

Exercice 1

1) Soient P et Q deux mesures sur l'espace mesurable $(E, \mathcal{F})$. Alors

$$d_{VT}(P, Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)|$$

$\langle A \text{ mesurable} \rangle$

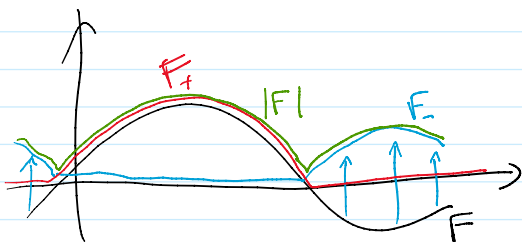
2) Soit μ une mesure (pas forcément de proba) sur $(E, \mathcal{F})$ et $\begin{cases} P \\ Q \end{cases}$ a une densité $\begin{matrix} f \\ g \end{matrix}$ par rapport à μ ^{supposons que}

Remarques: • $\forall A \in \mathcal{F}, P(A) = \int_A f(x) \mu(dx)$

• Pour toute fonction F à valeurs réelles,

$$F(x) = \underbrace{F_+(x)}_{\geq 0} - \underbrace{F_-(x)}_{\geq 0}$$

$$|F(x)| = F_+(x) + F_-(x)$$



$$\int (f-g)_+(x) \mu(dx)$$

Soit $A \in \mathcal{F}$.
$$P(A) - Q(A) = \int_A f(x) \mu(dx) - \int_A g(x) \mu(dx) = \int_A (f-g)(x) \mu(dx)$$

$$= \underbrace{\int_A (f-g)_+(x) \mu(dx)}_{\geq 0} - \underbrace{\int_A (f-g)_-(x) \mu(dx)}_{\geq 0}$$

$$\leq \int_A (f-g)_+(x) \mu(dx)$$

Donc
$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) \leq \sup_{A \in \mathcal{F}} \left(\int_A (f-g)_+(x) \mu(dx) \right)$$

Si $A \subset B$,
$$\int_A (f-g)_+(x) \mu(dx) \leq \int_B (f-g)_+(x) \mu(dx)$$

Pour tout $A \in \mathcal{F}$, $A \subset E$

donc

$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) \leq \int_E (f-g)_+(x) \mu(dx)$$

$\leftarrow$ pas obligé de préciser E

$$\int_A = \int \mathbf{1}_A \dots$$

Pour l'autre sens de l'inégalité, prenons $\tilde{A} = \{x : f(x) \geq g(x)\}$. $\tilde{A}$ est mesurable ($A \in \mathcal{F}$).

$$\begin{aligned} \int (f-g)_+(x) \mu(dx) &= \int (f-g)(x) \mathbf{1}_{f(x) \geq g(x)} \mu(dx) \\ &= \int (f-g)(x) \mathbf{1}_{\tilde{A}}(x) \mu(dx) = \int_{\tilde{A}} (f-g)(x) \mu(dx) \end{aligned}$$

D'où

$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) \geq P(\tilde{A}) - Q(\tilde{A}) = \int_{\tilde{A}} (f-g)_+(x) \mu(dx)$$

Donc

$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) = \int (f-g)_+(x) \mu(dx)$$

71/15

Donc

$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) = \int (f - g)_+(x) \mu(dx)$$

<< Soit $A \in \mathcal{F}$, alors $P(A) - Q(A) \leq qch$ donc $\sup_{A \in \mathcal{F}} (P(A) - Q(A)) \leq qch$.

Prends un $\tilde{A}$ bien choisi, alors

$$\sup_{A \in \mathcal{F}} (P(A) - Q(A)) \geq P(\tilde{A}) - Q(\tilde{A}) \Rightarrow$$

Pour la deuxième partie de la question,

$$\sup_{B \in \mathcal{F}} (Q(B) - P(B)) = \int (g - f)_+(x) \mu(dx) \underset{\substack{\uparrow \\ \text{1}^{ère} \text{ partie}}}{=} \int (f - g)_-(x) \mu(dx) \underset{F_+ = (-F)_-}{=} \int (f - g)_-(x) \mu(dx).$$

$$\begin{aligned} 3) \quad \int (f - g)_+(x) \mu(dx) - \int (f - g)_-(x) \mu(dx) &= \int ((f - g)_+(x) - (f - g)_-(x)) \mu(dx) \\ &= \int (f - g)(x) \mu(dx) \\ &= \int f \mu(dx) - \int g \mu(dx) \\ &= P(E) - Q(E) = 1 - 1 = 0 \end{aligned}$$

$\uparrow$ car P et Q sont des mesures de proba.

$$\int (f - g)_+(x) \mu(dx) + \int (f - g)_-(x) \mu(dx) = \int |f - g|(x) \mu(dx) = \|f - g\|_1.$$

$$\|F\|_1 = \int |F(x)| \mu(dx) \text{ et } \forall p \geq 1 \quad \|F\|_p^p = \int |F(x)|^p \mu(dx)$$

$$\begin{aligned} 4) \quad \text{D'après la question 3,} \quad \int (f - g)_+(x) \mu(dx) &= \int (f - g)_-(x) \mu(dx) \quad (\text{1}^{ère} \text{ Egalité}) \\ &= \frac{1}{2} \|f - g\|_1 \quad \leftarrow \int (f - g)_+ + \int (f - g)_- = 2 \int (f - g)_+ = \|f - g\|_1 \end{aligned}$$

D'après la question 3,

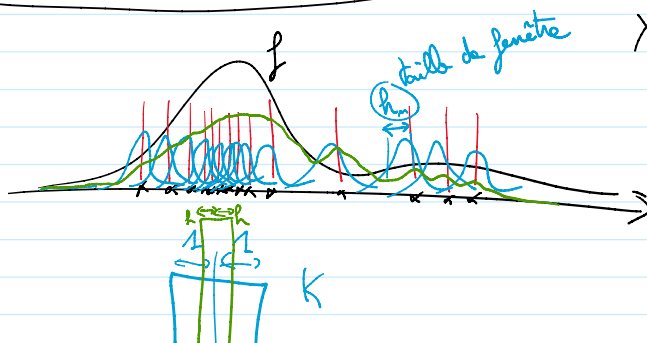
$$\begin{aligned} d_{TV}(P, Q) &\stackrel{q1}{=} \sup_{A \in \mathcal{F}} |P(A) - Q(A)| = \max \left(\sup_{A \in \mathcal{F}} (P(A) - Q(A)), \sup_{B \in \mathcal{F}} (Q(B) - P(B)) \right) \\ &\stackrel{q2}{=} \max \left(\int (f - g)_+(x) \mu(dx), \int (f - g)_-(x) \mu(dx) \right) \\ &\stackrel{\text{d'après } q4}{=} \frac{1}{2} \|f - g\|_1 \end{aligned}$$

Exercice 2

Pourquoi les estimateurs à noyau?

$$P_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$$

$$\hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i)$$



$X_i \sim f$

| dispo: marche pas

~: on est un peu myope



$$\begin{aligned}
 1) \quad \mathbb{E}[\hat{f}_{n,K}(x)] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n K_h(x - X_i)\right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[K_h(x - X_i)] \\
 &= \mathbb{E}[K_h(x - X_1)] \quad \text{car les } X_i \text{ sont iid} \\
 &= \int K_h(x - z) f(z) dz \\
 \text{chgt de var } z = x - y \quad &= \int K_h(y) f(x - y) dy = (K_h * f)(x).
 \end{aligned}$$

2) Pour toute v.a. réelle X et tout $a \in \mathbb{R}$,

$$\begin{aligned}
 \mathbb{E}[(X - a)^2] &= \mathbb{E}[(X - \mathbb{E}[X] + \mathbb{E}[X] - a)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2] + \mathbb{E}[(\mathbb{E}[X] - a)^2] \\
 &\quad + 2\mathbb{E}[(X - \mathbb{E}[X])(\mathbb{E}[X] - a)] \\
 &= \text{Var}(X) + (\mathbb{E}[X] - a)^2 + 2(\mathbb{E}[X] - a)(\mathbb{E}[X] - \mathbb{E}[X]) \\
 &= \text{Var}(X) + (\mathbb{E}[X] - a)^2
 \end{aligned}$$

Application: $X = \hat{f}_{n,K}(x)$ $a = f(x)$
 $\mathbb{E}[X] = (K_h * f)(x)$

$2(\mathbb{E}[X] - a)\mathbb{E}[X - \mathbb{E}[X]] = 2(\mathbb{E}[X] - f(x)) \cdot 0 = 0$

$$3) \quad f(x) = \int K_h(y) f(x) dy \quad \text{car } K_h \text{ est d'intégrale 1}$$

Conséquence du Théorème de Rolle: pour tout a, b , il existe c entre a et b t.q.

$$f(b) - f(a) = f'(c)(b - a).$$

$$\begin{aligned}
 |(K_h * f)(x) - f(x)| &= \left| \int K_h(y) f(x - y) dy - f(x) \right| = \left| \int K_h(y) f(x - y) dy - \int K_h(y) f(x) dy \right| \\
 &= \left| \int K_h(y) (f(x - y) - f(x)) dy \right| \\
 &\leq \int |K_h(y)| |f(x - y) - f(x)| dy \quad \text{inégalité triangulaire} \\
 &\leq \int K_h(y) L |y| dy \quad \text{car } K_h \geq 0 \text{ donc } |K_h| = K_h, \text{ Rolle } \oplus \|f'\|_\infty \leq L \\
 &= L \int |y| \frac{1}{h} K\left(\frac{y}{h}\right) dy \quad \text{chgt var } y = \frac{y}{h} \\
 &= L h \int |y| K(y) dy \leq L h M.
 \end{aligned}$$

$$\begin{aligned}
 4) \quad \mathbb{E}[(\hat{f}_{n,K}(x) - (K_h * f)(x))^2] &= \text{Var}(\hat{f}_{n,K}(x)) \quad (\text{que 1}) \\
 &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n K_h(x - X_i)\right) \\
 &= \frac{1}{n} \text{Var}(K_h(x - X_1)) \quad \text{var. iid.} \\
 &= \frac{1}{n} \text{Var}(K_h(x - X_1)) \\
 &= \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n K_h(x - X_i)\right) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(K_h(x - X_i)) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(K_h(x - X_1)) \\
 &= \frac{1}{n} \text{Var}(K_h(x - X_1)) \\
 &= \frac{1}{n} \text{Var}(K_h(x - X_1))
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \text{Var}(K_h(x-x_i)) \quad \text{var. i.i.d.} &= \frac{1}{n} \text{Var}(Z_1) \\
&= \frac{1}{n} \left[\underbrace{\mathbb{E}[K_h(x-x_i)^2]}_{\geq 0} - \underbrace{\mathbb{E}[K_h(x-x_i)]^2}_{\geq 0} \right] \\
&\leq \frac{1}{n} \mathbb{E}[K_h(x-x_i)^2] \\
&= \frac{1}{n} \int K_h(x-z)^2 f(z) dz & \text{chgt var } z = x - h_n y \\
&= \frac{1}{n} \int K_h(h_n y)^2 f(x - h_n y) h_n dy \\
&= \frac{1}{n} \int \frac{1}{h_n^2} K(y)^2 f(x - h_n y) h_n dy & |f(x - h_n y) - f(x)| \leq \|f'\| h_n |y| \\
&= \frac{1}{nh_n} \int K(y)^2 f(x - h_n y) dy & \text{Th Rolle: } f(x - h_n y) \leq f(x) + \|f'\| h_n |y| \\
&\leq \frac{1}{nh_n} \left(\int K(y)^2 L dy + \int K(y)^2 L h_n |y| dy \right) & \leq L + L h_n |y| \\
&\leq \frac{1}{nh_n} (LM + L h_n M)
\end{aligned}$$

5) 2, 3 et 4 ensemble :

$$\mathbb{E}[(\hat{f}_{n,K}(x) - f(x))^2] \leq \underbrace{(LMh_n)^2}_{\rightarrow 0} + \underbrace{\frac{1}{nh_n}}_{\rightarrow 0} \underbrace{(LM + L h_n M)}_{\text{borné}} \xrightarrow{n \rightarrow +\infty} 0 \quad \square$$

6) Th convergence dominée :

Après $d_{KT}(\hat{f}_{n,K}, f) \leq \text{constante} \int \underbrace{\mathbb{E}[(\hat{f} - f)^2(x)]}_{\xrightarrow{n \rightarrow +\infty} 0} dx$

$\leq 1_{[-(c,n), (c,n)]} \times 1$ intégrable.

à partir d'un certain rang ≤ 1