit $n \in \mathbb{N}$,

$$\left\{\|X_{(k_n)}(x) - x\| \geqslant \varepsilon\right\} \subset \left\{\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i) \leqslant \frac{k_n}{n}\right\}$$

Alors, comme $\frac{k_n}{n} \to 0$,

$$\bigcup_{N\in\mathbb{N}} \bigcap_{n\geqslant N} \left\{\frac{1}{N}\sum_{i=1}^{n} \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i) \leqslant \frac{k_n}{n}\right\} \subset \left\{\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i) = 0\right\} = A$$

Or par hypothèse $x \in \mathrm{supp}(\mu)$, alors par la LFGN :

$$\mathbb{P}\Bigg(\underbrace{\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i) = \mathbb{P}(\mathcal{B}(x,e))}_{B}\Bigg) = 1$$

$A \cap B = \emptyset$ et $\mathbb{P}(B) = 1$. Donc :

$$\mathbb{P}(A) \leqslant \underbrace{\mathbb{P}(A\cap B)}_{=0} + \underbrace{\mathbb{P}(A \cap B^C)}_{=0} = 0$$

Remarque. La convergence en probabilité est beaucoup plus simple. Soit $\varepsilon > 0$.

$$\begin{aligned}
\mathbb{P}(\|X_{(k_n)}(x) - x\| \geqslant \varepsilon) &\leqslant \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i) \leqslant \frac{k_n}{n}\right) \\
&\leqslant \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} [\mathbb{P}_X(\mathcal{B}(x,\varepsilon) - \mathbb{1}_{\mathcal{B}(x,e)}(X_i)] \geqslant \mathbb{P}_X(\mathcal{B}(x,\varepsilon)) - \frac{k_n}{n}\right) \\
&\underset{\text{BT}}{\leqslant} \frac{1}{\left(\mathbb{P}_X(\mathcal{B}(x,\varepsilon)) - \frac{k_n}{n}\right)^2} \mathrm{Var}\left[\frac{1}{n}\sum \mathbb{1}_{\mathcal{B}(x,\varepsilon)}(X_i)\right] \\
&\leqslant \frac{1}{n^2} \frac{n\mathbb{P}(\mathcal{B}(x,\varepsilon))\left(1 - \mathbb{P}_X(\mathcal{B}(x,\varepsilon))\right)}{\left(\mathbb{P}_X(\mathcal{B}(x,\varepsilon)) - \frac{k_n}{n}\right)^2} \xrightarrow[n\to+\infty]{} 0
\end{aligned}$$

3. On suppose $k_n = k$ pour tout n. Soit $\varepsilon > 0$.

$$\mathbb{P}(\|X_{(k_n)(X)-X}\| \geqslant \varepsilon) = \mathbb{E}\Big[\mathbb{1}_{X\in\mathrm{supp}(\mathbb{P}_X)} \underbrace{\mathbb{E}\left[\mathbb{1}_{\|X_{(k_n)}(X)-X\|\geqslant\varepsilon}|X\right]}_{\psi_n(X)}\Big]$$

D'après (1), on a pour tout $x \in \mathrm{supp}(\mathbb{P}_X)$, $\lim_{n\to+\infty} \psi_n(x) = 0$. Par le TCD on a donc que :

$$\mathbb{P}(\|X_{(k_n)}(X) - X\| \geqslant \varepsilon) \xrightarrow[n\to+\infty]{} 0$$

$\|X_{(k_n)}(X) - X)\| \xrightarrow[n\to+\infty]{} 0$ en $\mathbb{P}$ $(*)$. Si $k_n =$ alors $\{\|X_{(k)} - X\|\}_{n\in\mathbb{N}}$ est décroissante et minorée donc converge $\mathbb{P}-$p.s. vers Z. Par $(*)$, $Z = 0$.

Maintenant (k_n) peut varier.

$$Z_n(X) = \sup_{m\geqslant n} \|X_{(k_n)}(X) - X)\|$$

On montre de même que Z_n converge en probabilité vers 0 et est décroissante minore donc converge $\mathbb{P}-$p.s. vers 0.

Exercice 5.2

1. $$\hat{f}_n(x) = \mathbb{1}_{\{\hat{\eta}(x) > \frac{1}{2}\}}$$

   avec :

   $$\hat{\eta}_n(x) = \frac{1}{k_n}\sum_{j=1}^{k_n} y_{(j)}(x) \quad > \frac{1}{2} \Leftrightarrow \underbrace{\sum_{j=1}^{k_n} y_{(j)}(x) - \frac{k_n}{2}}_{\psi_n} > 0$$

   La règle NN est donc bien locale.

2. On suppose pour la suite que $Y_i = \mathbb{1}_{\{U_i \leqslant \eta(X_i)\}}$.

   Remarque. Si (X_i, Y_i) i.i.d., on définit :

   $$\tilde{Y}_i = \mathbb{1}_{\{U_i \leqslant \eta(X_i)\}}$$

   $X_i \sim \mathbb{P}_X$, $\tilde{Y}_i | X_i \sim \mathcal{B}(\eta(X_i))$. Donc les (X_i, Y_i) et $(X_i, \tilde{Y}_i)$ ont même loi.

   $Y_i'(X)|X \sim \mathcal{B}(\eta(X))$. Comme les U_i sont i.i.d. et indépendantes de X, $Y_i'(X)|X$ sont aussi i.i.d.

   $$\begin{aligned}
   & \mathbb{P}\Big(\psi_n(x, Y_{(1)}(x), \dots, Y_{(k_n)}(x)) \neq \psi_n(x, Y'_{(1)}(x), \dots, Y'_{(k_n)}(x))\Big) \\
   \leqslant \quad & \sum_{i=1}^{k_n} \mathbb{P}\big(Y_{(i)}(x) \neq Y'_{(i)}(x)\big) \\
   \leqslant \quad & \sum_{i=1}^{k_n} \mathbb{P}\big(\{\eta(X_i(x)) \leqslant U_i \leqslant \eta(x)\} \cup \{\eta(x) \leqslant U_i \leqslant \eta(X_{(i)}(x))\}\big) \\
   \leqslant \quad & \sum_{i=1}^{k_n} \mathbb{E}[\mathbb{P}(\dots|(X_i)] \\
   & = \sum_{i=1}^{k_n} \mathbb{E}[|\eta(X_{(i)}(x)) - \eta(x)|]
   \end{aligned}$$

   Alors :

   $$\begin{aligned}
   \mathbb{P}(f_n(X) \neq \hat{f}_n(X)) &= \mathbb{E}[\mathbb{P}(\hat{f}_n(X) \neq \hat{f}'_n(X)|X)] \\
   &\leqslant \mathbb{E}\big[\mathbb{P}\big(\psi_n(X, Y_{(1)}(X), \dots) \neq \psi_n(X, Y'_{(1)}(X), \dots)|X\big)\big] \\
   &\leqslant \sum_{i=1}^{k_n} \mathbb{E}[|\eta(X_{(i)}(x)) - \eta(x)|]
   \end{aligned}$$

3. $\mathbb{E}[[f[(X)] < +\infty$, $\frac{k_n}{n} \xrightarrow[n\to+\infty]{} 0$. Montrons que :

   $$\frac{1}{k_n}\sum_{j=1}^{k_n} \mathbb{E}[|f(X) - f(X_{(j)}(X))|] \xrightarrow[n\to+\infty]{} 0$$

Si μ est une mesure sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, pour tout $f \in \mathcal{L}^1(\mu)$, $\exists f^\varepsilon$ uniformément continue et bornée telle que :

$$\int |f - f^\varepsilon|(x)\mu dx \leqslant \varepsilon$$

Lemme de Stone :

$$\frac{1}{k_n}\sum_{j=1}^{k_n} \mathbb{E}[|f(X_{(j)}(x)|] \leqslant \gamma_d \mathbb{E}[|f|(X)]$$

4. Soit $\varepsilon > 0$, f^ε, A associé à l'UC de f^ε.

$$\frac{1}{k_n}\sum_{j=1}^{k_n} \mathbb{E}\big[|f(X) - f(X_{(j)}(X))|\big] \leqslant \frac{1}{k_n}\sum_{j=1}^{k_n} \underbrace{\mathbb{E}\big[|f(X) - f^\varepsilon(X)|\big]}_{\leqslant \varepsilon} + \underbrace{\mathbb{E}\big[|f^\varepsilon(X) - f(X_{(j)}(X))|\big]}_{B_j} + \underbrace{\mathbb{E}\big[|f^\varepsilon(X) - f^\varepsilon(X_{(j)}(X))|\big]}_{C_j}$$

On a

$$\frac{1}{k_n}\sum_{j=1}^{k_n} B_j \leqslant \gamma_d \mathbb{E}[|f - f^\varepsilon|(X)] \quad \text{par le lemme de Stone}$$

$$\begin{aligned} C_j &\leqslant \mathbb{E}\big[|f^\varepsilon(X) - f^\varepsilon(X_{(j)}(X))|\mathbb{1}_{|X - X_{(j)}(X)| \geqslant A} + |f^\varepsilon(X) - f^\varepsilon(X_{(j)}(X))|\mathbb{1}_{|X - X_{(j)}(X)| < A}\big] \\ &\leqslant \varepsilon + \mathbb{E}\big[\underbrace{|f^\varepsilon(X) - f^\varepsilon(X_{(j)}(X))|}_{\text{borné}}\underbrace{\mathbb{1}_{|X - X_{(j)}(X)| \geqslant A}}_{\longrightarrow 0 \ \mathbb{P}-\text{p.s.}}\big] \\ &\leqslant \varepsilon + u_n \end{aligned}$$

Alors,

$$\forall \varepsilon, \ \overline{\lim_n} \frac{1}{k_n}\sum_{j=1}^{k_n} \mathbb{E}[\ldots] \leqslant (2 + \gamma_d)\varepsilon$$

5. Montrons que :

$$\mathbb{E}\left(|\mathbb{1}_{\hat{f}'_n(X) \neq Y} - \mathbb{1}_{\hat{f}_n(X) \neq Y}\right) \xrightarrow[n \to +\infty]{} 0$$

$$\begin{aligned} \mathbb{E}\left(|\mathbb{1}_{\hat{f}'_n(X) \neq Y} - \mathbb{1}_{\hat{f}_n(X) \neq Y}\right) &= \mathbb{P}(\hat{f}'_n(X) \neq \hat{f}_n(X)) \\ &\leqslant \sum_{j=1}^{k} \mathbb{E}(|\eta(X) - \eta(X_{(j)} - X))|) \xrightarrow[n \to +\infty]{} 0 \end{aligned}$$

Exercice 5.3

1. $\hat{f}$ est locale donc d'après l'exercice précédent :

$$\lim_{n \to +\infty} \mathbb{E}\left[R_{\mathbb{P}}^{D_n}(\hat{f})\right] = \lim_{n \to +\infty} \mathbb{P}\big(\hat{f}'_n(X) \neq Y\big)$$

Soit $n \geqslant k$.

$$\begin{aligned}\mathbb{P}(\hat{f}'_n(X) \neq Y) &= \mathbb{P}(Y'_{(1)}(X) \neq Y) \\ &= \mathbb{E}\left[\mathbb{P}(Y'_{(1)}(X) \neq Y|X)\right]\end{aligned}$$

Or sachant X, on a $Y'_{(1)}(X) \perp Y$ et les deux sont des Bernouilli de paramètre $\eta(X)$. Donc

$$\mathbb{P}(Y'_{(1)}(X) \neq Y|X) = 2\eta(X)(1-\eta(X))$$

2. Z v.a. dans $I \subset \mathbb{R}$ $f, g \colon I \longrightarrow \mathbb{R}$ croissantes. Montrons que :

$$\mathbb{E}[f(Z)g(Z)] - \mathbb{E}[f(Z)]\mathbb{E}[g(Z)] \geqslant 0$$

$$\begin{aligned}\int f(z)g(z)\mathbb{P}_z(dz) - \int f(\tilde{z}\mathbb{P}_Z(d\tilde{z}) \int g(z)\mathbb{P}_z(dz) &= \int\int (f(z)-f(\tilde{z}))g(z)\mathbb{P}_Z(dz)\mathbb{P}_z(d\tilde{z}) \\ &= \int\int \mathbb{1}_{\tilde{z}<z}(f(z)-f(\tilde{z}))g(z)\mathbb{P}_Z(dz)\mathbb{P}_Z(d\tilde{z}) \\ &\quad + \int\int \mathbb{1}_{\tilde{z}>z}(f(z)-f(\tilde{z}))g(z)\mathbb{P}_z(dz)\mathbb{P}_z(d\tilde{z}) = B\end{aligned}$$

$$B = \int\int \mathbb{1}_{z>\tilde{z}}(f(\tilde{z})-f(z))g(\tilde{z})\mathbb{P}_Z(d\tilde{z})\mathbb{P}_Z(dz)$$

Finalement,

$$= \int \mathbb{1}_{z>\tilde{z}}(f(z)-f(\tilde{z}))(g(z)-g(\tilde{z}))d\mathbb{P}_Z(z)d\mathbb{P}_Z(\tilde{z}) \geqslant 0$$

3.

$$R^{\mathrm{NN}}_{\mathbb{P}} = 2\mathbb{E}[f(\eta(X))g(\eta(X))]$$

$f(p) = \min(p, 1-p)$ et $g = 1 - \min(p, 1-p)$. Alors par 2) :

$$R^{\mathrm{NN}}_{\mathbb{P}} \leqslant 2\mathbb{E}[f(\eta(X))]\mathbb{E}[1-\eta(X)] \leqslant 2R^*_{\mathbb{P}}(1-R^*_{\mathbb{P}})$$

Exercice 5.4

k impair.

$$\lim_{n\to+\infty} \mathbb{E}[R^{D_n}_{\mathbb{P}}(\hat{f})] = R^{k-\mathrm{NN}}_{\mathbb{P}}$$

$$\lim_{n\to+\infty} \mathbb{P}(Y \neq \hat{f}'_n(X)) = R^{k-\mathrm{NN}}_{\mathbb{P}}$$

$$\begin{aligned}\mathbb{P}(Y \neq \hat{f}'_n(X)) &= \mathbb{E}\big[\mathbb{P}(Y \neq f'_n(X) \mid X)\big] \\ &= \mathbb{E}\left[\mathbb{P}\left(Y=1, \sum Y'_{(j)} > \frac{k}{2} \mid X\right) + \mathbb{P}\left(Y=0, \sum Y'_{(j)} \leqslant \frac{k}{2} \mid X\right)\right]\end{aligned}$$

Sachant X, Y, $(Y_{(j)}(X))$ i.i.d. $\sim \mathcal{B}(\eta(X))$, on obtient :

$$= \mathbb{E}\Big[\sum_{e=\frac{k+1}{2}}^{k} \underbrace{\mathbb{P}(Y=1|X)}_{\eta(X)} \underbrace{\mathbb{P}(\sum Y'_{(j)}(X) = e|X)}_{\binom{k}{e}\eta(X)^e(1-\eta(X))^{k-e}}\Big] + \sum_{e=1}^{\frac{k-1}{2}} \mathbb{P}(Y=0|X)\dots$$