Estimation adaptative de densité par méthode de Goldenshluger-Lepski

Eddie Aamari
15/01/2014

Références : . Goldenshluger-Lepski 2010, 2013, 2011 ; Lepski 2013
. Cours N. Klutchnikoff (2013)

But : $X_1, \dots, X_n$ va réelles iid de loi à densité f / à Lebesgue, $x \in \mathbb{R}$ fixé

estimer $f(x)$, avec comme risque

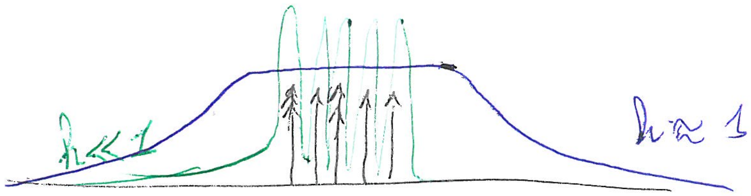
$$R_q(f, \hat{f})^q = \mathbb{E}|\hat{f}(x) - f(x)|^q \quad , \infty > q \geq 1$$

1/ Si l'on connaît la régularité de f.

2/ Si l'on ne la connaît pas.

Estimateur à noyau : $K:\mathbb{R}\to\mathbb{R}$, $h>0$

$$\hat f_h(x) := \frac{1}{n}\sum_{i=1}^{n} \underbrace{\frac{K\left(\frac{x-X_i}{h}\right)}{h}}_{K_h(x-X_i)}$$



Le choix de $h=h_n$ dépend de la régularité de f.

Def : $s>0$, $L\geq 0$. $s=\lfloor s\rfloor+\alpha$

$$\mathcal{C}^s(L)=\left\{ g:\mathbb{R}\to\mathbb{R} \;\middle|\; \begin{array}{l} \cdot\ g \text{ est une densité}\\ \cdot\ g\in\mathcal{C}^{\lfloor s\rfloor}\\ \cdot\ |g^{(\lfloor s\rfloor)}(y)-g^{(\lfloor s\rfloor)}(z)|\leq L|y-z|^{\alpha}\end{array}\right\}$$

est appelée classe de Hölder.

Rq : · hypothèses sur K à suivre
· tout ce qui est fait ici s'étend à un risque intégré $\left(\mathbb{E}\|\hat f - f\|_{L^p}^{q}\right)$

1/ Régularité connue

On suppose que $f \in \mathcal{C}^s(L)$.

$$\hat{f}_h(x) - f(x) = \underbrace{\left(\hat{f}_h(x) - \mathbb{E}\hat{f}_h(x)\right)}_{\text{terme stochastique centré}} + \underbrace{\left(\mathbb{E}\hat{f}_h(x) - f(x)\right)}_{\text{biais}}$$

Cas $q = 2$:

* Terme stochastique

$$\boxed{\mathbb{E}\left(\hat{f}_h(x) - \mathbb{E}\hat{f}_h(x)\right)^2} = \mathrm{Var}\left(\hat{f}_h(x)\right)$$

$$= \frac{1}{n}\mathrm{Var}\left(K_h(x - X_1)\right)$$

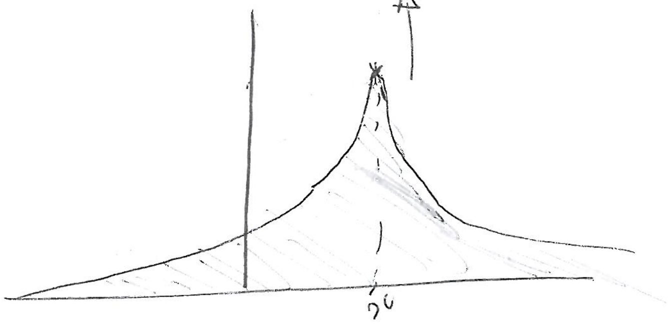
$$\leq \frac{1}{n}\mathbb{E}K_h(x - X_1)^2$$

$$= \frac{1}{n}\int K_h(x-y)^2 f(y)\,dy$$

$$\boxed{\leq \frac{\|f\|_\infty \|K\|_2^2}{nh}}$$

Rq : le $\|f\|_\infty$ ne coûte pas cher : Si $0 < s < 1$: $\forall h$, $f(\cancel{x}) \leq L|h|^s + f(x+h)$

⊕ Intégration dh



$*$ Biais : $\mathbb{E}\hat{f}_h(x) = \mathbb{E} K_h(x-X_1)$ $\qquad y = x+hv$

$$= \int K_h(x-y) f(y)\,dy = \int K(v) f(x+hv)\,dv$$

$$= K_h * f(x)$$

$$\boxed{\int K(y)\,dy = 1}$$

$$K_h * f(x) - f(x) = \int K(v)\big(f(x+hv) - f(x)\big)\,dv$$

Developpement de Taylor (avec reste integral ⊕ egalité de la moyenne)

$$f(x+hv) - f(x) = \sum_{k=1}^{\lfloor s\rfloor} \frac{f^{(k)}(x)}{k!}(hv)^k + \frac{f^{(\lfloor s\rfloor)}(\tilde{x}_v) - f^{(\lfloor s\rfloor)}(x)}{\lfloor s\rfloor !}(hv)^{\lfloor s\rfloor}$$

avec $x \leq \tilde{x}_v \leq x + hv$

$$\Downarrow$$

$$|\tilde{x}_v - x| \leq h|v|$$

$$\boxed{\forall 1 \leq k \leq \lfloor s\rfloor, \int K(u) u^k\,du = 0}$$

Rq : Ça force K à ne pas être ≥ 0 car $\int K(u)\underbrace{u^2}_{\geq 0}\,du = 0$

mais c'est pas bien grave

$$|K_h * f(x) - f(x)| \leq \frac{h^{\lfloor s\rfloor}}{\lfloor s\rfloor !}\int v^{\lfloor s\rfloor}|K(v)|\,\underbrace{|f^{(\lfloor s\rfloor)}(\tilde{x}_v) - f^{(\lfloor s\rfloor)}(x)|}\,dv$$

$$\leq L|\tilde{x}_v - x|^{s-\lfloor s\rfloor}$$

$$\leq L h^{s-\lfloor s\rfloor}|v|^{s-\lfloor s\rfloor}$$

$$\leq \frac{h^s}{\lfloor s\rfloor !} L \underbrace{\int |v|^s |K(v)|\,dv}_{K_s}$$

D'où $\boxed{|K_h * f(x) - f(x)| \leq \frac{L K_s}{\lfloor s\rfloor !} h^s}$

On obtient : $\mathbb{E}\,|\hat{f}_h(x) - f(x)|^2 \leq C\left(\frac{1}{nh} + h^{2s}\right)$

Avec un coup de Rosenthal on obtient le resultat pour $q \geq 2$
Pour $q \leq 2$ on utilise la convexité de $x \mapsto x^{q/2}$

$$\Rightarrow \quad \mathbb{E}\,|\hat{f}_h(x) - f(x)|^q \leq C_{q,L}\left(\left(\frac{1}{nh}\right)^{q/2} + h^{sq}\right)$$

Compromis lorsque $\left(\frac{1}{nh}\right)^{q/2} \asymp h^{sq}$, i.e $h_n \asymp n^{-\frac{1}{2s+1}}$

Pour un tel choix de $h_n \asymp n^{-\frac{1}{2s+1}}$, on obtient alors la borne

$$\boxed{\mathbb{E}\,|\hat{f}_h(x) - f(x)|^q \leq C\, n^{-\frac{sq}{2s+1}}}$$

<u>Rq</u> : . C est uniforme sur $\mathcal{C}^s(L)$

. Vitesse classique en non parametrique

. Moralement, pour $s \to \infty$, on retombe sur une vitesse parametrique

$\hookrightarrow$ Cas analytique : $\sqrt{\frac{\log n}{n}}$

. Si l'on ne connait pas s a priori, on est cuits : Méthode de G-L

$\Updownarrow$

Choisir une <u>bonne</u> fenêtre en fonction des données

2/ Régularité adaptative

Soit $\mathcal{H} \subset \mathbb{R}_+^*$ un ensemble fini de fenêtres

On va décrire une méthode pour choisir $h \in \mathcal{H}$ réalisant le meilleur compromis, sans avoir à connaître Δ.

$$R_q(\hat{f}_h, f)^q \leq C\left[\underbrace{|K_h * f(t) - f(t)|^q}_{\text{Biais} \leq ???} + \left(\frac{1}{nh}\right)^{q/2}\right]$$

Estimateur auxiliaire surlissé :

$$\forall h, \eta \in \mathcal{H}, \quad \hat{f}_{h,\eta}(x) = \frac{1}{n}\sum_{i=1}^{n}(K_h * K_\eta)(x - X_i)$$

Rq : $\cdot\ K_h * K_\eta = K_\eta * K_h$ donc $\hat{f}_{h,\eta} = \hat{f}_{\eta,h}$

. À $h \in \mathcal{H}$ fixé, on fait parcourir $\mathcal{H}$ à η et on compare $\hat{f}_h$ à $\hat{f}_{h,\eta}$

. Noyau K gaussien : $\hat{f}_{h,\eta}(x) = \hat{f}_{h+\eta}(x)$: Surlissage (on convole un coup de plus)

Procédure de Goldenshluger - Lepski

- $B_h(\eta) = \left[|\hat{f}_\eta(x) - \hat{f}_{h,\eta}(x)| - \chi(\eta) \right]_+ \qquad \left(\chi(\eta) \xrightarrow[\eta \to 0]{} \infty \right)$

- $B_h = \underset{\eta \in H}{\mathrm{Max}}\, B_h(\eta)$

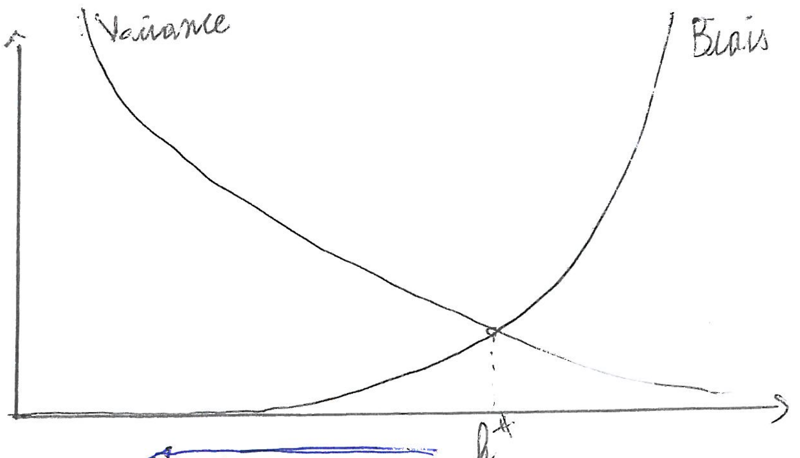
- $\hat{h} = \underset{h \in H}{\mathrm{argmin}} \left\{ B_h + \chi(h) \right\}$

- $\hat{f}(x) = \hat{f}_{\hat{h}}(x)$

Rq:

- χ est de l'ordre de l'écart type de $|\hat{f}_\eta(x) - \hat{f}_{h,\eta}(x)|$
- $B_h(\eta)$ est calculable et on espère qu'il est contrôlé par le biais $\mathbb{E}\hat{f}_\eta(x) - \mathbb{E}\hat{f}_{h,\eta}(x)$ avec grande probabilité (en fait, en norme L^q)
- Moralement équivalent à prendre

$$\hat{h} = \mathrm{Sup} \left\{ h \mid \forall \eta \leq h,\ |\hat{f}_\eta(x) - \hat{f}_{h,\eta}(x)| \leq C\sqrt{\mathrm{Var}(\hat{f}_\eta)} \right\}$$



Biais stabilisé
⊕ Variance qui explose

On décompose $|\hat{f}_{\hat{h}}(x) - f(x)|$ en passant par $f_{\hat{h},h}(x)$:

$$
\begin{aligned}
|\hat{f}_{\hat{h}}(x) - f(x)| \leq\ & \left\{ |\hat{f}_{\hat{h}}(x) - f_{h,\hat{h}}(x)| - \chi(\hat{h}) \right\}_+ + \chi(\hat{h}) \\
& + \left\{ |f_{h,\hat{h}}(x) - f_h(x)| - \chi(h) \right\}_+ + \chi(h) \\
& + |\hat{f}_h(x) - f(x)|
\end{aligned}
$$

symétrie

$$
\begin{aligned}
\leq\ & \max_{\eta \in H} B_h(\eta) + \chi(\hat{h}) \\
& + \max_{\eta \in H} B_{\hat{h}}(\eta) + \chi(h) \\
& + |\hat{f}_h(x) - f(x)|
\end{aligned}
$$

$$
\begin{aligned}
=\ & B_h + \chi(\hat{h}) \\
& + B_{\hat{h}} + \chi(h) \\
& + |\hat{f}_h(x) - f(x)|
\end{aligned}
$$

et $B_{\hat{h}} + \chi(\hat{h}) \leq B_h + \chi(h)$ par définition de $\hat{h}$

$$
|\hat{f}_{\hat{h}}(x) - f(x)| \leq 2\big(B_h + \underbrace{\chi(h)}_{\text{non aléatoire}}\big) + \underbrace{|\hat{f}_h(x) - f(x)|}_{\text{parfait, inégalité vraie pour tout } h \in H}
$$

$\longrightarrow$ Reste à borner $\mathbb{E}\, B_h^q$

$$B_h(\eta) = \left\{ \underbrace{\left|\hat{f}_\eta(x) - \hat{f}_{h,\eta}(x)\right|} - X(\eta) \right\}_+$$

$$\leq \underbrace{\left|\hat{f}_\eta(x) - \mathbb{E}\hat{f}_\eta(x)\right|}_{\xi_\eta} + \overbrace{\left|\mathbb{E}\hat{f}_\eta(x) - \mathbb{E}\hat{f}_{h,\eta}(x)\right|}^{\text{Différence des biais}}$$

$$+ \underbrace{\left|\mathbb{E}\hat{f}_{h,\eta}(x) - \hat{f}_{h,\eta}(x)\right|}_{\xi_{h,\eta}}$$

Donc,

$$B_h = \max_{\eta \in H} B_h(\eta)$$

$$\leq \max_{\eta \in H} \left\{ \xi_\eta - X_1(\eta) \right\}_+ \qquad \text{où } X = X_1 + X_2$$

$$+ \max_{\eta \in H} \left\{ \xi_{h,\eta} - X_2(\eta) \right\}_+$$

$$+ \underbrace{\max_{\eta \in H} \left| \mathbb{E}\hat{f}_\eta(x) - \mathbb{E}\hat{f}_{h,\eta}(x) \right|}_{\leq \max_{\eta \in H} \|K\|_1 \sup_{x-\eta \leq u \leq x+\eta} |K_h * f(u) - f(u)|}$$

Pour borner les deux premiers termes en $\mathbb{E}(\dots)^q$, on va utiliser de la concentration.

Prop : Concentration de Berstein (celle issue de Bennett)

Z_i indépendantes avec
- $Z_i \leq b$
- $\sum \mathbb{E} Z_i^2 \leq V$

On note
$$\begin{cases} S = \sum_{i=1}^{n} Z_i - \mathbb{E} Z_i \\ \lambda(u) = \sqrt{2Vu} + \frac{2bu}{3} \end{cases}$$

Alors $\mathbb{P}(S \geq \lambda(u)) \leq e^{-u}$

$$Z_i = \frac{K_\eta(x - X_i)}{n}$$

- $Z_i \leq \frac{\|K\|_\infty}{n\eta}$ $\boxed{b = \frac{\|K\|_\infty}{n\eta}}$

- $\sum \mathbb{E} Z_i^2 \leq \frac{\|f\|_\infty}{n} \|K_\eta\|_2^2 = \frac{\|K\|_2^2 \|f\|_\infty}{n\eta}$

$$\boxed{V = \frac{\|f\|_\infty \|K\|_2^2}{n\eta}}$$

$$Z_i = \frac{K_h * K_\eta(x - X_i)}{n}$$

- $Z_i \leq \frac{\|K_h * K_\eta\|_\infty}{n}$

Hölder

$$\leq \frac{\|K_\eta\|_\infty \|K_h\|_1}{n} = \frac{\|K\|_\infty \|K\|_1}{n\eta}$$

$$\boxed{b = \frac{\|K\|_\infty}{n\eta} \|K\|_1}$$

- $\sum \mathbb{E} Z_i^2 \leq \frac{1}{n} \|f\|_\infty \|K_h * K_\eta\|_2^2$

Young

$$\leq \frac{1}{n} \|f\|_\infty \|K_\eta\|_2^2 \|K_h\|_1^2 = \frac{\|f\|_\infty \|K\|_2^2}{n\eta} \|K\|_1^2$$

$$\boxed{V = \frac{\|f\|_\infty \|K\|_2^2}{nh} \|K\|_1^2}$$

$$\leq C \sum_{\eta \in \mathcal{H}} \left(\frac{1}{m\eta}\right)^{q/2} \eta^{\gamma} \int_0^{\infty} \mathrm{Max}(\sqrt{u}^{-1}, \sqrt{u})\, e^{-u}\, du$$

$$\leq \frac{C}{m^{q/2}} \sum_{\eta \in \mathcal{H}} \eta^{\gamma - \frac{q}{2}}$$

On pose $\boxed{\beta = \gamma - \frac{q}{2}}$

$$= \frac{C}{m^{q/2}} \sum_{\eta \in \mathcal{H}} \eta^{\beta}$$

- le terme $\mathbb{E} \max_{\eta \in \mathcal{H}} \left\{ \xi_{h,\eta} - \chi_2(\eta) \right\}_+^q$ se traite de la même manière, avec la même borne à la clef.

<u>Conclusion</u> : La majoration étant vraie pour tout $h \in \mathcal{H}$ on obtient, pour n assez grand :

$$\boxed{\begin{aligned} \mathbb{E}\left|\hat{f}_{\hat{h}}(x) - f(x)\right|^q \leq C \min_{h \in \mathcal{H}} \Bigg[& \mathbb{E}\left|f_h(x) - f(x)\right|^q + \left(\frac{|\log h|}{n h}\right)^{q/2} \\ & + \sup_u \left|K_h * f(u) - f(u)\right|^q \\ & + \frac{1}{n^{q/2}} \sum_{\eta \in \mathcal{H}} \eta^{\beta} \Bigg] \end{aligned}}$$

<u>Rq</u>: Cette majoration est vraie pour : . f densité bornée

. $K \in L^2 \cap L^\infty$ à support compact

Retour aux classes de Hölder

Il faut trouver $\mathcal{H} = \mathcal{H}_n$ tel que

- $\mathcal{H}_n$ est assez riche pour imiter les bons fenêtrages $h_n(s) = n^{-\frac{1}{2s+1}}$
- $\mathcal{H}_n$ pas trop gros pour que $\frac{1}{n^{q/2}} \sum_{\eta \in \mathcal{H}} \eta^{-\beta} = o\left(n^{-\frac{sq}{2s+1}}\right)$

On prend $\mathcal{H}_n = \left\{ 2^{-k},\ k \in \left\{0, \dots, \left\lfloor \frac{\log n}{\log 2} \right\rfloor \right\}\right\}$:

- $n^{-\frac{q}{2}} \sum_{\eta \in \mathcal{H}} \eta^{-\beta} \leq C\, n^{-\frac{q}{2}} = o\left(n^{-\frac{sq}{2s+1}}\right)$
- $\forall s \in]0, s^*]$, $\quad 2^{-(k_s+1)} \leq h_n(s) \leq 2^{-k_s}$

  où $k_s = \left\lfloor \frac{1}{2s+1} \cdot \frac{\log n}{\log 2} \right\rfloor$

En travaillant un peu plus on obtient une borne Sup :

$$\overline{\lim_{n \to \infty}} \sup_{s \in]0, s^*]} \sup_{f \in \mathcal{C}^s(L)} \left(\frac{\log n}{n}\right)^{-\frac{sq}{2s+1}} \mathbb{E} \left|\hat{f}_{\hat{h}}(x) - f(x)\right|^q < \infty$$

Rq : La perte du "log" est inévitable : borne inf concordante.

Majoration de $\mathbb{E}\left\{\xi_h - X_1(h)\right\}_+^q$:

Soit $\gamma \geq 0$ à choisir plus tard

On pose $\lambda(u) = \sqrt{\dfrac{2\|f\|_\infty \|K\|_2^2\, u}{nh}} + \dfrac{2\|K\|_\infty\, u}{3\, nh}$ $\left(\leq C\, \dfrac{\max(\sqrt{u}, u)}{\sqrt{nh}}\right.$

car on prend $h \geq \dfrac{\log n}{n}$

$\Rightarrow \sqrt{nh} \to \infty$

Posons $\boxed{X_1(h) = \lambda(\gamma|\log h|)}$

Rq: Tout ce qui vérifie $X_1(h) \geq \lambda(\gamma \log h)$ convient !
(Et qui n'est pas trop gros !)

- $\mathbb{E} \max\limits_{h\in\mathcal{H}} \left\{\xi_h - X_1(h)\right\}_+^q \leq \sum\limits_{h\in\mathcal{H}} \mathbb{E}\left\{\xi_h - X_1(h)\right\}_+^q$

$$= \sum_{h\in\mathcal{H}} \int_0^\infty \mathbb{P}\left(\xi_h - X_1(h) \geq t^{\frac{1}{q}}\right) dt$$

on pose $\lambda(u) = t^{\frac{1}{q}}$, et on utilise le fait que $\begin{cases}\lambda(a+b) \leq \lambda(a) + \lambda(b)\\ \lambda'(a) \leq \dfrac{\lambda(a)}{a}\end{cases}$

$$= \sum_{h\in\mathcal{H}} \int_0^\infty \mathbb{P}\left(\xi_h \geq \lambda(u) + \lambda(\gamma|\log h|)\right) \lambda(u)^{q-1} \lambda'(u)\, du$$

$$\leq \sum_{h\in\mathcal{H}} \int_0^\infty \mathbb{P}\left(\xi_h \geq \lambda(u + \gamma|\log h|)\right) \frac{\lambda(u)^q}{u}\, du$$

$$\leq \sum_{h\in\mathcal{H}} \int_0^\infty e^{-(u+\gamma|\log h|)} \frac{\lambda(u)^q}{u}\, du$$