Université A/Mira de Béjaia 2020/2021
Faculté des Sciences Exactes
Département de Mathématiques
Master 2 PAS

TD - Estimation non paramétrique

Exercice 1

Soit $X_1, ..., X_n$ une suite de variables aléatoires indépendantes et identiquement distribuées de densité de probabilité inconnue f sur $\mathbb{R}$. On considère le noyau K d'Epanechnikov.

1. Montrer que le noyau K est une densité de probabilité sur $\mathbb{R}$.
2. Montrer que K est symétrique.
3. Utiliser le noyau K pour estimer f.
4. Calculer le Biais, la variance de l'estimateur. En déduire le MSE associé à f.
5. Calculer l'erreur quadratique moyenne intégrée asymptotique optimale de l'estimateur.

Exercice 2

Soit $X_1, ..., X_n$ une suite de variables aléatoires indépendantes et identiquement distribuées de densité de probabilité inconnue f sur $\mathbb{R}$. On se propose d'estimer la densité f par un noyau K symétrique .

- Montrer que si f est bornée et le noyau K est de carré intégrable (*ie.* $\int_{\infty}^{+\infty} K^2(\nu)d\nu < \infty$), alors
$$\mathbb{V}ar\left(\hat{f}_h(x)\right) \leq \frac{\sup_x \mid f(x) \mid}{nh} \int_{-\infty}^{\infty} K^2(\nu)d\nu$$
et
$$\int_{\mathbb{R}} \mathbb{V}ar\left(\hat{f}_h(x)\right) dx \leq \frac{1}{nh} \int_{-\infty}^{\infty} K^2(\nu)d\nu.$$
- Supposons f de classe C^2 et telle que f'' soit bornée. Montrer que :
$$\mid \mathbb{B}iais(\hat{f}_h) \mid \leq \frac{h^2}{2} \sup \mid f''(x) \mid \int_{-\infty}^{+\infty} \nu^2 \mid K(\nu) \mid d\nu.$$
- Donner une majoration du risque quadratique ponctuel MSE.

Exercice 3

Soient T un intervalle de $\mathbb{R}$, et deux réels $\beta, L > 0$. La classe de Hölder $\Sigma(\beta, L)$ sur T est définie comme l'ensemble des fonctions $g : T \to \mathbb{R}$ telles que g est $l = \lfloor \beta \rfloor$ fois dérivale et telle que $\mid g^l(x) - g^l(y) \mid \leq \mid x - y \mid^{\beta - l}, \forall x, y \in T$. Soit f une densité et K un noyau tels que :f est bornée et dans une classe de Hölder $\Sigma(\beta, L)$ sur $\mathbb{R}$, K est un noyau d'ordre $l = \lfloor \beta \rfloor$ de carré intégrable et tel que $\int_{-\infty}^{+\infty} \mid \mu \mid^{\beta} \mid K(\mu) \mid d\mu < \infty$. Montrer que

1. $\mathbb{V}ar\left(\hat{f}_h(x)\right) \leq \frac{c_1}{nh}, c_1 > 0$.
2. $\mid Biais(f_h)(x) \mid \leq \sqrt{c_2} h^{\beta}$ avec $\sqrt{c_2} = \frac{1}{l!} \int_{-\infty}^{+\infty} \mid \mu \mid^{\beta} \mid K(\mu) \mid d\mu$.
3. $MSE\left(\hat{f}_h(x)\right) \leq \frac{c_1}{nh} + c_2 h^{2\beta}$.

4. Pour $h_{opt} = bn^{-\frac{1}{2\beta+1}}$, avec $b > 0$, il existe $C > 0$ tel que

$$MSE\left(\hat{f}_h(x)\right) \leq Cn^{-\frac{2\beta}{2\beta+1}}.$$

<u>**Rappel**</u> : On dit que $K : \mathbb{R} \to \mathbb{R}$ est un noyau d'ordre l si les fonctions

$$\begin{cases} \mathbb{R} \to \mathbb{R} \\ u \to u^j K(u) \end{cases}$$

sont intégrables pour $j = 0, 1, ..., l$ et vérfient $\int_{-\infty}^{+\infty} K(u)du = 1$, ainsi que

$$\int_{-\infty}^{+\infty} u^j K(u)du = 0, j =, 1, ..., l.$$

Exercice 4

Soit $X_1, X_2, ..., X_n$ une suite de variables aléatoires indépendantes et identiquement distribuées de densité de probabilité f inconnue. On se propose d'estimer f par un noyau gaussien. Supposons que $K \sim \mathcal{N}(0, 1)$

1. Donner la forme de l'estimateur .
2. Calculer la variance et le biais de l'estimateur .
3. Calculer l'erreur quadratique moyenne de l'estimateur.
4. Calculer l'erreur quadratique moyenne intégrée MISE de l'estimateur.
5. Donner la forme de l'estimateur du paramètre de lissage par la méthode de rul of thumb et la méthode de plugi-itéré.

Exercice 5

Soit $X_1, X_2, \ldots, X_n$ une suite de variables aléatoires indépendantes et identiquement distribuées de densité de probabilité f inconnue. Soit $\hat{f}_n$ l'estimateur de f par la méthode des fonctions orthogonales.
La base est donnée par :

$$\left\{e_0(x) = 1, e_k(x) = \sqrt{2}\cos(k\pi x)\right\}.$$

1. Montrer que la base est orthogonale dans $[0, 1]$.
2. Donner la forme de l'estimateur $\hat{f}_n$ de f.
3. Montrer que $\hat{f}_n$ est borné.
4. Donner les propriétés statistiques et asymptotiques des estimateurs des coefficients de Fourier de f.
5. Calculer le paramètre de lissage par la méthode de Kronmal-Tarter.