Chapitre 3 : Non paramétrique

I. Exemples de modèles et pbms.

Ex 1 : (X_i, Y_i), $Y_i = f(X_i) + \varepsilon_i$, ε_i iid $\sim \mathcal{N}(0, \sigma^2)$. f à trouver.

- $X_i = \frac{i}{n}$, $Y_i = f(\frac{i}{n}) + \varepsilon_i$,
- X_i iid, g s/ $[0,1]$

Ex 2 : $X_1, \dots, X_n$ iid de densité f / Lebesgue. Trouver f.

Ex 3 :
- Modèle bruit blanc : $dX_t^\varepsilon = f(t)dt + \varepsilon dW_t$, $(W_t, t \in [0,1])$ brownien.
- $dX_t = f(X_t) + \sigma(X_t) dW_t$, $X_1, \dots, X_T$. Estimer f, σ.

On peut vouloir reconstruire f en entier, $f(t_0)$ ou encore $T(f)$ une fonctionnelle de f (son intégrale, p/ ex).

Point de vue minimax :

On se donne une fc de perte $q(f) \to \hat{q}$; $\ell(\hat{q}, q(f))$.
un ensb $V = \{f \text{ possibles}\}$

$$\text{minimax} := \inf_{\hat{q}} \sup_{f \in V} \mathbb{E}_f [\ell(\hat{q}, q(f))]$$

On va tester 2 inégalités :
1. Borne inférieure $\exists c$ minimax $\geq c\, r(n)$
2. Borne sup $\exists d \geq c$ — $\leq d\, r(n)$

on borne ses ambi° $\quad r(n) \xrightarrow[n \to \infty]{} 0$

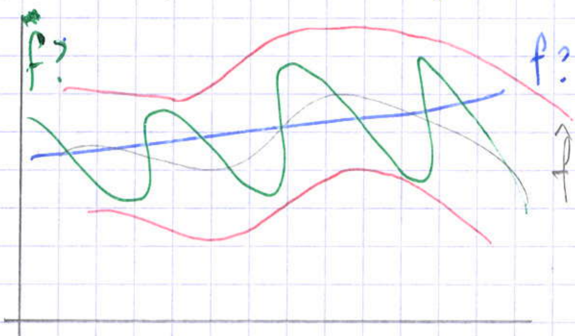
Pour calculer le sup, on trouve un estimateur q^* t.q.

$$\sup_{f \in V} \mathbb{E}_f [\ell(\hat{q}, q(f))] \leq D r(n).$$

La vitesse minimax n'est pas liée à la 'taille' de l'espace dans lequel on veut trouver f.

Quelle perte choisir ? $\ell(f, \hat{f}) = \|f - \hat{f}\|_2^2$, $\|f - \hat{f}\|_p^p$, $p < \infty$, $\|f - \hat{f}\|_\infty$

Si on considère une seule fc de perte, p/ ex $\|f - \hat{f}\|_\infty$, qu'on a $\|f - \hat{f}\|_\infty \leq \varepsilon_n$, on n'a pas idée du cptement de f :



Quel ensb V ? $V_\alpha(M) := \{ f:[0,1] \to \mathbb{R} \ / \ |f(0)| \leq M, \ |f(x)-f(y)| \leq M|x-y|^\alpha \ \forall x,y \in [0,1] \}$. $0<\alpha\leq 1$

1. Bornes inférieures.

Thm (borne inf) : $\inf_{\hat f} \sup_{f \in V_\alpha(M)} (\mathbb{E}_f \|\hat f - f\|_p^p) \geq c n^{-\frac{p\alpha}{1+2\alpha}}$

Rq : $(\_\_\_\_\_\_)^{1/p} \geq c n^{-\frac{\alpha}{1+2\alpha}}$ or $\alpha \mapsto \frac{\alpha}{1+2\alpha} \nearrow$, max en $\alpha = 1 \Rightarrow$ vaut $1/3$.

Strictement moins bonne qu'en paramétrique $n^{-1/2}$ (+ difficile).

Si $\alpha \to 0$, $\frac{\alpha}{1+2\alpha} \to 0$, indique que si on ne met pas de droite de régul. s/ f, on n'a pas vitesse $\to 0$.

Idée dém : Liée à 2 "ingrédients" :

$\to$ Géométrie des espaces de proba (mettre une dist^{ce} entre deux lois).

dist^{ce} d'Hellinger : P, Q déf sur un m̂ espace.

$$H^2(P,Q) = \frac{1}{2}\left(\int (\sqrt{dP} - \sqrt{dQ})^2 \right) \quad \left(:= \frac{1}{2} \int \left(\sqrt{\frac{dP}{d\mu}} - \sqrt{\frac{dQ}{d\mu}} \right)^2 \right)$$

où $\mu \gg P, Q$ - ne dépd pas de μ dominante (exo)

⚠ Pas tt à fait une distance.

On a $H^2(P,Q) = \frac{1}{2} \int (dP + dQ - 2\sqrt{dP}\sqrt{dQ}) = 1 - \int \sqrt{dPdQ} =: 1 - \rho(P,Q)$

Si on a affaire à un n-échantillon ⫫, on a

$$\begin{aligned} H^2(P^{\otimes n}, Q^{\otimes n}) &= 1 - \rho(P^{\otimes n}, Q^{\otimes n}) \\ &= 1 - \rho(P,Q)^n \\ &\leq (1 - \rho(P,Q))n \end{aligned}$$
(car $\rho(P,Q) \leq 1$ et $(1-u^n) = (1-u)(1+\dots+u^{n-1})$)

d'où $\boxed{H^2(P^{\otimes n}, Q^{\otimes n}) \leq nH^2(P,Q)}$.

Lemme (LeCam) : U, V deux v.a. réelles t.q. $U+V \geq a > 0$. Alors si $H^2(P,Q) \leq \frac{1}{2}$,

$$\int U dP + \int V dQ \geq \frac{a}{8}$$

Interpréta^o : φ un test, $\{P\}$ contre $\{Q\}$, $U = \varphi$, $V = 1-\varphi$

$\mathbb{E}_P \varphi$: risque de rejeter P
$\mathbb{E}_Q(1-\varphi)$: Q

Ce que dit le lemme : si $H^2(P,Q) \leq \frac{1}{2}$, alors $\mathbb{E}_P \varphi + \mathbb{E}_Q(1-\varphi) \geq \frac{1}{8}$ $(a=1)$.

$\Rightarrow$ on a du mal à estimer les risques.

Nota^o : $dP \wedge dQ := \inf\left(\frac{dP}{d\mu}, \frac{dQ}{d\mu}\right) d\mu$, μ dominante.

dém : On écrit $\int U dP + \int V dQ \geq \int U \, dP \wedge dQ + \int V \, dP \wedge dQ$

$$\geq a \int dP \wedge dQ$$

Lien ac $H^2(P,Q)$: $\int \sqrt{dPdQ} = \int \sqrt{dP \wedge dQ} \sqrt{dP \vee dQ}$

$$\underset{C.S.}{\leq} \left(\int dP \wedge dQ\right)^{1/2} \left(\int dP \vee dQ\right)^{1/2}$$
$$\leq \left(\int dP \wedge dQ\right)^{1/2} \underbrace{\left(\int dP + dQ\right)^{1/2}}_{= 2^{1/2}}$$

donc $\int dP\wedge dQ \geq \frac{1}{2}\left(\int\sqrt{dPdQ}\right)^2$, et si $H^2(P,Q)\leq\frac{1}{2}$, $\sqrt{dPdQ}\geq\frac{1}{2}$,

d'où $\int dP\wedge dQ \geq \frac{1}{8}$ ∎.

$\to$ $\sup_{f\in V}\mathbb{E}_f\|\hat{f}-f\|_p^p$, $\hat{f}$ arbitraire.

Pour mettre en échec $\hat{f}$ q ne donne pas la bonne vitesse, on prend par ex $p=1$,

$\psi\in V_\alpha(M)$ t.q. $\sup_{f\in V_\alpha(M)}\mathbb{E}_f\|f-\hat{f}\|_p^p \geq \sup_{f\in(\psi,-\psi)}\mathbb{E}_f\|\hat{f}-f\|$ $(*)$

Si de plus $\psi=\gamma g$, $0<\gamma\leq 1$, $g\in V_\alpha(M)$, $\int g^2=1$, $\int|g|=c>0$,

$$(*)\geq\frac{1}{2}\Big(\underbrace{\mathbb{E}_\psi\|\hat{f}-\psi\|_1}_{U}+\underbrace{\mathbb{E}_{-\psi}\|\hat{f}+\psi\|_1}_{V}\Big)$$

$\Rightarrow$ $U+V\geq 2\|\psi\|=2\gamma c$.

D'après le lemme, si $H^2(P_\psi^{\otimes n},P_{-\psi}^{\otimes n})\leq\frac{1}{2}$, alors

$$\sup_{f\in V_\alpha(M)}\mathbb{E}_f\|\hat{f}-f\|\geq\frac{\gamma c}{8}\left(=\frac{c'}{\sqrt{n}}\right)$$

Reste à vérifier cette majoraᵉ :

$$\rho(P_\psi,P_{-\psi})=\int_{\mathbb{R}}\int_{[0,1]}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-\psi(x))^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y+\psi(x))^2}\right)^{1/2}dx\,dy.$$

en dominant par $\mu=\lambda_{[0,1]}\times\lambda_{\mathbb{R}}$ $\left(\varepsilon_i\sim\mathcal{N}(0,1)\perp\!\!\!\perp, \perp\!\!\!\perp X_i\sim\mathrm{Unif}([0,1])\right)$.

$$=\int_{\mathbb{R}\times[0,1]}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(y-\frac{\psi(x)+\psi(x)}{2}\right)^2}e^{\frac{1}{4}(\psi(x)-(-\psi(x)))^2}$$

$$=\int_{[0,1]}e^{\frac{1}{4}\cdot 4\psi(x)^2}dx\underset{e^u\geq 1-u}{\geq}1-\int_{[0,1]}\psi^2(x)dx=1-\gamma^2$$

Par conséquent, $H^2(P_\psi^{\otimes n},P_{-\psi}^{\otimes n})\leq n(1-(1-\gamma^2))=n\gamma^2$

Il suffit de choisir γ t.q. $n\gamma^2\leq\frac{1}{2}$ i.e. $\gamma:=\frac{1}{\sqrt{2n}}$

En vrai, on prend des fc g peignes à supp dans $[0,1]$, et on les 'pince' :

$\sum_{k=0}^{2^j-1}\varepsilon_k 2^{j/2}g(2^j-k)$, $\varepsilon_k\in\{\pm1\}$.

L'espace 'restreint' en $(*)$ est alors $\Gamma:=\left\{f=\sum_{k=0}^{2^j-1}\gamma u_k 2^{j/2}g(2^jx-k), u_k\in\{\pm 1\}\right.$

On sépare alors $\int\|\hat{f}-f\|_p^p=\sum_k\int_{I_{jk}}\|\hat{f}-f\|_p^p$

2. Bornes supérieures.

Dans la famille des estimateurs, il y a 2 familles :
- estimateurs à noyaux
- par séries orthogonales.

$X_1, \dots, X_n$ iid de densité f sur $[0,1]$.

$$f(x) = \lim_{h \to 0} \frac{1}{2h} \mathbb{P}(x-h \leq X \leq x+h)$$

ça nous donne un candidat $\hat{f}$:

$$\hat{f}(x) := \frac{1}{2h} \cdot \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[x-h, x+h]}(X_i) = \frac{1}{n} \sum_i K_h(x - X_i)$$

où $K_h(x) = \frac{1}{2h} \mathbb{1}_{[-h,h]}(x) = \frac{1}{2h} \mathbb{1}_{[-1,1]}\left(\frac{x}{h}\right)$.

On va ensuite imposer des régularités à K_h.