



République Algérienne Démocratique et Populaire  
Ministères de l'Enseignement Supérieur  
et de la Recherche Scientifique

Centre Universitaire Salhi Ahmed De Naama  
Faculté des Sciences & Technologies  
Département De Mathématique Et Informatique



## *Mémoire de Master*

**Domaine :** Mathématiques et Informatique

**Filière :** Mathématiques

**Option :** Probabilité, Statistiques et Application

### **Thème**

---

*Estimation non paramétrique par la méthode De noyau de la fonction de densité conditionnelle avec une variable aléatoire fonctionnelle sur Des données indépendantes identiquement distribuées*

---

**Présenté par :**

HARKATI MAWLOUD

**Sous la direction de :**

DR DAOUDI HAMZA

**Membres du jury :**

M<sup>r</sup> DR LAALA Zineb

M<sup>r</sup> DR BELGUERNA Abderrahmane

M<sup>r</sup> DR DAOUDI HAMZA

MCA C.U. Salhi Ahmed - Naama-

MCA C.U. Salhi Ahmed -Naama-

MCA Univ Iben Kheldoun -Tiaret -

**Présidente.**

**Examineur.**

**Encadreur.**

# Remerciements



*Je voudrais d'abord et avant tout Remercier mon encadreur  
Docteur DAOUDI Hamza.*

*Pour l'aide qu'il a fournie et sa patience et la précision de ses  
indications.*

*Je tiens aussi à remercier tous mes professeurs de département  
de mathématique de centre universitaire Salhi Ahmed de  
NAAMA.*

*Pour aider et de ses efforts depuis de nombreuse années. Nous  
avons beaucoup profité de lui. Tous les remerciements et sa  
gratitude à lui.*

*J'aimerais présenter mes remerciements à tous ceux qui m'ont  
soutenu de près ou de loin avec un peu ou beaucoup.*

*Merci à vous tous.*

# Dédicace



*Je dédie ce travail :*

*Dieu soit loué, qui nous a appris ce que nous ne savons pas, il est le signe de l'invisible.*

*A l'aime de mon cher père HARKATI cheikh.  
Mon père, qui a planté en moi l'amour de la connaissance et de l'apprentissage, en particulier la sanctification et la perfection du travail.*

*A ma mère mon amour et la lumière de ma vie.*

*A mon frère Abderrahmane. A mes sœurs Amina et Aicha.*

*A tous ma grande famille Harkati-Talbi.*

*A tous mes professeurs.*

*A mes amis par leur nom, surtout mon ami  
Brahimi Nouredine.*

*A tous mes collègues du domain de l'éducation.*

# Table des matières

|  |           |
|--|-----------|
| <b>Introduction</b>  | <b>8</b>  |
| <b>1 Estimation fonctionnelle</b>                                    | <b>11</b> |
| 1.1 Les variables aléatoires fonctionnelles . . . . .                | 11        |
| 1.2 Ensembles de données fonctionnels . . . . .                      | 12        |
| 1.3 Estimation paramétrique et non paramétrique . . . . .            | 12        |
| 1.4 Estimation non-paramétrique d'une densité . . . . .              | 14        |
| 1.4.1 L'estimateur de la fonction de densité . . . . .               | 14        |
| <b>2 La Méthode de noyau</b>   | <b>18</b> |
| 2.1 Introduction . . . . .   | 18        |
| 2.2 Définition . . . . .   | 19        |
| 2.3 Présentation générale de la méthode du noyau . . . . .           | 20        |
| 2.3.1 Noyaux usuels . . . . .  | 20        |
| 2.3.2 Type de noyau . . . . .  | 21        |
| 2.3.3 Utilisation du noyau pour les données fonctionnelles . . . . . | 22        |
| 2.4 Estimateur à noyau . . . . .                                     | 23        |
| 2.4.1 Définition de la méthode de Parzen-Rozenblatt . . . . .        | 23        |
| 2.4.2 Estimateur à noyau . . . . .                                   | 23        |
| 2.4.3 Les propriétés de l'estimateur à noyau . . . . .               | 25        |
| 2.4.4 Les conditions sur le noyau $K$ . . . . .                      | 26        |
| 2.4.5 Choix de paramètre de lissage $h$ . . . . .                    | 26        |

## TABLE DES MATIÈRES

---

|          |   |           |
|----------|---|-----------|
| <b>3</b> | <b>La méthode de noyau de la fonction de densité conditionnelle</b>         | <b>27</b> |
| 3.1      | Présentation des modèles . . . . .  | 27        |
| 3.1.1    | Estimateur simple . . . . .   | 27        |
| 3.1.2    | Quelques résultats théoriques de base . . . . .                             | 30        |
| 3.2      | Estimateur à noyau de la densité conditionnelle . . . . .                   | 31        |
| 3.2.1    | Estimateur à noyau . . . . .  | 31        |
| 3.2.2    | Etude du biais et de la variance . . . . .                                  | 31        |
| 3.3      | Convergence presque complète . . . . .                                      | 39        |
| 3.3.1    | Inégalité de Bernshtein Fréchet . . . . .                                   | 39        |
| 3.3.2    | Convergence presque complète . . . . .                                      | 42        |
| 3.3.3    | La vitesse de Convergence presque complète . . . . .                        | 42        |
| <b>4</b> | <b>Application et exploration des observations sur des données simulées</b> | <b>45</b> |
| 4.1      | Simulée les données fonctionnelles . . . . .                                | 46        |
| 4.1.1    | Programme sous R d'un courbe aléatoire simulée . . . . .                    | 46        |
| 4.1.2    | Résultat de simulation d'un variable aléatoire fonctionnelle . . . . .      | 47        |
| 4.1.3    | Programme sous R de plusieurs courbes aléatoires simulées . . . . .         | 48        |
| 4.1.4    | Résultat de plusieurs variables aléatoires fonctionnelles . . . . .         | 49        |
| 4.2      | Simulation les variables de réponse réels . . . . .                         | 50        |
| 4.2.1    | Algorithme 1 . . . . .  | 50        |
| 4.2.2    | Programme sous R . . . . .  | 50        |
| 4.3      | Simulation l'estimateur de densité conditionnelle . . . . .                 | 51        |
| 4.3.1    | Algorithme 2 . . . . .  | 52        |
| 4.3.2    | Programme sous R . . . . .  | 52        |
| 4.4      | Comparation . . . . .   | 53        |
| 4.4.1    | Programme sous R . . . . .  | 54        |
| 4.5      | Résultats des simulations . . . . .   | 55        |
| 4.5.1    | Simulation de l'estimateur de densité conditionnelle . . . . .              | 55        |

|                      |  |           |
|----------------------|--|-----------|
| 4.5.2                | Simulation de la fonction $p$ . . . . .                | 56        |
| 4.5.3                | La fonction $p$ et l'estimateur $f$ au même plan . . . | 57        |
| 4.5.4                | Résultat final . . . . .                               | 57        |
| <b>Conclusion</b>    |  | <b>58</b> |
| <b>Bibliographie</b> |  | <b>59</b> |
| <b>Resumé</b>        |  | <b>62</b> |

# Table des figures

|     |  |    |
|-----|--|----|
| 1   | Les données fonctionnelles . . . . .   | 10 |
| 1.1 | Allures des noyaux : Triangulaire, Biweight, Gaussien et Epanechnikov . . . . .          | 17 |
| 4.1 | Un courbe aléatoire simulée . . . . .  | 47 |
| 4.2 | Plusieurs courbes aléatoires simulées . . . . .  | 49 |
| 4.3 | Comportement graphique de l'estimateur de densité conditionnelle $\hat{f}_Y^X$ . . . . . | 55 |
| 4.4 | Courbe de la fonction $p$ . . . . .  | 56 |
| 4.5 | Courbe de la fonction $p$ et l'estimateur de $f$ . . . . .                               | 57 |

## Notations et Préliminaires

- $X$  : Variable aléatoire.
- $E(X)$  : Espérance de la variable aléatoire  $X$ .
- $V(X)$  : La variance de la variable aléatoire  $X$ .
- p.co : Convergence presque complète.
- i.i.d : Indépendantes et identiquement distribuées.
- MSE : Erreur quadratique moyenne (Mean square Error).
- MISE : Erreur quadratique moyenne intégrée  
(Mean Integrated square Error).
- CV : Cross Validation (validation croisée)
- $O_{p.co}$  : vitesse de convergence presque complète
- $h$  : le paramètre de lissage
- $K$  : une fonction asymétrique du noyau
- $r$  : opérateur de régression non linéaire

# Introduction

L'objet principal de la statistique est de faire, à partir d'observations d'un phénomène aléatoire, une inférence au sujet de la loi générant ces observations en vue d'analyser le phénomène ou de prévoir un événement futur.

Pour réduire la complexité du phénomène étudié, nous pouvons utiliser deux approches statistiques : non-paramétrique et paramétrique. Dans la première approche, un problème récurrent en statistique est celui de l'estimation d'une densité  $f$  ou d'une fonction de répartition  $F$  à partir d'un échantillon de variables aléatoires réelles  $X_1, X_2, \dots, X_n$  indépendantes et de même loi inconnue.

Les fonctions  $f$  et  $F$ , tout comme la fonction caractéristique, décrivent complètement la loi de probabilité des observations et en connaître une estimation convenable permet de résoudre nombre de problèmes statistiques.

Cette estimation tient donc naturellement une place importante dans l'étude de nombreux phénomènes de nature aléatoire.

Pour estimer n'importe quel paramètre fonctionnel il suffit d'estimer la fonction de répartition  $F$  par la fonction de répartition empirique  $F_n$ , et par conséquent l'estimateur de  $\theta_n$  est  $T(F_n)$  où  $T$  est la fonctionnelle statistique. La fonction de répartition empirique donc joue un rôle fondamental dans l'estimation fonctionnelle plus précisément dans l'estimation de la densité  $f$ , pour qu'on puisse tirer plus d'information sur la loi parente.

La connaissance de l'estimateur de  $F$  et  $f$  mènent à résoudre un autre problème fondamental de la statistique non paramétrique, c'est le problème de la régression.

Les estimateurs non-paramétriques classiques ont été introduit par Roseblatt (1956) pour estimer la fonction de densité, a été reprise simultanément par Waston (1964) et Nadaraja (1964) pour estimer une fonction de régression.

Le comportement asymptotique de ces estimateurs a été étudié par de nombreux auteurs tel que Tsybakov (2004).

Ainsi, le but de ce travail est de définir les estimateurs à noyau associé et d'établir les propriétés relatives.

La base et l'objectif de cet mémoire est l'estimation non paramétrique par la méthode de noyau de la fonction de densité conditionnelle avec une variable aléatoire fonctionnelle sur des données indépendantes identiquement distribuées(i.i.d),après on va étudier la convergence presque complète et la vitesse de convergence.

Dans le premier chapitre, nous donnerons quelques notions élémentaires, définitions et exemples dans l'estimation paramétrique et non paramétrique.

Dans le deuxième chapitre, nous donnerons quelques notions élémentaires, définitions de la méthode de noyau et leur utilisation pour les données fonctionnelles, et l'estimation par cette méthode.

Dans le troisième chapitre, nous donnerons présentation des modèles et l'estimateur à noyau de la densité conditionnelle.

Dans le dernier chapitre le travail des simulations sont faites pour illustrer les résultats théoriques établis sur la densité de probabilité.

Les applications et les commentaires, où nous avons confirmé que les variables aléatoires fonctionnelle sont des courbes (sous logiciel R), nous illustrons la convergence presque complète en effectuant des simulations sur l'estimateur de densité conditionnelle par la méthode de noyau pour les variables fonctionnelles (qui présenté par Ferraty), et sur autre densité, ensuite nous examinons les résultats.

Nous allons comparer les résultats, cet application montre que la convergence presque complète de l'estimateur de densité conditionnelle est confirmée.

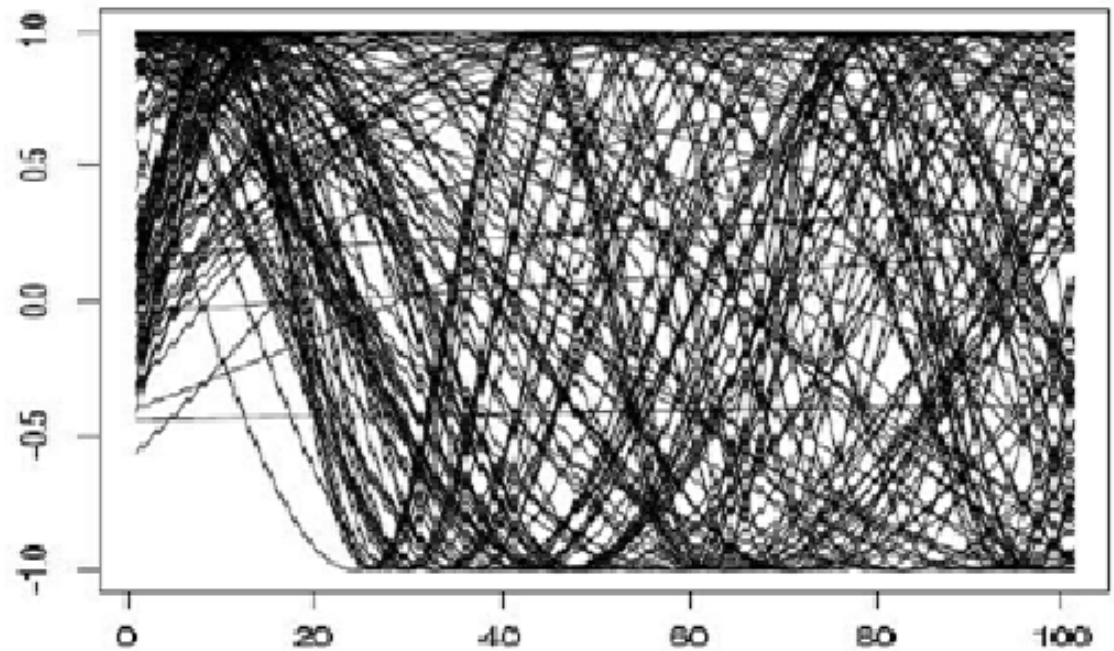


FIGURE 1 – Les données fonctionnelles

# Chapitre 1

## Estimation fonctionnelle

Dans ce chapitre, nous donnerons quelques notions élémentaires, définitions et exemples dans l'estimation paramétrique et non paramétrique.

### 1.1 Les variables aléatoires fonctionnelles

Une variable aléatoire  $\mathcal{X}$  est appelée variable fonctionnelle (v.f.) si elle prend des valeurs dans un espace de dimension infini (ou espace fonctionnel).

Une observation  $\chi$  de  $\mathcal{X}$  s'appelle une donnée fonctionnelle.

Notez que, lorsque  $\mathcal{X}$  (resp.  $\chi$ ) désigne une courbe aléatoire (ou son observation), nous impliquons implicitement l'identification suivante  $\mathcal{X} = \{\mathcal{X}(t); t \in T\}$  (resp.  $\chi = (\chi(t)); t \in T$ ).

Dans cette situation, la caractéristique fonctionnelle provient directement des observations la situation où la variable est une courbe est associée à un ensemble unidimension  $T \subset \mathbb{R}$ .

Ici, il est important de remarquer que la notion de variable fonctionnelle couvre une plus grande surface que l'analyse des courbes.

## 1.2 Ensembles de données fonctionnels

Un ensemble de données fonctionnelles  $\chi_1, \dots, \chi_n$  est l'observation de  $n$  variables fonctionnelles  $\mathcal{X}_1, \dots, \mathcal{X}_n$  est identique à  $\mathcal{X}$ .

Cette définition couvre de nombreuses situations, le plus populaire étant les ensembles de données courbes.

Ici, nous supposons que nous disposons d'un échantillon de données fonctionnelles.

## 1.3 Estimation paramétrique et non paramétrique

**Définition 1.3.1. Estimation paramétrique :** Si l'on sait à priori que  $h$  appartient à une famille paramétrée  $\{h(x, \theta), \theta \in \Theta\}$  où  $\Theta \subset \mathbb{R}^s$  et  $h(.,.)$  est une fonction connue, on parle alors d'estimation paramétrique, car estimer  $h$  revient à estimer le paramètre fini dimensionnel  $\theta$ .

**Définition 1.3.2. Estimation non paramétrique :** Si l'on sait seulement que  $h$  appartient à  $\mathbb{P}$  ensemble des lois de probabilités qui est un espace de dimension infinie, alors on dit que l'on fait de l'estimation non paramétrique ou de l'estimation fonctionnelle

Dans ce qui suit, on suppose que l'on a observé un échantillon  $X_1, X_2, \dots, X_n$  à valeurs dans  $\mathbb{R}^s$  muni de sa tribu borélienne  $\mathcal{B}$ .

De plus, on suppose que les  $\{X_i, i = 1, \dots, n\}$  sont indépendantes et identiquement distribuées (i.i.d)  $\mu \in P_0$  une famille de loi sur  $(\mathbb{R}^s; \mathcal{B})$ .

1. **La densité de probabilité :** Si  $P_0$  est une famille de loi dominée par une loi  $\lambda$ , donc elle admet (théorème de Radon-Nykadim) une densité  $f = \frac{d\mu}{d\lambda}$  c'est un paramètre dans  $L^1$ .

Si  $\frac{d\mu}{d\lambda}$  admet une version bornée (respectivement continue et bornée) alors on peut la considérer comme un paramètre dans  $L^2$  (respectivement dans  $C_b(\mathbb{R}^s)$ ).

Enfin, si  $f_\mu$  est différentiable, on définit de nouveaux paramètres fonctionnels : les dérivées partielles de  $f_\mu$  :

2. La fonction de répartition  $C'$  est la fonction définie par

$$F_\mu(x_1, \dots, x_s) = \mu \left( \prod_{i=1}^s ]-\infty; x_i] \right), (x_1, \dots, x_s) \in \mathbb{R}^s.$$

3. La fonction des quantiles Pour  $s = 1$ , la fonction quantile d'ordre  $p$  définie par

$$F_\mu^{-1}(p) = Q(p) = \inf \{ t \in \mathbb{R}; F_\mu(t) \geq p \} \quad 0 < p < 1$$

$F_\mu^{-1}$  est un paramètre à valeur dans l'espace de fonctions réelles définies sur  $]0; 1[$  monotones non décroissantes et continues à gauche.

4. La fonction caractéristique Elle est définie par

$$\hat{\mu}(t) = E_\mu[\exp\{i\langle t, x \rangle\}] \quad \text{où } t, x \in \mathbb{R}^s.$$

$\hat{\mu}$  est un paramètre dans  $C_b(\mathbb{R}^s)$ .

5. Le paramètre de régression Supposons que l'on observe un échantillon  $\{(X_i, Y_i), i = 1, \dots, n\}$  d'un couple  $(X, Y)$  à valeurs dans  $\mathbb{R}^{s_1} \times \mathbb{R}^{s_2}$  est soit  $\mu_Y^x, x \in \mathbb{R}^{s_1}$  une famille de versions des lois conditionnelles de  $Y$  sachant  $X = x$  : Toute fonction de la forme  $r : x \rightarrow r(\mu_Y^x)$

est un paramètre de régression. Les plus usuels sont :

- (a) L'espérance conditionnelle (qui est la fonction de régression),
- (b) La densité conditionnelle,
- (c) Le mode conditionnel,
- (d) La fonction de répartition conditionnelle,
- (e) Le quantile conditionnel.

## 1.4 Estimation non-paramétrique d'une densité

### 1.4.1 L'estimateur de la fonction de densité

Supposons que nous observons  $n$  variables aléatoires indépendantes et identiquement distribués  $X_1, \dots, X_n$  de densité de probabilité par rapport à la mesure de Lebesgue une fonction inconnue  $f$  de  $\mathbb{R}$  dans  $[0, +\infty[$ .

L'objectif de notre étude est la construction d'un estimateur de  $f$ , c'est-à-dire une fonction  $\hat{f}_n(x) = f_n(x, X_1, \dots, X_n)$  mesurable par rapport à la tribu engendrée par  $(X_1, \dots, X_n)$ .

Notons  $F(x) = P(X_1 \leq x)$  la fonction de répartition de la loi de  $X_1$  et considérons la fonction de répartition empirique

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}; \quad \forall x \in \mathbb{R}$$

La loi forte des grands nombres permet d'affirmer que  $F_n$  est un estimateur de  $F$ .

Il est même possible d'obtenir des intervalles de confiance et de tester l'adéquation des données à différentes lois.

Néanmoins, il n'est pas évident d'utiliser  $\hat{F}_n$  pour estimer  $f$ .

Une des premières idées intuitives est de considérer pour  $h > 0$  petit

$$\hat{f}_n(x) = \frac{\hat{F}_n(x+h) - \hat{F}_n(x-h)}{2h} = \frac{1}{2nh} \sum_{i=1}^n \mathbf{1}_{\{-h \leq X_i - x \leq h\}}$$

L'idée la plus naturelle est d'estimer  $f$  en un point  $x$  et de voir ce qui se passe au voisinage de  $x$ .

Une des premières idées intuitives est de considérer pour  $h > 0$  petit

$$\begin{aligned} f(x) &= \lim_{h \rightarrow 0} \frac{F(x+h) - F(x-h)}{2h} \\ &= \lim_{h \rightarrow 0} \frac{F(x+h) - f(x)}{2h} + \frac{F(x) - F(x-h)}{2h}. \end{aligned}$$

En remplaçant alors  $F$  par  $F_n$ , on obtient

$$\hat{f}_n(x) = \frac{F_n(x+h) - F_n(x-h)}{2h} \quad (1.4.1)$$

qui peut aussi s'écrire sous la forme

$$\begin{aligned} \hat{f}_n(x) &= \frac{1}{2h} \frac{1}{n} \sum_{i=1}^n \{ \mathbf{1}_{\{X_i \leq x+h\}} - \mathbf{1}_{\{X_i \leq x-h\}} \} \\ &= \frac{1}{2nh} \sum_{i=1}^n \mathbf{1}_{\{x-h \leq X_i \leq x+h\}} \\ &= \frac{1}{2nh} \sum_{i=1}^n \mathbf{1}_{\{-1 \leq \frac{X_i - x}{h} \leq 1\}} \\ &=: \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right) \end{aligned}$$

Cet estimateur, appelé estimateur de Rosenblatt (1956), c'est le premier exemple d'estimateur à noyau construit à l'aide d'une fonction appelée noyau :  $K(t) := \frac{1}{2}\mathbf{1}_{\{-1 \leq t \leq 1\}}$ . Définissons maintenant plus généralement la notion d'estimateur à noyau.

**Définition 1.4.1.** Soit  $K : \mathbb{R} \longrightarrow \mathbb{R}$  une fonction intégrable, positive et telle que  $\int K(u)du = 1$ ,  $K$  est appelée noyau.

Soit  $h := h_n > 0$  un paramètre de lissage (fenêtre) qui dépend de la taille de l'échantillon  $n$ .

Pour toute  $n \in \mathbb{N}^*$ , l'estimateur à noyau de la densité de la v.a.  $\mathbf{X}$  au point  $x$ , noté  $\hat{f}_{n,X}(x)$  est donné par :

$$\hat{f}_{n,X}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right). \quad (1.4.2)$$

### 1.4.1.1 Exemples de noyaux classiques

Voici quelques exemples de noyaux classiques :

- Noyau Triangulaire :  $K(t) = (1 - |t|)\mathbf{1}_{\{|t| \leq 1\}}$ ,
- Noyau Biweight :  $K(t) = \frac{15}{16} (1 - t^2)^2 \mathbf{1}_{\{|t| \leq 1\}}$ ,
- Noyau Gaussien :  $K(t) = \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\}$ ,  $t \in \mathbb{R}$ ,
- Noyau d'Epanechnikov :  $K(t) = \frac{3}{4} (1 - t^2) \mathbf{1}_{\{|t| \leq 1\}}$ ,

La figure (1.1) si-après présente l'allure des quatre noyaux cités.

## CHAPITRE 1. ESTIMATION FONCTIONNELLE

---

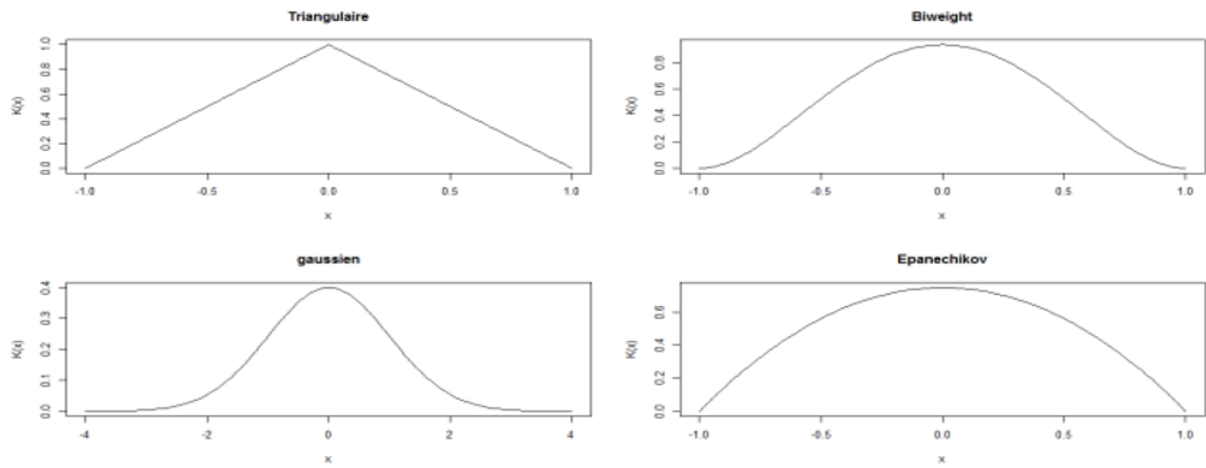


FIGURE 1.1 – Allures des noyaux : Triangulaire, Biweight, Gaussien et Epanechnikov

# Chapitre 2

## La Méthode de noyau

Dans ce chapitre, nous donnerons quelques notions élémentaires, définitions de la méthode de noyau et leur utilisation pour les données fonctionnelles, et l'estimation par cette méthode.

### 2.1 Introduction

En raison de leur simplicité, les praticiens ont utilisé les noyaux classiques (symétriques), introduits par Rosenblatt [1956] et Parzen [1962] pour l'estimation de la fonction densité de probabilité à support non borné.

Citons le noyau gaussien, uniforme, biweight ainsi que le noyau optimal d'Epanechnikov [1969].

L'estimateur à noyau classique d'une densité de probabilité  $f$ , étant donné un  $n$ -échantillon  $X_1, \dots, X_n$  i.i.d. issu d'une v.a.  $X$  de fonction densité inconnue  $f$ , est de la forme

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (2.1.1)$$

où  $K$  est la fonction noyau vérifiant généralement les conditions suivantes :

$$\int_{\mathcal{S}} K(u) du = \mathbf{1}, \int_{\mathcal{S}} uK(u) du = \mathbf{0} \text{ et } \int_{\mathcal{S}} u^2K(u) du = \sigma_K^2 < \infty$$

et  $\mathcal{S}$  le support de la fonction noyau  $K$ .

## 2.2 Définition

Les définitions suivantes présentent les notions du noyau associé et de l'estimateur à noyau associé pour la fonction de densité  $f$  inconnue sur le support  $T$ .

**Définition 2.2.1.** : Soit  $x \in \mathbb{T}$  et  $h > 0$  avec  $\mathbb{T} \subseteq \mathbb{R}$  le support de la densité  $f$  à estimer. On appelle noyau associé  $K_{x,h}$  toute densité paramétrique de probabilité de support  $\mathcal{S}_{x,h}$ , vérifiant les trois conditions suivantes :

$$x \in \mathcal{S}_{x,h} \tag{2.2.1}$$

$$E \{ \mathcal{K}_{x,h} \} = x + A(x, h) \tag{2.2.2}$$

$$V \{ \mathcal{K}_{x,h} \} = B(x, h) \tag{2.2.3}$$

où  $\mathcal{K}_{x,h}$  est une variable aléatoire de densité  $K_{x,h}$  et les termes  $A(x, h)$  et  $B(x, h)$  convergent vers 0 quand  $h \rightarrow 0$ .

**Définition 2.2.2.** Soit  $X_1, \dots, X_n$  un  $n$ -échantillon aléatoire indépendant et identiquement distribué issu d'une variable aléatoire  $X$  de fonction de densité inconnue  $f$  sur  $\mathbb{T} \subseteq \mathbb{R}$ . L'estimateur à noyau associé  $\hat{f}_h$  de  $f$  utilisant  $K_{x,h}$  est défini par

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_{x,h}(X_i), \quad x \in \mathbb{T} \tag{2.2.4}$$

où  $K_{x,h}$  est le noyau associé continu de cible  $x$  et de fenêtre (paramètre de lissage)  $h$  sur  $\mathcal{S}_{x,h}$ . La cible  $x$  et le paramètre  $h$  interviennent dans la définition intrinsèque du noyau associé  $K_{x,h}$ .

## 2.3 Présentation générale de la méthode du noyau

Un noyau est une fonction de pondération utilisée dans les techniques d'estimation non paramétrique.

Les noyaux interviennent dans l'estimation à noyau pour estimer la densité de probabilité d'une variable aléatoire, elle se base sur un échantillon d'une population statistique et permet d'estimer la densité en tout point du support.

En ce sens, cette méthode généralise astucieusement la méthode d'estimation par un histogramme. Cette méthode est aussi appelée méthode de Parzen-Rozenblatt.

### 2.3.1 Noyaux usuels

Il existe plusieurs types de noyaux, citons les plus utilisés.

$$K_1(u) = \begin{cases} \frac{1}{2}, & \text{si } |u| \leq 1 \\ 0, & \text{si } |u| > 1 \end{cases} \quad (\text{noyau rectangulaire})$$

$$K_2(u) = \begin{cases} (1 - |u|), & \text{si } |u| \leq 1 \\ 0, & \text{si } |u| > 1. \end{cases} \quad (\text{noyau triangulaire})$$

$$K_3(u) = \begin{cases} \frac{3}{4}(1 - u^2), & \text{si } |u| \leq 1; \\ 0, & \text{si } |u| > 1 \end{cases} \quad (\text{noyau parabolique ou d'Epanechnikov});$$

$$K_4(u) = \begin{cases} \frac{15}{16}(1 - u^2)^2 & \text{si } |u| \leq 1 \\ 0, & \text{si } |u| > 1 \end{cases} \quad (\text{noyau quadratique})$$

$$K_5(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \quad (\text{noyau gaussien})$$

$$K_6(u) = \frac{1}{2} \exp(-|u|/\sqrt{2}) \sin(|u|/\sqrt{2} + \pi/4)$$

(noyau de Silverman).

### 2.3.2 Type de noyau

Nous allons considérer deux sortes de noyaux : noyaux de type I et noyau de type II.

La famille du noyau de type I contient les noyaux usuels discontinus, tandis que la seconde famille contient les noyaux standards continus.

**Définition 2.3.1.** Une fonction  $K : \mathbb{R} \rightarrow \mathbb{R}^*$  telle que  $\int K = 1$  est dite noyau de type I si il existe deux constantes réelles  $0 < C_1 < C_2 < \infty$  telle que :

$$C_1 \mathbf{1}_{[0,1]} \leq K \leq C_2 \mathbf{1}_{[0,1]}$$

**Définition 2.3.2.** Une fonction  $K : \mathbb{R} \rightarrow \mathbb{R}^*$  telle que  $\int K = 1$  est dite noyau de type II si son support est  $[0, 1]$  et si sa dérivé  $K'$  existe sur l'intervall  $[0, 1]$  et cette dérivé vérifie la condition suivante : si il existe deux constante réelles  $C_1$  et  $C_2$  telle que :

$$-\infty < C_2 < C_1 < 0$$

alors

$$C_2 \leq K' \leq C_1$$

Afin de simplifier notre objectif, pour la pondération locale des variables aléatoires réelles nous allons définir le noyau suivant :

**Définition 2.3.3.** Une fonction  $K : \mathbb{R} \rightarrow \mathbb{R}^*$  telle que :  $\int K = 1$  sur un support compact  $[-1, 1]$  et  $\forall u \in (0, 1), K(u) > 0$  est appelé noyau de type  $0[1]$ .

### 2.3.3 Utilisation du noyau pour les données fonctionnelles

Les méthodes du noyau sont utilisées de manière intensive par la communauté des non-paramétrés car elles constituent un moyen utile de faire une pondération locale. Nous commençons par rappeler quelle est la pondération locale du noyau dans les cas réels et multivariés avant de l'étendre au contexte fonctionnel.

#### 2.3.3.1 Cas fonctionnelle

Soit  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$   $n$  v.a.f. évalué dans  $E$  et soit  $\chi$  un élément fixe de  $E$ .

une extension fonctionnelle des idées de pondération locale du noyau multivarié serait transforme le  $n$  v.a.f.  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$ , les  $n$  quantités

$$\frac{1}{V(h)} K \left( \frac{d(\chi, \mathcal{X}_i)}{h} \right),$$

Où  $d$  est semi-métrique sur  $E$ ,  $K$  est un noyau réel (asymétrique). Dans cette expression  $V(h)$  serait le volume de

$$B(\chi, h) = \{\chi' \in E, d(\chi, \chi') \leq h\}$$

Qui est la boule, par rapport à la topologie induite par  $d$ , centrée à  $\chi$  et au rayon  $h$ . Cependant, cette approche naïve demande de définir  $V(h)$ . il faut avoir une mesure sur  $E$ .

C'est la principale différence avec les cas réels et multivariés pour lesquels la mesure de Lebesgue est implicitement utilisée alors que dans l'espace fonctionnel  $E$ .

Par conséquent, afin de se libérer d'un choix d'une mesure particulière, nous construisons la normalisation en utilisant directement la distribution de probabilité de la v.a.f. les variables pondérées locales

du noyau fonctionnel sont définies par :

$$\Delta_i = \frac{K\left(\frac{d(\mathcal{X}, \mathcal{X}_i)}{h}\right)}{\mathbb{E}\left(K\left(\frac{d(\mathcal{X}, \mathcal{X}_i)}{h}\right)\right)} \quad (2.3.1)$$

Si nous retournons au cas multivarié, nous avons, pour un  $C$  constant selon  $K$  et sur la norme  $\|\cdot\|$  Utilisé dans  $\mathbf{R}_p$ ,

$$\mathbb{E}K(\|x - \mathbf{X}_i\|/h) \sim Cf(x)h^p \quad (2.3.2)$$

Tant que  $\mathbf{X}_i$  a une densité par rapport à la mesure de Lebesgue qui est continue et tel que  $f(x) > 0$  (ce type de résultat est connu dans la littérature comme théorème de type de Bochner.

Ainsi, il est clair maintenant que (2.3.1) est une extension de la pondération locale du noyau multivarié dans le cadre fonctionnel.

## 2.4 Estimateur à noyau

### 2.4.1 Définition de la méthode de Parzen-Rozenblatt

L'estimation par la méthode noyau (ou méthode de Parzen-Rozenblatt) est une méthode non paramétrique d'estimation de la densité d'une variable aléatoire.

Cette méthode permet d'obtenir une densité continue et constitue en ce sens une généralisation de la méthode de l'histogramme.

En effet, la fonction indicatrice utilisée pour l'histogramme est ici remplacée par une fonction continue (le noyau) et une somme de fonctions continues reste continue.

### 2.4.2 Estimateur à noyau

On va s'intéresser au modèle de régression, où

$$Y_i = f(x_i) + \varepsilon_i, i = 1, \dots, n \quad (2.4.1)$$

## CHAPITRE 2. LA MÉTHODE DE NOYAU

---

Ici,  $x$  est un v.a et les  $x_i$  sont  $n$ -échantillons de  $x$ , les  $\varepsilon_i$  sont i.i.d., et on cherche à estimer  $f$ . Noter aussi que les  $Y_i$  sont des variables aléatoires indépendantes mais pas identiquement distribuées.

On suppose que  $f$  est continue, telle que :

$$\forall t \in [0, 1] \quad F(t) = \int_t^1 f(x) dx.$$

On peut estimer la fonction de répartition  $F$  de manière empirique par  $\hat{F}$  donnée par :

$$\forall t \in [0, 1] \quad \hat{F}(t) = \frac{1}{n} \sum_{i=1}^n Y_i \mathbf{1}_{x_i \leq t},$$

puis par  $\hat{f}$  un taux d'accroissement

$$\hat{f}(x) = (\hat{F}(x+h) - \hat{F}(x-h))/2h.$$

On a alors

$$\hat{f}(x) = \frac{1}{2nh} \sum_{i=1}^n Y_i \mathbf{1}_{x-h < x_i \leq x+h}.$$

$\hat{f}(x)$  est encore la moyenne locale des  $Y_i$  mais sur une fenêtre "glissante" centrée sur  $x$

Si on pose  $K(x) = \frac{1}{2} \mathbf{1}_{-1 < x \leq 1}$  on a :

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n Y_i \frac{1}{h} K\left(\frac{x_i - x}{h}\right). \quad (2.4.2)$$

Remarquons que :

$$\sum_{i=1}^n \frac{1}{h} K\left(\frac{x_i - x}{h}\right) = \frac{1}{2} \sum_{i=1}^n \mathbf{1}_{x-h < x_i \leq x+h} \approx nh$$

et on peut choisir

$$\hat{f}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} \quad (2.4.3)$$

Le paramètre  $h > 0$  est appelé fenêtre, c'est un paramètre de lissage. Dans les deux cas, il s'agit d'un estimateur linéaire tel que

$$\hat{f}(x) = \sum_{i=1}^n W_{n_i,x} Y_i$$

avec

$$W_{n_i,x} = \frac{1}{nh} K\left(\frac{x_i - x}{h}\right)$$

si on choisit (2.5.1), ou

$$W_{n_i,x} = K\left(\frac{x_i - x}{h}\right) / \sum_{k=1}^n K\left(\frac{x_k - x}{h}\right)$$

si on choisit (2.2.2). Pour généraliser la "moyenne locale avec fenêtre glissante", on peut choisir d'autres types de noyaux, i.e. de fonctions  $K$ , que l'on choisit souvent positives, à support compact, paires, d'intégrale 1.

### 2.4.3 Les propriétés de l'estimateur à noyau

Si  $k$  est une densité de probabilité, alors  $\hat{f}$  est aussi une densité de probabilité.

$\hat{f}$  a les mêmes propriétés de continuité et de différentiabilité que  $k$  :

-Si  $k$  est continue, sera une fonction continue.

-Si  $k$  est différentiable,  $\hat{f}$  sera une fonction différentiable.

-Si  $k$  peut prendre des valeurs négatives, alors  $\hat{f}$  pourra aussi prendre des valeurs négatives.

### 2.4.4 Les conditions sur le noyau $K$

On suppose que  $K$  vérifie les 4 conditions suivantes :

1.  $\int_{\mathbb{R}} K = 1$
2.  $K$  est une fonction paire ou, plus généralement,  $\int_{\mathbb{R}} uK(u)du = 0$
3.  $\int_{\mathbb{R}} u^2K(u) < \infty$
4.  $\int_{\mathbb{R}} K(u^2) < \infty$ .

### 2.4.5 Choix de paramètre de lissage $h$

Le choix de  $h$  est une étape importante lors de l'estimation par noyau dans le sens où  $h$  optimal nécessite de connaître la densité que l'on cherche à estimer, ce qui n'est en pratique pas le cas.

Par ailleurs, si  $h$  est trop petit, le biais de l'estimateur devient petit devant sa variance et l'estimateur trop fluctant.

On obtient un phénomène de sous-lissage.

Dans le cas contraire, lorsque  $h$  est trop grand, le biais prend l'ascendant sur la variance et l'estimateur varie peu : on obtient un phénomène de sur-lissage. En pratique, on utilise souvent la méthode de la validation croisée pour choisir automatiquement  $h$ .

# Chapitre 3

## La méthode de noyau de la fonction de densité conditionnelle

Dans ce chapitre, nous donnerons présentation des modèles et l'estimateur à noyau de la densité conditionnelle. L'objectif de notre étude dans ce chapitre est la construction d'un estimateur de  $f$ , c'est-à-dire une fonction  $\hat{f}_n(x) = f(x, X_1, X_2, \dots, X_n)$  par la méthode du noyau, et de l'étude de la convergence presque complète, et la vitesse de convergence presque complète.

### 3.1 Présentation des modèles

#### 3.1.1 Estimateur simple

Rappelons que la densité de probabilité  $f$  est égale à la dérivée de la fonction de répartition  $F$  (si cette dérivée existe).

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

On peut donc écrire, quand  $h \rightarrow 0$

$$\begin{aligned} f(x) = F'(x) &= \frac{F(x+h) - F(x-h)}{2h} \\ &= \frac{P[x-h < X \leq x+h]}{2h} \end{aligned}$$

Un estimateur de  $f(x)$  est alors

$$\begin{aligned} f_n(x) &= \frac{F_n(x+h) - F_n(x-h)}{2h} \\ &= \frac{1}{2nh} \sum_{i=1}^n \mathbf{I}_{[X_i \in [x-h, x+h]]} (X_i) \\ &= \frac{1}{nh} \sum_{i=1}^n \frac{1}{2} \mathbf{I}_{\{-1 \leq \frac{x-X_i}{h} < 1\}} (X_i) \end{aligned}$$

Notons que cette estimateur peut encore s'écrire comme

$$f_n(x) = \frac{1}{nh} \sum_{i=1}^n \mathbf{W} \left( \frac{x - X_i}{h} \right)$$

Où

$$\mathbf{W}(y) = \begin{cases} 1/2 & \text{si } y \in [-1, 1] \\ 0 & \text{sinon} \end{cases}$$

Cet estimateur, appelé estimateur de Rosenblatt (1956), est le premier exemple d'estimateur à noyau construit à l'aide du noyau

$$\mathbf{W}(y) = \frac{1}{2} \mathbf{I}_{\{-1 \leq y < 1\}}$$

notion que nous allons étudier plus tard. Quelles sont les propriétés de l'estimateur simple  $f_n(x)$  ?

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

Remarquons que

$$f_n(x) = \frac{F_n(x+h) - F_n(x-h)}{2h}$$

avec  $F_n$  la fonction de répartition empirique.

Le paramètre de lissage  $h$  dépend de la taille de l'échantillon  $n$ , c'est-à-dire  $h = h_n$ . Nous savons que

$$n F_n(x) = \sum_{i=1}^n \mathbf{I}_{\{X_i \leq x\}} (X_i) \sim \text{Bin}(n, F(x))$$

et

$$\begin{aligned} 2nh_n f_n(x) &= n F_n(x+h_n) - n F_n(x-h_n) \sim \text{Bin}(n, F(x+h_n) - F(x-h_n)) \\ &\Rightarrow \mathbf{E}[2nh_n f_n(x)] = n [F(x+h_n) - F(x-h_n)] \\ &\Rightarrow \mathbf{E}[f_n(x)] = \frac{1}{2h_n} [F(x+h_n) - F(x-h_n)] \end{aligned}$$

**Pour la variance nous trouvons**

$$\begin{aligned} \text{Var}[2nh_n f_n(x)] &= n [F(x+h_n) - F(x-h_n)] [1 - F(x+h_n) - F(x-h_n)] \\ &\Rightarrow \text{Var}[f_n(x)] = \frac{1}{4nh_n^2} [F(x+h_n) - F(x-h_n)] [1 - F(x+h_n) - F(x-h_n)] \end{aligned}$$

Remarquons que, si  $n \rightarrow \infty$  et  $h_n \rightarrow 0$ , alors

$$\mathbf{E}[f_n(x)] = f(x)$$

et

$$nh_n \text{Var}[f_n(x)] \rightarrow \frac{1}{2} f(x)$$

**Remarque 3.1.1.** Quand  $nh_n \rightarrow \infty$ , l'expression de la variance devient

$$\text{Var}[f_n(x)] = \frac{1}{2nh_n} f(x)$$

Donc,

$$\text{Var}[f_n(x)] \rightarrow 0$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

L'erreur quadratique moyen de l'estimateur  $f_n(x)$  de  $f(x)$  est donné par :

$$\mathbf{E} [f_n(x) - f(x)]^2 = \mathbf{Var} [f_n(x)] + [\text{Biais } \{f_n(x)\}]^2$$

Donc, si  $h_n \rightarrow 0$  et  $nh_n \rightarrow \infty$  quand  $n \rightarrow \infty$ , on a que

$$\mathbf{E} [f_n(x) - f(x)]^2 \rightarrow 0$$

pour tout point  $x$ . L'estimateur simple  $f_n(x)$  est alors un estimateur consistant de  $f(x)$ .

### 3.1.2 Quelques résultats théoriques de base

Comme l'idée du noyau du poids local fonctionnel est le foyer de toutes les méthodes non paramétriques fonctionnelles qu'on va étudier, on utilise les deux résultats suivants :

**Lemme 3.1.2.** *Si  $\mathbf{K}$  est un noyau de type I, alors il existe deux constantes réelles non négatives  $\mathbf{C}$  et  $\mathbf{C}'$  telles que :*

$$\mathbf{C}\varphi_x(h) \leq \mathbf{E} \left( \mathbf{K} \left( \frac{d(x, \mathbf{X})}{h} \right) \right) \leq \mathbf{C}'\varphi_x(h) \quad (3.1.1)$$

**Lemme 3.1.3.** *Si  $\mathbf{K}$  est un noyau de type II, et si  $\varphi_x(h)$  satisfait*

$$\exists \mathbf{C} > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > \mathbf{C} \text{ffl} \varphi_x(\epsilon) \quad (3.1.2)$$

et si il existe deux constantes réelles non négatives  $\mathbf{C}$  et  $\mathbf{C}$  alors :

$$\mathbf{C}\varphi_x(h) \leq \mathbf{E} \left( \mathbf{K} \left( \frac{d(x, \mathbf{X})}{h} \right) \right) \leq \mathbf{C}'\varphi_x(h) \quad (3.1.3)$$

## 3.2 Estimateur à noyau de la densité conditionnelle

### 3.2.1 Estimateur à noyau

L'estimateur  $f_n(x)$  peut être généralisé en remplaçant la fonction de poids  $W(y)$  (la densité de probabilité uniforme) par une fonction de poids plus générale  $K$  (par exemple une densité de probabilité quelconque). Ceci donne le résultat qui suit.

#### 3.2.1.1 Définition et construction

**Définition 3.2.1.** Soit  $(\Omega, \mathbf{A}, \mathbf{P}_\Omega)$  un espace de probabilité. Soit  $X_1, X_2, \dots, X_n$  un échantillon i.i.d. de f.d.r  $F$  et d'une densité  $f$ . L'estimateur à noyau de la fonction de densité, notée  $\hat{f}_n(x)$  est définie par

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)$$

Où  $K$  est appelé fonction de poids (weight function) ou noyau (kernel function), et  $h$  est appelé paramètre de lissage (smoothing parameter) ou fenêtre (window width).

### 3.2.2 Etude du biais et de la variance

Lorsqu'on définit un estimateur à noyau, on a non-seulement le choix de la fenêtre  $h > 0$  mais aussi celui du noyau  $K$ . Il y a un certain nombre de conditions qui sont considérées comme usuelles pour les noyaux et qui permettent d'analyser le risque de l'estimateur à noyau qui en résulte.

**Hypothèse K :** On suppose que  $K$  vérifie les 4 conditions suivantes :

1.  $\int_{\mathbb{R}} K(u) du = 1$

CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

2.  $K$  est une fonction paire ou, plus généralement,  $\int_{\mathbb{R}} u K(u) du = 0$

3.  $\int_{\mathbb{R}} u^2 |K(u)| du < \infty$

4.  $\int_{\mathbb{R}} K(u)^2 du < \infty$

**Proposition 3.2.2.** *Si les trois premières conditions de l'hypothèse  $\mathbf{K}$  sont remplies, alors*

$$\text{Biais} \left[ \hat{f}_n(x) \right] = \frac{1}{2} f''(x) \mu_2 h^2 + o(h^2)$$

où  $\mu_2 = \int_{\mathbb{R}} K(u) u^2 du$  Si, de plus, la condition 4 de l'hypothèse  $\mathbf{K}$  est satisfaite, alors

$$\text{Var} \left[ \hat{f}_n(x) \right] = \frac{1}{nh} f(x) R(x) + o\left(\frac{1}{nh}\right)$$

où  $R(x) = \int_{\mathbb{R}} K^2(u) du$

**Preuve 3.2.3.** *Commençons par calculer le biais : Considérons l'estimateur à noyau*

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

où nous avons introduit la notation

$$K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$$

pour une version transformée de  $\mathbf{K}$ .  $E \left[ \hat{f}_n(x) \right] = E [K_h(x - X)]$  car les

CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

$X_i$  sont i.i.d.

$$\begin{aligned}
 &= \int_{\mathbb{R}} K_h(x - y) f(y) dy \\
 &= \int_{\mathbb{R}} K(u) f(x - uh) du \quad \text{avec } u = \frac{x - y}{h} \text{ et } du = -\frac{1}{h} dy \\
 &= \int_{\mathbb{R}} K(u) \left[ f(x) - f'(x)uh + \frac{1}{2}f''(x)u^2h^2 + \dots \right] du \quad \text{par Taylor} \\
 &= f(x) \int_{\mathbb{R}} K(u) du - f'(x)h \int_{\mathbb{R}} K(u)u du + \frac{1}{2}f''(x)h^2 \int_{\mathbb{R}} K(u)u^2 du + o(h^2)
 \end{aligned}$$

Alors

$$\begin{aligned}
 \text{Biais } \left[ \hat{f}_n(x) \right] &= \mathbf{E} \left[ \hat{f}_n(x) \right] - f(x) \\
 &= \frac{1}{2}f''(x)h^2 \int_{\mathbb{R}} K(u)u^2 du + o(h^2) \quad (3.2.1) \\
 &= \frac{1}{2}f''(x)\mu_2h^2 + o(h^2)
 \end{aligned}$$

Avec  $\mu_2 = \int_{\mathbb{R}} K(u)u^2 du$ . d'où la première assertion de la proposition.  
 Pour la variance on calcule :

$$\begin{aligned}
 \text{Var} \left[ \hat{f}_n(x) \right] &= \mathbf{E} \left[ \hat{f}_n^2(x) \right] - \left\{ \mathbf{E} \left[ \hat{f}_n(x) \right] \right\}^2 \\
 &= \frac{1}{n} \left\{ \mathbf{E} \left[ K_h^2(x - \mathbf{X}) \right] - \left( \mathbf{E} \left[ K_h(x - \mathbf{X}) \right] \right)^2 \right\} \\
 &= \frac{1}{nh^2} \int_{\mathbb{R}} K^2 \left( \frac{x - y}{h} \right) f(y) dy \quad \text{avec } u = \frac{x - y}{h} \\
 &= \frac{1}{nh} \int_{\mathbb{R}} K^2(u) f(x - uh) du \quad du \\
 &= \frac{1}{nh} \int_{\mathbb{R}} K^2(u) |f(x) - f'(x)hu + \dots| du - f'(x) \int_{\mathbb{R}} K^2(u)u du + o(1) \\
 &= \frac{1}{nh} f(x) \int_{\mathbb{R}} K^2(u) du + o(1)
 \end{aligned}$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

Donc nous trouvons que :

$$\begin{aligned}\text{Var} \left[ \hat{f}_n(x) \right] &= \frac{1}{nh} f(x) \int_{\mathbb{R}} K^2(u) du + o\left(\frac{1}{nh}\right) \\ &= \frac{1}{nh} f(x) \mathbf{R}(x) + o\left(\frac{1}{nh}\right)\end{aligned}\quad (3.2.2)$$

avec  $\mathbf{R}(x) = \int_{\mathbb{R}} K^2(u) du$

**Remarque 3.2.4.** Si  $h = h_n \rightarrow 0$  quand  $n \rightarrow \infty$ , alors

$$\text{Biais} \left[ \hat{f}_n(x) \right] \rightarrow 0$$

Si  $h = h_n \rightarrow 0$  et  $nh_n \rightarrow \infty$  quand  $n \rightarrow \infty$ , alors

$$\text{Var} \left[ \hat{f}_n(x) \right] \rightarrow 0$$

— Si  $h$  décroît alors le (Biais) et la variance ?

— Si  $h$  augmente alors le (Biais) / et la variance

Il faut donc essayer de choisir un  $h$  qui fasse un compromis entre le (Biais)<sup>2</sup> et la variance.

### 3.2.2.1 Expression d'erreur quadratique moyenne (MSE)

Les expressions asymptotiques du biais et de la variance nous permettent de trouver l'expression asymptotique pour la (MSE) et l'erreur quadratique moyenne intégrée (MISE), Ces expressions ont été obtenues sous certaines conditions / sur  $K$  et en supposant que la densité de probabilité  $f$  avait toutes les dérivées (continues) nécessaires. A partir de (3.2.1) et (3.2.2) on peut obtenir facilement l'expression suivante pour la MSE et la MISE.

$$\text{MSE} \left[ \hat{f}_n(x) \right] = \frac{1}{4} h^4 \mu_2^2 (f''(x))^2 + \frac{1}{nh} f(x) \mathbf{R}(x) + o\left(h^4 + \frac{1}{nh}\right) \quad (3.2.3)$$

$$\text{MISE} \left[ \hat{f}_n(x) \right] = \frac{1}{4} h^4 \mu_2^2 \int (f''(x))^2 dx + \frac{1}{nh} f(x) \mathbf{R}(x) + o\left(h^4 + \frac{1}{nh}\right) \quad (3.2.4)$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

sous des conditions appropriées d'intégrabilité de  $f$  et ses dérivées. On note l'approximation asymptotique de la MSE par

$$\text{AMSE} \left[ \widehat{f}_n(x) \right] = \frac{1}{4} h^4 \mu_2^2 (f''(x))^2 + \frac{1}{nh} f(x) \mathbf{R}(x) \quad (3.2.5)$$

et l'approximation asymptotique de la MISE par

$$\text{AMISE} \left[ \widehat{f}_n(\cdot) \right] = \frac{1}{4} h^4 \mu_2^2 \mathbf{R}(f'') + \frac{1}{nh} f(x) \mathbf{R}(x) \quad (3.2.6)$$

### 3.2.2.2 L'estimateur à noyau de la densité conditionnelle

Soit  $(X_1, Y_1), \dots, (X_n, Y_n)$  un échantillon aléatoire du couple  $(X, Y)$  indépendant, identiquement distribué qui est à valeurs dans  $\mathcal{H} \times \mathbb{R}$ , où  $\mathcal{H}$  est un espace de Hilbert muni du produit scalaire  $\langle \cdot, \cdot \rangle$ , et dont la norme associée est notée  $\| \cdot \|$ . L'estimateur à noyau de la densité conditionnelle  $f(y | x)$  noté  $\widehat{f}(y | x)$  est défini par :

$$\widehat{f}(y | x) = \frac{\sum_{i=1}^n \mathbf{K} \left( \frac{d(X_i, x)}{h} \right) \frac{\partial}{\partial y} \mathbf{H} \left( \frac{y - Y_i}{g} \right)}{\sum_{i=1}^n \mathbf{K} \left( \frac{d(X_i, x)}{h} \right)}, \forall y \in \mathbb{R} \quad (3.2.7)$$

où  $\mathbf{H}$  est défini par

$$\forall u \in \mathbb{R}, \mathbf{H}(u) = \int_{-\infty}^u \mathbf{K}_0(v) dv$$

et

$$d(X_i, x) = \|X_i - x\|.$$

La fonction  $\mathbf{K}$  est un noyau de type I ou de type II et la fonction  $\mathbf{K}_0$  est un noyau de type 0 et  $h = h(n)$  ( resp.  $g = g(n)$ ) est une suite de nombres réels positifs qui tend vers zéro lorsque  $n$  tend vers l'infinie.

Il est aussi appelé le paramètre de lissage ou largeur de fenêtre. Tout au long de notre travail, nous noterons par  $C$  et  $C'$  deux constantes génériques et strictement positives.

Afin d'établir la convergence presque complète (p.co.) de notre estimateur on considère les hypothèses suivantes.

Soient  $x$  ( resp.  $y$  ) un élément de  $\mathcal{H}$  ( resp. de  $\mathbb{R}$  ),  $N_x \subset \mathcal{H}$  un voisinage de  $x$  et  $S$  un sous ensemble compact de  $\mathbf{R}$  tels que :

$$\mathbf{P}(d(\mathbf{X}, x) < h) = \varphi_x(h) > 0 \quad (3.2.8)$$

$$\exists C > 0, \forall (x, x') \in \mathbb{R} \times \mathbb{R}, |\mathbf{K}_0(x) - \mathbf{K}_0(x')| \leq C |x - x'|,$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

$$\lim_{n \rightarrow \infty} \frac{\log n}{ng\varphi_x(h)} = 0 \text{ et } \exists \alpha > 0, \lim_{n \rightarrow \infty} gn^\alpha = \infty$$

$$\exists C > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > C\epsilon\varphi_x(\epsilon),$$

$$\exists C_x > 0 \text{ tel que } \forall (y_1, y_2) \in \mathbb{S}^2, \forall (x_1, x_2) \in \mathbb{N}_x \times \mathbb{N}_x,$$

$$|f(y_1 | x_1) - f(y_2 | x_2)| \leq C_x \left( d^{\beta_1}(x_1, x_2) + |y_1 - y_2|^{\beta_2} \right), \beta_1 > 0, \beta_2 > 0.$$

**Théorème 3.2.5.** *Sous les conditions (3.2.4), (3.2.5), (3.2.6), (3.2.7) et (3.2.8), nous avons pour tout nombre réel fixé  $y$  :*

$$\hat{f}(y | x) - f(y | x) = o\left(h^{\beta_1}\right) + o\left(g^{\beta_2}\right) + o_{p.ca} \left( \sqrt{\frac{\log n}{ng\varphi_x(h)}} \right) \quad (3.2.9)$$

**Preuve 3.2.6.** *La preuve est basée sur la décomposition qui suit*

$$\begin{aligned} \hat{f}(y | x) - f(y | x) &= \frac{(\hat{r}_3(x, y) - E\hat{r}_3(x, y)) - (f(y | x) - E\hat{r}_3(x, y))}{\hat{r}_1(x)} \\ &\quad - \frac{f(y | x)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \end{aligned} \quad (3.2.10)$$

où  $\hat{r}_1$  est défini par avec

$$\begin{aligned} \hat{r}_1(x) &= \frac{1}{n} \sum_{i=1}^n \Delta_i \\ \Delta_i &= \frac{\mathbf{K}\left(\frac{d(x, x_i)}{h}\right)}{\mathbf{E}\left(\mathbf{K}\left(\frac{d(x, X_i)}{h}\right)\right)} \end{aligned} \quad (3.2.11)$$

et où

$$\hat{r}_3(x, y) = \hat{r}_1(x) \hat{f}(y | x) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Gamma_i(y) \quad (3.2.12)$$

avec

$$\Gamma_i(y) = \frac{1}{g} K_0\left(\frac{y - Y_i}{g}\right) \quad (3.2.13)$$

Ainsi, la preuve est une conséquence directe des résultats qui suivent.

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

**Lemme 3.2.7.** *Sous les hypothèses (3.2.5) et (3.2.8), lorsque  $n$  tend vers l'infini, nous avons :*

$$\mathbf{E}_3(x, y) - f(y | x) = \mathbf{0} \left( h^{\beta_1} \right) + \mathbf{0} \left( g^{\beta_2} \right). \quad (3.2.14)$$

**Preuve 3.2.8.** *Puisque  $\mathbf{E}\Delta_i = 1$  et puisque  $\mathbf{K}_0$  est une fonction intégrable, nous avons :*

$$\begin{aligned} \mathbf{E}\widehat{r}_3(x, y) - f(y | x) &= \mathbf{E}\Delta_1\Gamma_1(y) - f(y | x) \\ &= \mathbf{E} \left( \Delta_1 \left( \mathbf{E} \left( \Gamma_1(y) | \mathbf{X} \right) - f(y | x) \right) \right) \\ &= \mathbf{E} \left( \Delta_1 \int_{\mathbf{R}} \frac{1}{g} \mathbf{K}_0 \left( \frac{y-u}{g} \right) f(u | \mathbf{X}) du - f(y | x) \right) \\ &= \mathbf{E} \left( \Delta_1 \int_{\mathbf{R}} \frac{1}{g} \mathbf{K}_0 \left( \frac{y-u}{g} \right) (f(u | \mathbf{X}) - f(y | x)) du \right) \\ &= \mathbf{E} \left( \Delta_1 \int_{\mathbf{R}} \frac{1}{g} \mathbf{K}_0(v) (f(y - vg | \mathbf{X}) - f(y | x)) dv \right) \\ &= \mathbf{E} \left( \mathbf{1}_{\beta(x, h)} \Delta_1 \int_{\mathbf{R}} \frac{1}{g} \mathbf{K}_0(v) (f(y - vg | \mathbf{X}) - f(y | x)) dv \right) \end{aligned} \quad (3.2.15)$$

Compte tenu du fait que le support  $\mathbf{K}_0 = [-1, 1]$  et puisque  $h$  et  $g$  tendent vers zéro, la condition de Hölder (3.2.8) permet d'écrire que :

$$\sup_{v \in [-1, 1]} |f(y - vg | \mathbf{X}) - f(y | x)| = \mathbf{0} \left( h^{\beta_1} \right) + \left| \mathbf{0} \left( g^{\beta_2} \right) \right|^{\beta_2} \quad (3.2.16)$$

Ainsi, le résultat (3.2.14) découle directement en combinant (3.2.15), (3.2.16) et le fait que  $\mathbf{E}\Delta_1 = 1$ .

**Lemme 3.2.9.** *Sous les hypothèses du théorème, lorsque  $n$  tends vers l'infini, nous avons :*

$$\widehat{r}_3(x, y) - \mathbf{E}\widehat{r}_3(x, y) = \mathbf{0} \text{ p.co. } \left( \sqrt{\frac{\log n}{ng\varphi_x(h)}} \right) \quad (3.2.17)$$

**Preuve 3.2.10.** *Pour cela, nous utilisons la décomposition suivante :*

$$\widehat{r}_3(x, y) - \mathbf{E}\widehat{r}_3(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbf{Z}_i \quad (3.2.18)$$

Où

$$\mathbf{Z}_i = (\mathbf{T}_i - \mathbf{E}\mathbf{T}_i) \quad \text{et} \quad \mathbf{T}_i = \Delta_i \Gamma_i(y) \quad (3.2.19)$$

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

Afin d'appliquer une inégalité de type-Bernstein, nous commençons par montrer

$$|\mathbf{T}_i| \leq \frac{\mathbf{C}}{g\varphi_x(h)} \quad \text{et} \quad \mathbf{E}\mathbf{T}_i^2 \leq \frac{\mathbf{C}}{g\varphi_x(h)} \quad (3.2.20)$$

En utilisant le Lemme 3.2.1 ou le Lemme 3.2.2 et en tenant compte de l'hypothèse (3.2.7) et du fait que  $\mathbf{K}$  est de type I ou de type II, nous avons :

$$\mathbf{C}\varphi_x(h) \leq \mathbf{E} \left( \mathbf{K} \left( \frac{d(\mathbf{X}, x)}{h} \right) \right) \leq \mathbf{C}'\varphi_x(h) \quad (3.2.21)$$

et en utilisant le dernier résultat et puisque  $\mathbf{K}_0$  est borné, nous obtenons

$$|\mathbf{T}_i| \leq \frac{\mathbf{C}}{g\varphi_x(h)}$$

Le second moment des variables  $\mathbf{T}_i$  peut être calculé en utilisant l'intégration par changement de variable :

$$\begin{aligned} \mathbf{E}\mathbf{T}_i^2 &= \mathbf{E} \left( \Delta_i^2 \frac{1}{g^2} \mathbf{K}_0^2 \left( \frac{y - \mathbf{Y}_i}{g} \right) \right) = \mathbf{E} \left( \mathbf{E} \left( \Delta_i^2 \frac{1}{g^2} \mathbf{K}_0^2 \left( \frac{y - \mathbf{Y}_i}{g} \right) \mid \mathbf{X} = x \right) \right) \\ &= \frac{1}{g^2} \mathbf{E} \left( \Delta_i^2 \int_{\mathbb{R}} \mathbf{K}_0^2 \left( \frac{y - u}{g} \right) f(u \mid \mathbf{X}) du \right) \\ &= \frac{g}{g^2} \mathbf{E} \left( \Delta_i^2 \int_{\mathbb{R}} \mathbf{K}_0^2(v) f(y - vg \mid \mathbf{X}) dv \right) \\ &= \frac{1}{g} \mathbf{E}\Delta_i^2 \end{aligned} \quad (3.2.22)$$

Puisque  $0 < \int \mathbf{K}^2 < \infty$ , si  $\mathbf{K}$  est de type I (resp.II)

alors  $\frac{\mathbf{K}^2}{\int \mathbf{K}^2}$  est aussi de type I (resp.II).

Ainsi, en appliquant le Lemme 3.2.1 ou le Lemme 3.2.2 on trouve

$$\mathbf{C}\varphi_x(h) \leq \mathbf{E} \left( \mathbf{K}^2 \left( \frac{d(\mathbf{X}_i, x)}{h} \right) \right) \leq \mathbf{C}'\varphi_x(h) \quad (3.2.23)$$

et en utilisant le dernier résultat, on écrit que

$$\frac{\mathbf{C}}{\varphi_x(h)} \leq \mathbf{E}\Delta_i^2 \leq \frac{\mathbf{C}'}{\varphi_x(h)} \quad (3.2.24)$$

ce qui implique que

$$\mathbf{E}\mathbf{T}_i^2 \leq \frac{\mathbf{C}}{g\varphi_x(h)} \quad (3.2.25)$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

En tenant compte de (3.2.20), on peut appliquer l'inégalité de type-Bernstein-type donné par le corollaire A.9 ( voir Ferraty [2]), et on obtient :

$$\forall \epsilon \geq 0, \mathbf{P} [ |\hat{r}_3(x, y) - \mathbf{E}\hat{r}_3(x, y)| > \epsilon ] \leq 2 \exp \frac{\epsilon^2 n g \varphi_x(h)}{2C'(1 + \epsilon)} \quad (3.2.26)$$

Puisque la suite  $\frac{\log n}{n g \varphi_x(h)}$  tend vers zéro,

en choisissant  $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{n g \varphi_x(h)}}$  dans le résultat (3.2.26) nous obtenons directement

$$\begin{aligned} \mathbf{P} \left[ |\hat{r}_3(x, y) - \mathbf{E}\hat{r}_3(x, y)| > \epsilon_0 \sqrt{\frac{\log n}{n g \varphi_x(h)}} \right] &\leq 2 \exp \frac{\epsilon_0^2 \log n}{2C' \left( 1 + \epsilon_0 \sqrt{\frac{\log n}{n g \varphi_x(h)}} \right)} \\ &\leq 2n^{-C\epsilon_0^2}, \end{aligned}$$

### 3.3 Convergence presque complète

#### 3.3.1 Inégalité de Bernshtein Fréchet

**Lemme 3.3.1.** *Soit  $X_1, X_2, \dots, X_n$  une suite de variables aléatoires indépendantes, tel que,  $\alpha_i \leq X_i \leq \beta_i$  et  $\alpha_i, \beta_i \in \mathbb{R}$  Alors,  $\forall t > 0$  on a :*

$$\mathbf{P} \left[ \left| \sum_{i=1}^n (X_i - \mathbf{E}(X_i)) \right| \geq t \right] \leq 2 \exp \left( \frac{-2t^2}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} \right)$$

**Preuve 3.3.2.** *On montre que :*

$$\mathbf{I}_{\mathbf{A}} = \left( \sum_{i=1}^n (X_i - \mathbf{E}(X_i)) - t \geq 0 \right) \leq \exp \left( h \sum_{i=1}^n (X_i - \mathbf{E}(X_i)) - t \right) \quad (3.3.1)$$

et en déduire l'Inégalité de Bernshtein Fréchet pour  $\forall t > 0$

$$\mathbf{P} \left[ \left| \sum_{i=1}^n (X_i - \mathbf{E}(X_i)) \right| \geq t \right] \leq 2 \exp \left( \frac{-2t^2}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} \right)$$

$$\mathbf{I}_{\mathbf{A}}(w) = \begin{cases} 1 & \text{si } w \in \mathbf{A} \\ 0 & \text{sinon} \end{cases}$$

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

★Si $I_A = 0$

$$\sum_{i=1}^n (X_i - E(X_i)) - t \leq 0$$

★Si $I_A = 1$

$$\sum_{i=1}^n (X_i - E(X_i)) - t \geq 0$$

$$\exp(h \sum_{i=1}^n (X_i - E(X_i)) - t) \geq 1$$

Alors (3.3.1) est vrai pour les deux cas. On sait que  $E[I_A(\mathbf{X})] = P(A)$

$$\begin{aligned} P(A) &= P\left(\sum_{i=1}^n (X_i - E(X_i)) - t \geq 0\right) \\ &\leq E\left[\exp\left(h \sum_{i=1}^n (X_i - E(X_i)) - t\right)\right] \\ &\leq \exp(-ht) \cdot E\left[\exp\left(h \sum_{i=1}^n (X_i - E(X_i))\right)\right] \\ &\leq \exp(-ht) \cdot \prod_{i=1}^n E[\exp(h(X_i - E(X_i)))] \\ &\leq \prod_{i=1}^n E[\exp(-hE(X_i))] \cdot E[\exp(hX_i)] \cdot \exp(-ht) \end{aligned}$$

Pour  $(hX_i)$  on va utiliser le fait que cette fonction est convexe, on pose

$$\varphi(X_i) = \exp(hX_i)$$

où  $\varphi$  est convexe vérifie

$$\varphi(\alpha x + \beta y) \leq \alpha \varphi(x) + \beta \varphi(y) \quad \text{où } \alpha, \beta \in \mathbb{R}$$

Posons :

$$\alpha = \frac{\beta_i - X_i}{\beta_i - \alpha_i} \quad \text{et} \quad \beta = \frac{X_i - \alpha_i}{\beta_i - \alpha_i}$$

Il est clair que  $\alpha + \beta = 1$

On pose

$$X_i = \alpha x + \beta y$$

$$x = \alpha_i$$

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

$$y = \beta_i$$

$$\begin{aligned} \varphi(\alpha x + \beta y) &= \exp(h(\alpha x + \beta y)) \\ &\leq \alpha \varphi(x) + \beta \varphi(y) \\ \Rightarrow \exp(hX_i) &\leq \frac{\beta_i - X_i}{\beta_i - \alpha_i} \exp(h\alpha_i) + \frac{X_i - \alpha_i}{\beta_i - \alpha_i} \exp(h\beta_i) \\ \Rightarrow E[\exp(hX_i)] &\leq \frac{\beta_i - E[X_i]}{\beta_i - \alpha_i} \exp(h\alpha_i) + \frac{E[X_i] - \alpha_i}{\beta_i - \alpha_i} \exp(h\beta_i) \end{aligned} \quad (3.3.2)$$

Alors

$$E[\exp(-hE(X_i))] \cdot E[\exp(hX_i)] \leq E[\exp(-hE(X_i))] \cdot \left[ \frac{\beta_i - E[X_i]}{\beta_i - \alpha_i} \exp(h\alpha_i) + \frac{E[X_i] - \alpha_i}{\beta_i - \alpha_i} \exp(h\beta_i) \right] \quad (3.3.3)$$

On essaye de mettre (3.3.3) sous la forme  $\exp(\psi(h_i))$  tel que :  $h_i = h(\beta_i - \alpha_i)$  c'est à dire ;

$$(3.3.2)(h) = \exp(\psi(h_i))$$

D'après le developpement limitée de  $\psi(h_i)$  on a :

$$\psi(h_i) = \psi(0) + \psi'(0)h_i + \frac{1}{2}\psi''(\xi)h_i^2 \quad \text{où } \xi \in [0, h_i]$$

On trouve que :

$$\psi(0) = 0 \quad , \quad \psi'(0) = 0 \quad \text{et} \quad \psi''(h_i) \leq \frac{1}{4}$$

Donc :

$$|\psi(h_i)| \leq \frac{1}{8}h_i^2 = \frac{h^2(\beta_i - \alpha_i)^2}{8}$$

On peut dire que :

$$(3.3.2)(h_i) = \exp\left(\frac{h^2(\beta_i - \alpha_i)^2}{8}\right)$$

celà veut dire

$$\begin{aligned} P\left(\sum_{i=1}^n (X_i - E(X_i)) - t \geq 0\right) &\leq \exp(-ht) \cdot \prod_{i=1}^n E[\exp(h(X_i - E(X_i)))] \\ &\leq \exp(-ht) \cdot \exp\left[\frac{h^2 \sum_{i=1}^n (\beta_i - \alpha_i)^2}{8}\right] \\ &\leq \exp\left(-ht + \frac{h^2 \sum_{i=1}^n (\beta_i - \alpha_i)^2}{8}\right) \end{aligned}$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

La relation est vraie pour  $h \geq 0$ . On pose :  $h = \frac{4t}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} \geq 0$

$$\begin{aligned} \mathbf{P} \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{E}(\mathbf{X}_i)) - t \geq 0 \right) &\leq \exp \left( \frac{-4t^2}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} + \frac{16t^2}{8 \sum_{i=1}^n (\beta_i - \alpha_i)^2} \right) \\ &\leq \exp \left( \frac{-2t^2}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} \right) \end{aligned}$$

On trouve le même résultat pour :

$$\mathbf{A} = \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{E}(\mathbf{X}_i)) - t \leq 0 \right)$$

Enfin on conclut :

$$\mathbf{P} \left[ \left| \sum_{i=1}^n (\mathbf{X}_i - \mathbf{E}(\mathbf{X}_i)) \right| \geq t \right] \leq 2 \exp \left( \frac{-2t^2}{\sum_{i=1}^n (\beta_i - \alpha_i)^2} \right)$$

### 3.3.2 Convergence presque complète

**Définition 1.3** La convergence presque complète. Soit  $(\mathbf{X}_n)_n, n \in \mathbb{N}$ , une suite de variables aléatoires. On dit que  $\mathbf{X}_n$  converge presque complètement vers  $\mathbf{X}$  si :

$$\forall \varepsilon > 0$$

$$\sum_{n=0}^{\infty} \mathbf{P} [|\mathbf{X}_n - \mathbf{X}| > \varepsilon] < \infty$$

ç veut dire,

$$\mathbf{X}_n \rightarrow \mathbf{X} \text{ quand } n \rightarrow \infty$$

### 3.3.3 La vitesse de Convergence presque complète

**Définition 3.3.3.** vitesse de convergence en p.c.o. On dit que  $\mathbf{X}_n = 0$  ( $\mathbf{Y}_n$ ) en p.c.o ( $\mathbf{X}_n$  converge vers 0 pour une vitesse  $\mathbf{Y}_n$ ) Si  $\exists \varepsilon > 0$ ,

$$\sum_{n=0}^{\infty} \mathbf{P} [|\mathbf{X}_n| > \varepsilon |\mathbf{Y}_n|] < \infty$$

## CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

**Théorème 3.3.4.** Soit  $X_1, X_2, \dots, X_n$  un  $n$ -chantillon de  $X$  de fonction de répartition  $F$  et  $F_n$  la fonction empirique. Alors, pour tout  $x$  on a :

$$F_n(x) - F(x) = o\left(\sqrt{\frac{\log n}{n}}\right) \text{ p.c.o}$$

**Preuve 3.3.5.** D'après la définition précédent de vitesse de convergence, il suffit de montrer que :  $\exists \varepsilon > 0$ ,

$$\sum_{n=0}^{\infty} \mathbf{P} \left[ |F_n(x) - F(x)| > \varepsilon \sqrt{\frac{\log n}{n}} \right] < \infty$$

on note :

$$A = \sum_{n=0}^{\infty} \mathbf{P} \left[ |F_n(x) - F(x)| > \varepsilon \sqrt{\frac{\log n}{n}} \right]$$

on a :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{]-\infty, x]}(X_i) \quad \text{et } F(x) = \mathbf{E}[F_n(x)]$$

Alors A devient :

$$\begin{aligned} A &= \sum_{n=0}^{\infty} \mathbf{P} \left[ \left| \frac{1}{n} \sum_{i=1}^n I_{]-\infty, x]}(X_i) - \frac{1}{n} \mathbf{E} \left[ \sum_{i=1}^n I_{]-\infty, x]}(X_i) \right] \right| > \varepsilon \sqrt{\frac{\log n}{n}} \right] \\ &= \sum_{n=0}^{\infty} \mathbf{P} \left[ \sum_{i=1}^n \left| I_{]-\infty, x]}(X_i) - \mathbf{E} \left[ I_{]-\infty, x]}(X_i) \right] \right| > \varepsilon \sqrt{n \log n} \right] \end{aligned}$$

On utilise l'inégalité de Bernshtein Fréchet . On sait que :

$$0 \leq I_{]-\infty, x]}(X_i) \leq 1$$

Possons  $t = \varepsilon \sqrt{n \log n}$ . donc, par identification  $\alpha_i = 0$  et  $\beta_i = 1$ . On trouve :

$$\begin{aligned} A &= \sum_{n=0}^{\infty} \mathbf{P} \left( \sum_{i=1}^n \left| I_{]-\infty, x]}(X_i) - \mathbf{E} \left[ I_{]-\infty, x]}(X_i) \right] \right| > t \right) \leq \sum_{n=0}^{\infty} 2 \exp \left( \frac{-2\varepsilon^2 n \log n}{n} \right) \\ &\leq \sum_{n=0}^{\infty} 2 \exp \left( \log n^{-2\varepsilon^2} \right) \\ &\leq \sum_{n=0}^{\infty} 2n^{-2\varepsilon^2} \\ &\leq \sum_{n=0}^{\infty} \frac{2}{n^{2\varepsilon^2}} \end{aligned}$$

### CHAPITRE 3. LA MÉTHODE DE NOYAU DE LA FONCTION DE DENSITÉ CONDITIONNELLE

---

Donc,  $\exists \varepsilon > 0$  tel que :

$$2\varepsilon^2 > 1 \Rightarrow \varepsilon > \frac{1}{\sqrt{2}}$$

Pour que la série converge. D'où

$$\sum_{n=0}^{\infty} \mathbf{P} \left[ |\mathbf{F}_n(x) - \mathbf{F}(x)| > \varepsilon \sqrt{\frac{\log n}{n}} \right] < \infty$$

De cela, nous concluons que

$$\mathbf{F}_n(x) - \mathbf{F}(x) = o \left( \sqrt{\frac{\log n}{n}} \right) \text{ p.c.o}$$

## Chapitre 4

# Application et exploration des observations sur des données simulées

La densité conditionnelle est un outil fondamental pour décrire la relation entre deux variables aléatoires.

Dans ce chapitre, nous allons déterminer ce lien par la méthode d'estimation non paramétrique de noyau.

L'objectif principal est de montrer à l'aide de données simulées l'applicabilité de cette méthode dans le cadre fonctionnel.

Dans un premier temps, nous illustrons les variables aléatoires fonctionnelle sont des courbes.

Ensuite, nous proposons une implémentation de l'estimateur de la densité conditionnelle proposé.

## 4.1 Simulée les données fonctionnelles

Nous récoltons des observations fonctionnelles générales à l'aide du processus suivant :

$$x_i(t) = \cos \left( w_i + \pi \left( \frac{2t}{100} - 1 \right) \right)$$

ou  $w_i$  est distribué selon la loi uniforme  $U \left( 0, \frac{\pi}{4} \right)$  Nous effectuerons une étude de simulation à l'aide du logiciel statistique *R*

### 4.1.1 Programme sous *R* d'un courbe aléatoire simulée

Un courbe aléatoire simulée :

```
x <- matrix(nrow = 100, ncol=100, byrow=TRUE)
xx <- (1 :100)
w= runif(100,0,pi/4)
# échantillon de loi uniforme
for (i in 1 :100) {
for (j in 1 :100) {
x[i,j]=cos(w[i] + pi *((2 * j/100) - 1))^4
}
}
# boucle pour les observations de v.a.f
yy <- x[1, ] #Vecteur de premier v.a.f
plot(yy,type="l",col="blue",xlab="t", ylab="X(t)",
main="Un courbe aléatoire
simulée")
# Tracer un courbe qui représente un v.a.f
```

### 4.1.2 Résultat de simulation d'un variable aléatoire fonctionnelle

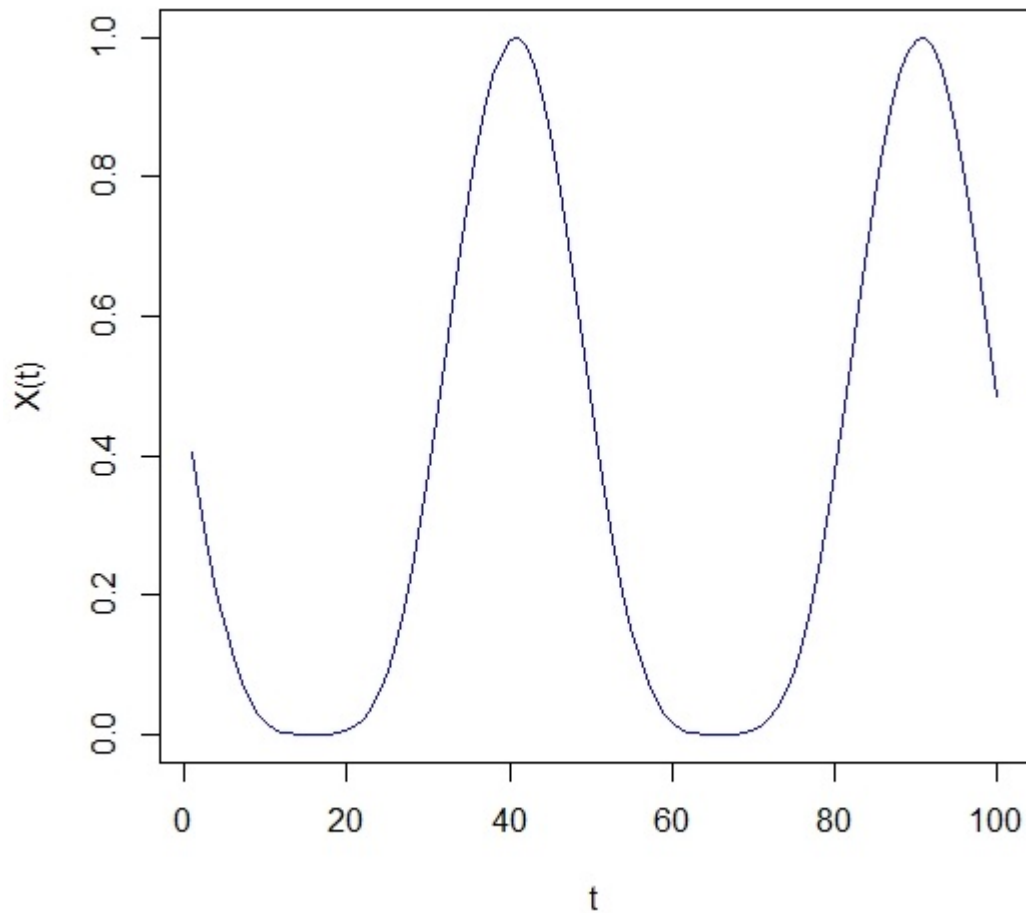


FIGURE 4.1 – Un courbe aléatoire simulée

## CHAPITRE 4. APPLICATION ET EXPLORATION DES OBSERVATIONS SUR DES DONNÉES SIMULÉES

---

### 4.1.3 Programme sous R de plusieurs courbes aléatoires simulées

#### Plusieurs courbes aléatoires simulées :

```
x <- matrix(nrow = 100, ncol=100, byrow=TRUE)
xx <- (1 :100)
w= runif(100,0,pi/4)
# échantillon de loi uniforme
for (i in 1 :100) {
for (j in 1 :100) {
x[i,j]=cos(w[i] + pi *((2 * j/100) - 1))^4
}
}
# boucle pour les observations de v.a.f
yy <- x[1, ]
#Vecteur de premier v.a.f

plot(yy,type="l",col="blue",xlab="t", ylab="X(t)",
main="Un courbe aléatoire
simulée")
# Tracer un courbe qui représente un v.a.f
plot(xx,yy,type="l",col="blue",xlab="t", ylab="X(t)",
main="plusieurs courbes
aléatoires simulées")
for (j in 1 :100) {
lines(xx,x[j,], col="black")
}
# Tracer les courbes qui représente les v.a.f.
```

#### 4.1.4 Résultat de plusieurs variables aléatoires fonctionnelles

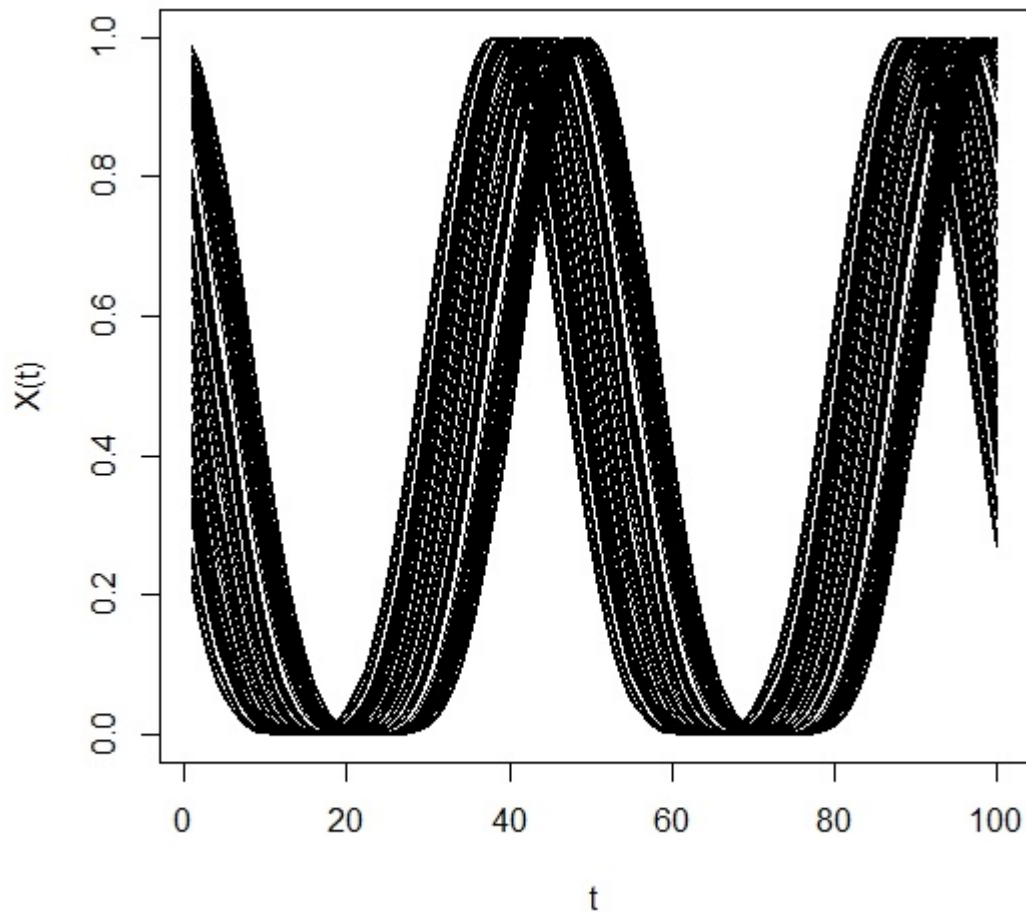


FIGURE 4.2 – Plusieurs courbes aléatoires simulées

## 4.2 Simulation les variables de réponse réels

Pour trouver les variables de réponses réels que nous donnerons le terme de régression multiple suivante :

$$Y_i = r(X_i) + \varepsilon_i$$

tell que les  $\mathcal{X}_i$  est une matrice, et pour chaque  $i$  nous avons un vecteur qui présente une variables aléatoire fonctionnelle.

Et l'erreur  $\varepsilon_i$  est centrée réduit.

On utilise les variables aléatoires fonctionnelles qui sont utilisé dans la section précédente.

On suit les étapes de l'algorithme suivant :

### 4.2.1 Algorithme 1

1. Générer un échantillon de taille  $n$  d'une variable normale centrée réduit qui présent l'erreur  $\varepsilon_i$  ( $\varepsilon = \mathbf{rnorm}(n)$ )
2. Nous choisirons cette fonction pour représenter la régression :

$$r(\mathcal{X}_i) = \mathcal{X}_i^2 + 2$$

3. Nous prendrons les valeurs de variables de réponse.

### 4.2.2 Programme sous R

```
n<-100
err<-rnorm(100,0,1) # èchantillon de loi N(0; 1)
f<-function(x)
for (i in 1 : 100) {
f[i]<-(x[i,]^2 + 2)
}
```

## CHAPITRE 4. APPLICATION ET EXPLORATION DES OBSERVATIONS SUR DES DONNÉES SIMULÉES

---

```
yr<-f[i]+err[i]
for (i in 1 : 100) {
yr[i]=f[i]+err[i]
}
plot(f,type="l",col="blue",xlab="x", ylab="f(x)",
main="courbe de f")
for (i in 1 :100){
lines(f[i], col="red")
}
```

### 4.3 Simulation l'estimateur de densité conditionnelle

Dans cette section, nous calculons et représentons graphiquement l'estimateur de densité conditionnelle.

Il s'agit de l'estimateur proposé par Ferraty et Vieu(2006).

Cette estimation pour le couple  $(\chi, y)$  de  $(\mathcal{X}, \mathcal{Y})$  est donnée par :

$$\hat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i)) \frac{1}{g} HK_0(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))}$$

Afin d'estimer doit choisir  $K$  un type de noyaux présentée dans le chapitre 1 , nous choisirons le noyau gaussien sous forme :

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp^{-u^2/2}$$

et le noyau  $K_0$  est un noyau symétrique, nous choisirons le noyau rectangulaire :

$$K_0(u) = \frac{1}{2} \mathbf{1}_{[-1,1]}(u)$$

Pour estimer plus facilement nous calculerons les noyaux composes  $HK_0(y - Y_i)$ , tel que :

$$\begin{aligned} \forall \mu \in \mathbb{R}, H(\mu) &= \int_{-\infty}^{\mu} K_0(v) dv \\ &= \int_{-1}^{\mu} \frac{1}{2} \mathbf{1}_{[-1,1]}(v) dv \\ &= \frac{\mu + 1}{2} \end{aligned}$$

## CHAPITRE 4. APPLICATION ET EXPLORATION DES OBSERVATIONS SUR DES DONNÉES SIMULÉES

---

et aussi :

$$\begin{aligned} K_0(y - Y_i) &= \frac{1}{2} \mathbf{1}_{[-1,1]}(y - Y_i) \\ &= \begin{cases} \frac{1}{2} & \text{si } -1 \leq y - Y_i \leq 1 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

alors les noyaux composés  $HK_0(y - Y_i)$  est sous forme :

$$\begin{aligned} H(K_0(y - Y_i)) &= \frac{1}{2} (K_0(y - Y_i)) + \frac{1}{2} \\ &= \begin{cases} \frac{3}{4} & \text{si } -1 \leq y - Y_i \leq 1 \\ \frac{1}{2} & \text{sinon} \end{cases} \end{aligned}$$

On suit les étape de l'algorithme suivant pour simulée l'estimateur  $\widehat{f}_Y^X$  à couple  $(\chi, y)$  (on fixée  $(\chi_1, y_1)$ ) :

### 4.3.1 Algorithme 2

1. On choisit la valeur de paramètre de lissage  $h$  et  $g$ .
2. Calcule  $y - Y_i$
3. On insérer  $K(h^{-1}d(\chi, \mathcal{X}_i))$
4. On insérer  $H(K_0(y - Y_i))$
5. On doit calculer l'estimateur  $\widehat{f}_Y^X$
6. On présente  $\widehat{f}_Y^X$

### 4.3.2 Programme sous R

```
h <- 0.44 # paramètre de lissage
z <- yr[1]-yr[i]
zz <- (x[1,]-x[i,])/h
for (i in 1 :100){
z[i]=yr[1]-yr[i] zz[i]=(x[1,]-x[i,])/h
}
g <- 0.34 # paramètre de lissage
kz <- fonction(z)
for (i in 1 :100) {
kz[i]=(1/sqrt(2*pi))*exp(-(1/2)*((z[i])^2)) }
```

## CHAPITRE 4. APPLICATION ET EXPLORATION DES OBSERVATIONS SUR DES DONNÉES SIMULÉES

---

```
hk <- function(z) # noyau composèe HK0
for (i in 1 :100){
if(z[i]<=1 z[i]>=-1)
hk[i]<- 3/4
else hk[i]<- 1/2
}
s<-function(kz,hk) for (i in 1 :100) {
s[i]<-(sum((1/g)*(kz[i]*hk[i])))
}
ss<-function(kz)
for(i in 1 :100)
ss[i]<-sum(kz[i])
}
fchp <- numeric(n) # l'estimateur  $\hat{f}_Y^X$ 
Y
for (i in 1 :100){ fchp[i]<-sum(s[i]*ss[i])/(sum(ss[i])) }
plot(fchp,type="l",col="black",xlab="t", ylab="f,fchp",xlim=c(0,100) ,ylim=c(0,2),
lwd="2",main="Courbe de f et fchp" )
for(i in 1 :100) {
lines(fchp, col="black",lwd="2")
} # présentèe b fX
```

### 4.4 Comparison

Afin de faire la comparaison, nous devons proposer une autre fonction, qui doit vérifier les conditions de la fonction de densité (i.e  $\int_{-\infty}^{+\infty} f = 1$  et  $f$  continue,) On choisit la fonction suivante :

$$p(\chi, y) = 2 \exp -(\chi + y) \mathbf{1}_{\{0 \leq \chi \leq y\}}$$

## CHAPITRE 4. APPLICATION ET EXPLORATION DES OBSERVATIONS SUR DES DONNÉES SIMULÉES

---

On présent programme de  $p$  sous R :

### 4.4.1 Programme sous R

```
p<-function(x)
for (i in 1 :100) {
if(x[i]<=yr[i] x[i]>=0)
p[i]<- ( 2*exp(-(x[i]+yr[i])) )
else p[i]<- 0
}
plot(p,col="blue")
plot(fchp,type="l",col="black",xlab="t", ylab="f,fchp",xlim=c(0,100)
,ylim=c(0,5),main="Courbe de f et fchp" )
for (i in 1 :100){
lines(p, col="blue",lwd="2")
}
```

## 4.5 Résultats des simulations

### 4.5.1 Simulation de l'estimateur de densité conditionnelle

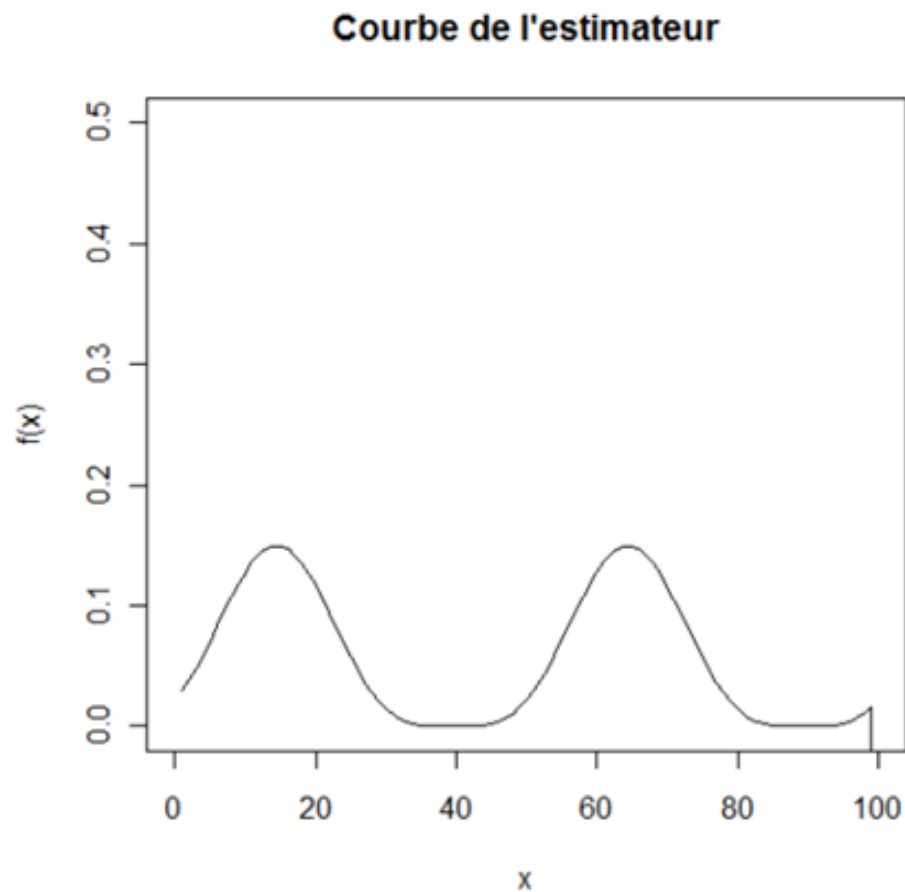


FIGURE 4.3 – Comportement graphique de l'estimateur de densité conditionnelle  $\hat{f}_Y^X$

### 4.5.2 Simulation de la fonction $p$

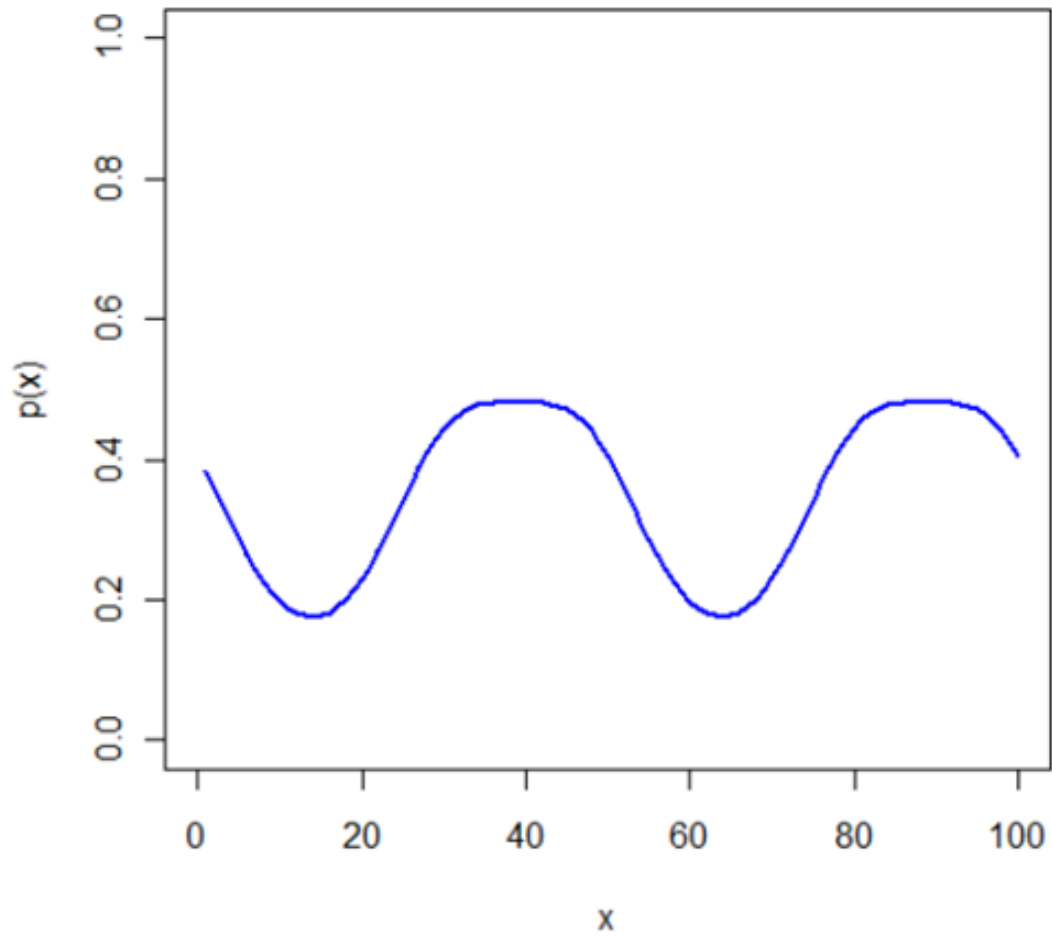


FIGURE 4.4 – Courbe de la fonction  $p$

### 4.5.3 La fonction $p$ et l'estimateur $f$ au même plan

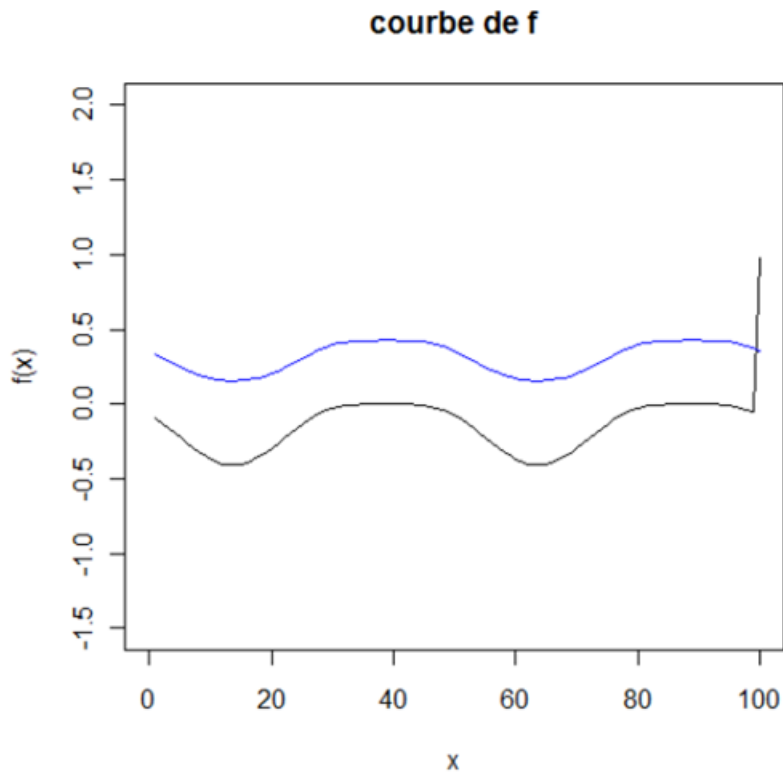


FIGURE 4.5 – Courbe de la fonction  $p$  et l'estimateur de  $f$

### 4.5.4 Résultat final

Après les résultats des simulations, nous avons obtenu les deux courbes sont proches, donc les approximations sont bonnes.

# Conclusion

Dans cet mémoire nous avons présenté d'une part des outils mathématique existants dans la littérature, et qui sont en particulier nécessaire pour la résultat du problème posé par cette dimension infinie.

Nous avons étudié l'estimation non paramétrique de la densité conditionnelle par la méthode de noyau dans le cadre fonctionnelle, où les variables sont dans le cas i.i.d. où nous avons conclu après l'étude théorique et pratique que nous avons fait, que l'estimateur de densité conditionnelle converge presque complète vers les valeurs simulées.

Cet mémoire montre que la convergence presque complète de l'estimateur de densité conditionnelle qui présentée par Ferraty et vieu (2006) est confirmée, et c'est la meilleure approximation.

# Bibliographie

- [1] MECHAB WARDIA CHAHINEZ, "Analyse non paramétrique par régression relative", Sidi bel abbes :these de doctorat, université Sidi bel abbes, 2016/2017.
- [2] F. FERRATY et P. VIEU : Nonparametric functional data analysis : Theory and practice. Springer, 2006.
- [3] G. COLLOMB, W. HÄRDLE et S. HASSANI : A note on prediction via estimation of the conditional mode function. Journal of Statistical Planning and Inference, 15 : 227 – 236, 1987.
- [4] G. COLLOMB, S. HASSANI, P. SARDA et P. VIEU : Estimation non paramétrique de la fonction de hasard pour des observations dépendantes. Statistique et analyse des données, 10(3) : 42 – 49, 1985.
- [5] Bashtannyk, D.M. and Hyndman, R.J. (2001). Bandwidth selection for kernel conditional density estimation. Comput. Statist. Data Anal., 36, Pages 279-298.
- [6] Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. Pub. Inst. Stat. Univ. Paris, 3, Pages 27-42.
- [7] Collomb, G. Estimation nonparamétrique de la régression (in french). PhD Université Paul Sabatier, Toulouse (1976).
- [8] Rahmani Asmaa, "Estimation non paramétrique de la régression par la méthode du noyau", Ouargla :Mémoire de master, université Ouargla, 2019.
- [9] BOUKHAMES Oussama et ZAIDI Mohamed, "Estimation Non Paramétrique De La Densité De Probabilité Et De La Fonction De Régression", Boumerdes :Mémoire de master, université Boumerdes, 2016.
- [10] BELGUERNA Abderrahmane, Eléments de la Statistique Inférentielle, Un polycopié destiné aux étudiants de licence LMD Mathématiques (L3), 2015
- [11] MENACEUR Oussama et AMARA Abdeljalil, "Estimation de la fonction de densité discrète par la méthode du noyau associé discret", Béjaïa :Mémoire de master, université Béjaïa, 2021.

## BIBLIOGRAPHIE

---

- [12] Hsu, P.L. And Robbins, H. (1974). complete Convergence And The Law Of Large Numbers. *proc. Nat. Acad. Sci. Usa* 33 25-31.
- [13] Lagodowski, Z.A. And Rychlik, Z. (1985). complete Convergence And Convergence Rates For Randomly Indexed Sums Of Random Variables With Multidimensional Indices. *bull. Acad. Sci. Polon. Ser. Math.* 32 219-223.
- [14] Szynal, D. (1972). on Almost Complete Convergence For The Sum Of A Random Number Of Independent Random Variables. *bull. Acad. Sci. Polon. Ser. Math. Astronom. Phys.* 20 571-574.
- [15] Hsu, P.L. And Robbins, H. (1974). complete Convergence And The Law Of Large Numbers. *proc. Nat. Acad. Sci. Usa* 33 25-31.
- [16] Gut, A. (1993). complete Convergence And Ces Summation For I.I.D. Random Variables. *probab. Th. Rel. Fields* 97 169-178.
- [17] Mustapha REFAI , "Estimation Fonctionnelle de la Densité Conditionnelle", Mostaganem : Mémoire de master, université Mostaganem, 2018.
- [18] Erdos, P. (1950). remark On My Paper "On A Theorem Of Hsu And Robbins". *ann. Math. Statist.* **21138**.
- [19] Erdos, P. (1949). on A Theorem Of Hsu And Robbins. *ann. Math. Statist.* 20 **286 – 291**
- [20] Asmussen, S. And Kurtz, T.G. (1980). necessary And Sufficient Conditions For Complete Convergence In The Law Of Large Numbers. *ann. Probab.* **8176 – 182**.
- [21] GHETTAB SARAH , "Estimation Non Paramétrique Par La Méthode Du Noyau : Applications à Des Données Censurées", Alger : Mémoire de magister, université Alger, 2014.
- [22] Gut, A. (1993). complete Convergence And Ces Summation For I.I.D. Random Variables. *probab. Th. Rel. Fields* 97 169-178.
- [23] Attouch, M., Laksaci, A., Ould-Saïd, E. (2009). Asymptotic Distribution of Robust Estimator for Functional Nonparametric Models. *Communications in Statistics : Theory and Methods.* **38, 1317 – 1335**.
- [24] Bosq, D., Lecoutre, J. P. (1987). *Théorie de l'estimation fonctionnelle*. ECONOMICA, Paris.
- [25] Burba, F., Ferraty, F., Vieu, P. (2008). Convergence de l'estimateur à noyau des  $k$  plus proches voisins en régression fonctionnelle non-paramétrique. *C. R. Acad. Sci. Paris.* **346, 339 – 342**.
- [26] Carbon, M., Francq, C., Tran, L. T. (2007). Kernel regression estimation for random fields. *J. Statist. Plann. Inference.* 137, 778-798.

## BIBLIOGRAPHIE

---

- [27] Cardot, H., Crambes, C., Sarda, P. (2004). Spline estimation of conditional quantiles for functional covariates, *C. R. Math. Acad. Sci. Paris.* 339, 141-144.
- [28] Collomb, G. (1985). Nonparametric regression : an up to date bibliography *Statistics.* **16**, 309 – 324.
- [29] ZOUTAT Khadidja et FOURAR Nadjette , "Estimation fonctionnelle de la densité conditionnelle avec la methode du noyau", Béjaïa :Mémoire de master, université Béjaïa, 2013.
- [30] Deheuvels, P., Mason, D. (2004). General asymptotic confidence bands based on kernel type function estimators. *Statistical Inference for Stochastic Processes*, 7, 225-277.
- [31] Ferraty, F., Goia, A., Vieu, P. (2002). Functional Nonparametric Model for Time Series : A Fractal Approach for Dimension Reduction. *TEST*, 11, 317-344.
- [32] Ferraty, F., Laksaci, A., Tadj, A., Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference.* **140**, 335 – 352.
- [33] MEDJIR Aouatef , "Estimation de la fonction de densité de probabilité et le différentes Techniques de détection Radar", Blida :Mémoire de master, université Blida, 2020/2021.
- [34] Ferraty, F., Laksaci, A., Tadj, A. Vieu, P. (2011). Kernel regression with functional response. *Electronic Journal of Statistics.* 5, 159-171.
- [35] Ferraty, F. Laksaci, A., Vieu, P. (2005). Functional time series prediction via conditional mode estimation. *C. R. Math. Acad. Sci. Paris.* 340, 389-392.
- [36] Györfi, L., Härdle, W., Sarda, P. et Vieu, P. (1989). Nonparametric curve estimation for time series, *Lecture Notes in Statistics.* 60, Springer- Verlag.
- [37] Laksaci, A., Madani, F., Rachdi, M. (2010). Kernel conditional density estimation when the regressor is valued in a semi metric space. *International Statistical Review.* (In press).
- [38] Li, W. V., Tran, L. T. (2009). Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference.* 139, 164-175.

# Resumé

L'estimation de la régression joue un rôle très important dans la prédiction.

L'objet de mémoire est d'établir une synthèse des résultat limite de l'estimation non paramétrique de la densité conditionnelle pour les variables aléatoires fonctionnelles (c'est-à-dire des variables qui prennent leurs valeurs dans espaces de dimension infinis), et dans le cas i.i.d (indépendant et identiquement distribuée), et en utilise la méthode de noyau.

Dans ce mémoire on a 4 chapitres :

Dans le premier chapitre, nous donnerons quelques notions élémentaires, définitions et exemples dans l'estimation paramétrique et non paramétrique.

Dans le deuxième chapitre, nous donnerons quelques notions élémentaires, définitions de la méthode

de noyau et leur utilisation pour les données fonctionnelles, et l'estimation par cette méthode.

Dans le troisième chapitre, nous donnerons présentation des modèles et l'estimateur à noyau de la densité conditionnelle

Par la suite, Nous présentons dans le dernier chapitre le travail des simulations sont faites pour illustrer les résultats théoriques établis sur la densité de probabilité.

En dernier, Les résultats obtenus sont écrits sous la forme d'une conclusion.

**Mots clé :** non paramétrique, méthode du noyau, paramètre de lissage, la convergence presque complète.

# Abstract

Regression estimation plays a very important role in prediction. The object of the thesis is to establish a synthesis of the limiting results of the nonparametric estimation of the conditional density for the functional random variables (i.e. variables that take their values in infinite dimensional spaces), and in the case i.i.d (independent and identically distributed), and uses the kernel method.

In this dissertation we have 4 chapters :

In the first chapter, we will give some basic notions, definitions and examples in parametric and non-parametric estimation.

In the second chapter, we will give some elementary notions, definitions of the kernel method and their use for functional data, and estimation by this method.

In the third chapter, we will give presentation of the models and the kernel estimator of the conditional density. Thereafter, we present in the last chapter the work of the simulations are made to illustrate the theoretical results established on the density of probability.

Lastly, The results obtained are written in the form of a conclusion.

**Keywords :** nonparametric, kernel method, smoothing parameter, Almost complete convergence

# ملخص

يلعب تقدير الانحدار دورًا مهمًا جدًا في التنبؤ

الهدف من الأطروحة هو إنشاء توليفة من النتائج المقيدة للتقدير اللامعلمي للكثافة الشرطية محدودة للمتغيرات العشوائية الوظيفية (أي المتغيرات التي تأخذ قيمها في مسافات أبعاد غير ، والموزعة بشكل مماثل)، ونستخدم في ذلك طريقة النواة

في هذه المذكرة لدينا 4 فصول

في الفصل الأول ، سنقدم بعض المفاهيم الأساسية والتعريفات والأمثلة في التقدير البارامترية وغير البارامترية

في الفصل الثاني ، سنقدم بعض المفاهيم الأولية، وتعريفات طريقة النواة واستخدامها للبيانات الوظيفية، والتقدير بهذه الطريقة

في الفصل الثالث سوف نقدم النماذج ومقدر النواة للكثافة الشرطية

وفي الفصل الأخير قمنا بعمل المحاكاة لتوضيح النتائج النظرية المحددة على كثافة الاحتمال

أخيرًا ، تتم كتابة النتائج التي تم الحصول عليها في شكل استنتاج