

# Table des matières

<b>Introduction</b>	<b>5</b>
<b>1 Présentation</b>	<b>7</b>
1.1 Notations et définitions . . . . .	7
1.2 Outils . . . . .	8
<b>2 Régression non-paramétrique vectorielle</b>	<b>11</b>
2.1 Modèle . . . . .	11
2.2 Convergence presque complète ponctuelle . . . . .	12
2.2.1 Cas indépendant . . . . .	12
2.2.2 Hypothèses . . . . .	12
2.2.3 Cas dépendant . . . . .	16
2.2.4 Hypothèses . . . . .	16
2.2.5 Résultat . . . . .	16
2.3 Convergence presque complète uniforme . . . . .	20
2.3.1 Cas i.i.d . . . . .	20
2.3.2 Hypothèses . . . . .	20
2.3.3 Cas $\alpha$ -mélange . . . . .	29
2.3.4 Propriétés asymptotiques . . . . .	29
<b>3 Régressoin semi paramétrique fonctionnelle</b>	<b>33</b>
3.1 Modèle . . . . .	33
3.2 Convergence presque complète ponctuelle . . . . .	34
3.2.1 Cas i.i.d . . . . .	34
3.2.2 Cas de mélange forte . . . . .	37
3.2.3 Hypothèses . . . . .	37
3.2.4 Propriétés asymptotiques . . . . .	37

<b>4 Simulation</b>	<b>41</b>
4.1 Algorithme . . . . .	41
4.1.1 Résultats . . . . .	42
<b>Conclusion</b>	<b>45</b>

# Remerciements

*Ce mémoire n'aurait pas été possible sans l'intervention, d'un grand nombre de personnes. Nous souhaitons ici les en remercier.*

*Nous tenons d'abord à remercier très chaleureusement notre encadreur de mémoire de fin d'études mademoiselle Hachemi, pour ses précieux conseils et son orientation tout au long de notre recherche.*

*Les conseils qu'elle nous a prodigué, la patience, la confiance qu'elle nous a témoignés ont été déterminants dans la réalisation de notre travail de recherche.*

*Nos remerciements s'étendent également à tous nos enseignants durant les années des études. A nos familles et nos amis qui par leurs prières et leurs encouragements, on a pu surmonter tous les obstacles.*

*Enfin, nous tenons à remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.*



# Introduction

Etudier les liens entre deux variables aléatoires est une question très importante en statistique. En conséquence, on peut dire qu'en statistique, les outils naturels pour faire la prévision sont les modèles liés à la distribution conditionnelle de ces variables aléatoires. Parmi ces modèles les plus sollicités citons la régression classique, la régression modale (mode conditionnel), les quantiles de régressions (la médiane conditionnelle).

La régression est un ensemble de méthodes statistiques très utilisées pour analyser cette relation. Pendant longtemps, la régression d'une variable aléatoire  $y$  sur le vecteur de variables aléatoires  $x$  désignait la moyenne conditionnelle de  $y$  sachant  $x$ . Aujourd'hui, le terme de régression désigne tout élément de la distribution conditionnelle de  $y$  sachant  $x$  considérée comme une fonction de  $x$ . Dans notre contexte non paramétrique, les premiers résultats ont été obtenus par Tukey ([9], 1961). Tandis que l'estimation par la méthode du noyau a été utilisée pour la première fois en 1964 séparément par Nadaraya et Waston. Les premiers travaux en ce sens eurent essentiellement pour l'objet de chercher à utiliser les estimateurs non paramétrique (linéaire le plus souvent) de nouveaux estimateurs. Notons que, les questions dans les espaces de dimension éventuellement infinie sont particulièrement intéressante. L'étude de modèle de régression d'indice simple a été introduit par par Ichimura ((1993),[6]), il a proposé un estimateur par la méthode des moindres carrés pour l'indice du modèle et a étudié la convergence en probabilité et la normalité asymptotique de l'estimateur construit. Depuis plusieurs décennies nombreux sont les statisticiens qui ont développé des applications permettent le traitement de variables aléatoires fonctionnelles. En effet, Ferraty et al. (2003, [4]) ont obtenus la convergence presque complète lorsque les observations sont indépendants et identiquement distribuées et ont été généralisés par Saidi et al. (2005, [1]). La littérature est très abondante dans ce cas par Bosq (2000, [2]), Ramsay et Silverman (2002, [8]), Ferraty et Vieu (2004, [3]). Cette thèse est organisée de la manière suivante :

Le premier chapitre est consacré à la présentation des notations et outils techniques de nos résultats. Puis, cette travail se divise en trois parties.

La première partie s'intéresse à une variable de réponse réelle donnée une variable aléatoire vectorielle dans le cas où les observations sont indépendantes et identiquement distribuées, on établit sous des conditions générales la convergence presque complète de la fonction de régression et sous des hypothèses un peu plus restrictives dans le cas mélange fort qui donne la même vitesse de convergence. On trouve aussi la convergence uniforme.

Dans la deuxième partie, nous nous intéressons à la situation où la variable d'explicative est fonctionnelle. On généralise les derniers résultats lorsqu'on suppose que l'indice du modèle est fixé en quantifiant les propriétés asymptotiques d'un estimateur à noyau pour la fonction de lien.

Enfin, une troisième partie propose par des données simulées une comparaison entre les trois méthodes de régression, la régression linéaire, non-paramétrique et la régression semi-paramétrique.

# Chapitre 1

## Présentation

### 1.1 Notations et définitions

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité et  $\{\Delta_i\}_{i \in \mathcal{Z}}$  une famille des variables aléatoires définie sur  $(\Omega, \mathcal{A}, \mathbb{P})$  é valeurs dans un espace probabilisable  $(\mathbb{E}, \xi)$ . On note  $(\sigma_i^j)_{i \neq j}$  dans  $\mathbb{Z} \cup \{-\infty, +\infty\}$ , la tribu engendrée par  $\{\Delta_k, i < k < j\}$  et par  $L_2(\sigma_i^j)$  l'espace des variables aléatoires  $\sigma_i^j$ -mesurable et de carrée sommable.

**Définition 1.1.1** Soit  $\{\Delta_i\}_{i \in \mathcal{Z}}$  une famille des variables aléatoires définie sur  $(\Omega, \mathcal{A}, \mathbb{P})$  é valeurs dans un espace probabilisable  $(\mathbb{E}, \xi)$ . On dit que la famille  $\{\Delta_i, i \in \mathcal{Z}\}$  est  $\alpha$ -mélangeant si la suite

$$\alpha_n = \sup_{\{k \in \mathcal{Z}, A \in \sigma_{-\infty}^k, B \in \sigma_{n+k}^{+\infty}\}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$$

tend vers 0 quand  $n$  tend vers l'infinie. La suite  $\alpha_n$  est appelée coefficient de mélange forte.

**Définition 1.1.2** Une fonction  $K$  de  $\mathbb{R}^p$  dans  $\mathbb{R}$  est dite noyau d'ordre  $k$ ,  $k \in \mathbb{N}^*$ , si :

$$T_{i_1, \dots, i_p}(K) = 0, \forall (i_1, \dots, i_p) \in \mathbb{R}_*^p \text{ vérifiant } i_j < k, 1 \leq j \leq p.$$

et

$$T_j(K) \neq 0, \forall j \leq k.$$

oé

$$T_{i_1, \dots, i_p}(K) = \int_{\mathbb{R}} u_1^{i_1}, \dots, u_p^{i_p} K(u_1, \dots, u_p) du_1, \dots, du_p.$$

et

$$T_j(K) = \int_{\mathbb{R}} u_j^k K(u_1, \dots, u_p) du_1, \dots, du_p.$$

## 1.2 Outils

**Lemme 1.2.1** [3] Soit  $\Delta_1, \dots, \Delta_n$  des variables aléatoires réelles centrées, indépendantes et identiquement distribuées, telles qu'il existe deux réels positifs  $d$  et  $\delta$  vérifiant :

$$|\Delta_1| \leq d \text{ et } E\Delta_1^2 \leq \delta^2.$$

Alors, pour tout  $\epsilon \in ]0, \frac{\delta^2}{d}[$  on a

$$p \left[ n^{-1} \left| \sum_{i=1}^n \Delta_i \right| > \epsilon \right] \leq 2e^{-\frac{n\epsilon^2}{4\delta^2}}.$$

cette inégalité a été donnée par W.Hoeffding en (1963, [3]). Les lemmes suivants donnent les deux version de l'inégalité de Fuk Nagaev.

**Lemme 1.2.2** [3] Soit  $\{\Delta_i, i \in \mathbb{N}\}$  une suite des variables aléatoires réelles  $\alpha$ -mélangeante, de coefficient de mélange  $\alpha_n$  vérifiant :

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \quad \alpha(n) \leq cn^{-a}$$

et si  $\forall i, \|\Delta_i\|_{\infty} < \infty$ , alors, pour tout  $\epsilon > 0$  et  $r > 0$ , on a

$$\mathbb{P} \left[ \left| \sum_{k=1}^n \Delta_k \right| > 4\epsilon \right] \leq 4 \left( 1 + \frac{\epsilon^2}{rS_n^2} \right)^{\frac{-r}{2}} + 2ncr^{-1} \left( \frac{2r}{\epsilon} \right)^{a+1}. \quad (1.1)$$

**Lemme 1.2.3** [1] Soit  $\{\Delta_i, i \in \mathbb{N}\}$  une suite des variables aléatoires réelles  $\alpha$ -mélangeante, de coefficient de mélange  $\alpha_n$  vérifiant :

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \quad \alpha(n) \leq cn^{-a}$$



et si  $\|\Delta_i\|_\infty < \infty, \forall i$ , alors, pour tout  $\lambda > 0$  et  $r > 1$ , on a

$$\begin{aligned} \mathbb{P} \left[ \left| \sum_{i=1}^n \Delta_i \right| > 4\lambda n \mathbb{E}K_1(x) \right] &= \mathbb{P} [|s_n| > 4\lambda n \mathbb{E}K_1(x)] \\ &\leq C \left( 1 + \frac{\lambda^2 n^2 (\mathbb{E}K_1(x))^2}{r s_{n,l}} \right)^{\frac{-r}{2}} \\ &\quad + C \frac{n}{r} \left( \frac{r}{\lambda n \mathbb{E}K_1(x)} \right)^{\frac{p(\alpha+1)}{a+p}}. \end{aligned}$$

où

$$s_n = \sum_{i=1}^n \Delta_i$$

Pour calculer l'expression de  $S_n^2$ , définie dans le lemme précédent, on utilise le lemme suivant :

**Lemme 1.2.4** [1] Soit  $\{\Delta_i, i \in \mathbb{N}\}$  une suite des variables aléatoires réelles  $\alpha$ -mélangeante, de coefficient de mélange  $\alpha_n$  vérifiant :

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \quad \alpha(n) \leq cn^{-a}$$

et si  $\|\Delta_i\|_\infty < \infty, \forall i$ , alors, pour tout  $\epsilon > 0$  et  $r > 0$

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\text{cov}(\Delta_i, \Delta_j)|.$$

pour calculer l'expression de  $S_n^2$ , définie dans le lemme précédent, on utilise le lemme suivant :

**Lemme 1.2.5** ([1], [3]) Soit  $\{\Delta_i, i \in \mathbb{N}\}$  une suite des variables aléatoires réelle  $\alpha$ -mélangeante, de coefficient des mélange  $\alpha_n$ , telle que  $\|\Delta_i\|_\infty < \infty, \forall i$ . On a pour tout  $i \neq j$  :

$$|\text{cov}(\Delta_i, \Delta_j)| \leq 4\|\Delta_i\|_\infty \|\Delta_j\|_\infty \alpha_{|i-j|}.$$

et dans le cas fonctionnelle on introduit

$$|\text{cov}(\Delta_i, \Delta_j)| \leq C(\alpha_{|i-j|})^{\frac{(p-2)}{p}}.$$



## Chapitre 2

# Régression non-paramétrique vectorielle

Dans ce chapitre, on s'intéresse à un modèle de régression non paramétrique dans le cas où la variable explicative est multi-dimensionnelle. Ce chapitre est dévisé en deux sections. Dans la première section, on présente notre modèle et son estimateur. La deuxième section traite la convergence presque complète ponctuelle sur un compact fixé.

### 2.1 Modèle

Soit  $(X_i, Y_i)_{i=1, \dots, n}$  un n-échantillon de variables aléatoire définie sur l'espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  à valeurs dans  $\mathbb{R}^{p+1}$ ,  $p$  est un entier strictement positif. En outre, nous supposons qu'il existe une version régulière de la probabilité conditionnelle de  $Y$  donné  $X$ , ce qui à une densité par rapport à la mesure de Lebesgues noté  $f$ . Le modèle de régression étudié est donné par l'équation

$$Y = r(X) + \varepsilon$$

où  $\varepsilon$  est une variable aléatoire réelle centrée est indépendante de  $X$  et  $r$  est une fonction de  $\mathbb{R}^p$  à valeur dans  $\mathbb{R}$  supposée de classe  $C^k$ . La fonction  $r$  peut être exprimé par

$$\mathbb{E}(Y/X) = r(x)$$

$$\hat{r}_{N,W}(x) = \frac{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)}, \quad x \in \mathbb{R}^p$$

où  $K$  est une fonction de  $\mathbb{R}^p$  à valeurs dans  $\mathbb{R}$  et  $h$  est suite de nombres réel positifs tel que  $(h_n = h)$  s'appelle le paramètre de lissage.

## 2.2 Convergence presque complète ponctuelle

On se propose d'étudier la convergence presque complète de l'estimateur  $\hat{r}(x)$  dans les deux cas indépendant et dépendant.

### 2.2.1 Cas indépendant

Nous allons commencer par donner des résultats de convergence presque complète sous des modèles de régularité du type :

### 2.2.2 Hypothèses

- (H1) Les fonctions  $r$  et  $f$  sont  $k$ -fois continûment dérivable au voisinage de  $x$
- (H2) la densité  $f$  de la variable explicative est strictement positive au point  $x$ .
- (H3) Le paramètre de lissage est tel que :

$$\lim_{n \rightarrow \infty} h_n = 0, \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{nh^p}{\log n} = \infty$$

- (H4) le noyau  $K$  est d'ordre  $k$ , supposé borné et à support compact.
- (H5) La variable de réponse est telle que :  $|Y| < M < \infty$ .

**Théorème 2.2.1** *Sous les les hypothèses (H1) – (H5), on a*

$$\hat{r}(x) - r(x) = o(h^k) + o\left(\sqrt{\frac{\log(n)}{nh^p}}\right), \quad p.co \tag{2.1}$$

**preuve** On a

$$\widehat{r}(x) = \frac{\widehat{g}(x)}{\widehat{f}(x)}$$

où

$$\widehat{g}(x) = \frac{1}{nh^p} \sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h}\right)$$

et

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

la preuve est basé sur la decomposition suivante :

$$\begin{aligned} \widehat{r}(x) - r(x) &= \frac{1}{\widehat{f}(x)} (\widehat{g}(x)) - \mathbb{E}(\widehat{g}(x)) + \mathbb{E}(\widehat{g}(x)) - g(x) \\ &+ \frac{r(x)}{\widehat{f}(x)} (\widehat{f}(x) - f(x)) \end{aligned}$$

**Lemme 2.2.1** *Sous les hypothèses  $(H_1), (H_3), (H_4), (H_5)$ , on a*

$$\mathbb{E}(\widehat{g}(x)) - g(x) = o(h^k) \tag{2.2}$$

$$\mathbb{E}(\widehat{f}(x)) - f(x) = o(h^k) \tag{2.3}$$

**Lemme 2.2.2** *Sous les hypothèses  $(H_1), (H_3), (H_4), (H_5)$*

$$\mathbb{E}(\widehat{g}(x)) - \widehat{g}(x) = o\left(\sqrt{\frac{\log n}{nh^p}}\right) \tag{2.4}$$

$$\mathbb{E}(\widehat{f}(x)) - \widehat{f}(x) = o\left(\sqrt{\frac{\log n}{nh^p}}\right) \tag{2.5}$$

**Lemme 2.2.3** *Sous les conditions du lemme (2.2.2)*

$$\exists \delta > 0, \quad \sum_{i=1}^{\infty} \mathbb{P}(|\widehat{f}(x)| \leq \delta) < \infty \tag{2.6}$$

**Démonstration du lemme 2.2.1** On a, par équistribution des observations,

$$\begin{aligned}
\mathbb{E}(\widehat{g}(x)) &= \mathbb{E} \left( \frac{1}{nh^p} \sum_{i=1}^n Y_i K \left( \frac{x - X_i}{h} \right) \right) - g(x) \\
&= \frac{1}{nh^p} \sum_{i=1}^n \mathbb{E} \left( Y_i K \left( \frac{x - X_i}{h} \right) \right) - g(x) \\
&= \frac{1}{h^p} \left( \mathbb{E} \left( Y_1 K \left( \frac{x - X_1}{h} \right) \right) - h^p g(x) \right) \\
&= \frac{1}{h^p} \left( \int_{\mathbb{R}^p} r(u) K \left( \frac{x - u}{h} \right) f(u) du - h^p g(x) \right) \\
&= \frac{1}{h^p} \left( \int_{\mathbb{R}^p} g(u) K \left( \frac{x - u}{h} \right) du - h^p g(x) \right)
\end{aligned}$$

En posant  $(x - u)/h = z$ , tel que,  $du = h^p dz$  dans le calcul d'intégrale à l'expression :

$$\begin{aligned}
\mathbb{E}(\widehat{g}(x)) - g(x) &= \int_{\mathbb{R}^p} [g(x - hz) - g(x)] K(z_1 \cdots z_p) dz_1 \cdots dz_p \\
&= \int_{\mathbb{R}^p} [g(x - hz) - g(x)] K(z) dz
\end{aligned}$$

Il suffit alors de développer la fonction  $g$  au voisinage du point  $x$ , on obtient dans ce cas,

$$g(x - hz) - g(x) = \sum_{j=1}^k \frac{(-h)^j}{j!} \sum_{i_1 \cdots i_p = j} \left( (z_1^{i_1} \cdots z_p^{i_p}) \left( \frac{\partial^j g}{\partial x_1^{i_1} \cdots \partial x_p^{i_p}} \right) (x) + o(h^j) \right)$$

D'après l'hypothese( $H_4$ ), on a

$$\mathbb{E}(\widehat{g}(x)) - g(x) = \frac{(-h)^k}{k!} \sum_{j=1}^p \left( \frac{\partial^k g}{\partial x_j^k} \right) (x) T_k(j) + O(h^k)$$

Ceci achever la preuve de (2.2).

la preuve de (2.3) suit les mêmes étapes de (2.2). En fait on établit un résultat plus précis que celui recherché, à savoir que

$$\mathbb{E}(\widehat{f}(x)) - f(x) = \frac{(-h)^k}{k!} \sum_{j=1}^p \left( \frac{\partial^k f}{\partial x_j^k} \right) (x) T_k(j) + O(h^k)$$

**Démonstration du lemme 2.2.2**

on pose

$$\Delta_i = \frac{1}{h^p} \left[ Y_i K\left(\frac{x - X_i}{h}\right) - \mathbb{E} \left( Y_i K\left(\frac{x - X_i}{h}\right) \right) \right]$$

Pour appliqué l'inégalité de Hoeffding (lemme (1.2.1)) aux variables  $\Delta_i$ , il faut majorer les deux termes :  $|\Delta_i|$  et  $\mathbb{E}\Delta_i^2$ . En effet, d'après les hypothèses  $(H_4)$ ,  $(H_6)$  on a,

$$|\Delta_i| \leq \frac{C}{h^p}, \quad \text{et} \quad \mathbb{E}\Delta_i^2 \leq \frac{C}{h^p}$$

Le Lemme (1.2.1) permet alors d'écrire que pour tout  $\epsilon$ , on a :

$$\begin{aligned} \mathbb{P} \left( |\mathbb{E}(\widehat{g}(x)) - \widehat{g}(x)| > \epsilon_0 \sqrt{\frac{\log n}{nh^p}} \right) &= 2 \exp\left(\frac{-n\epsilon_0^2 \frac{\log n}{nh^p}}{4\frac{C}{h^p}}\right) \\ &\leq 2n^{-\epsilon_0^2 C} \end{aligned}$$

donc

$$\sum_{i=1}^n \mathbb{P} \left( |\mathbb{E}(\widehat{g}(x)) - \widehat{g}(x)| > \epsilon_0 \sqrt{\frac{\log n}{nh^p}} \right) < +\infty$$

Finallement on a ;

$$\mathbb{E}(\widehat{g}(x)) - \widehat{g}(x) = o\left(\sqrt{\frac{\log n}{nh^p}}\right) \quad p.co$$

la preuve de (2.5) se démontre de la même manière que (2.4)

**Démonstration du lemme 2.2.3**

on a,

$$\mathbb{E}(\widehat{f}(x)) - 1 = 0$$

alors,

$$|\widehat{f}(x)| \leq \frac{f(x)}{2} \implies |\widehat{f}(x) - f(x)| > \frac{f(x)}{2}$$

on peut écrire

$$\sum_{i=1}^{\infty} \mathbb{P} \left( |\hat{f}(x)| < \frac{f(x)}{2} \right) \leq \sum_{i=1}^{\infty} \mathbb{P} \left( |\hat{f}(x) - f(x)| > \frac{f(x)}{2} \right) < \infty$$

Ainsi, le résultat est une conséquence du lemme précédente.

### 2.2.3 Cas dépendant

le but de cette section est la généralisation du résultat donné dans le cas précédent à des observations mélangées.

### 2.2.4 Hypothèses

On garde les mêmes hypothèses du cas i.i.d et on ajouté les conditions ci-dessous permettent de trouver la même vitesse de convergence.

- $(H_1)$  le couple des variables aléatoires  $(X_i, X_j)$  admet une densité notée  $f_{i,j}$ , pour tout  $i \neq j$
- $(H_2)$  Le coefficient des mélange est tel que

$$\exists c \in \mathbb{R}^{**}, \quad \text{et} \quad a \in \mathbb{R}^{**} \text{telles que} \quad \alpha(n) \leq cn^{-a}$$

- $(H_3)$  Le paramètre de lissage est vérifié pour  $a > (5 + \sqrt{17})/2$  tel que

$$\exists \theta > 0, \exists c_1 > 0, \exists c_2 > 0, \quad c_2 n^{(\theta + \frac{3-a}{a+1})} \leq h^p \leq c_1 n^{(\frac{1}{1-a} - \theta)}$$

### 2.2.5 Résultat

**Théorème 2.2.2** *Si les conditions  $(H1) - (H4), (H_1) - (H_3)$  sont réalisées alors*

$$\hat{r}(x) - r(x) = o(h^k) + o\left(\sqrt{\frac{\log(n)}{nh^p}}\right), \quad P.co$$

**Preuve :**

on obtient le même résultat pour cela on va démontré seulement la partie de dispersion



**Lemme 2.2.4** *Si les conditions  $(H_2)$ ,  $(H_3)$  sont vérifiées alors il existe  $\exists v > 0$  et  $\epsilon > 0$  tel que l'on ait :*

$$\mathbb{P} \left( \left| \mathbb{E}(\widehat{f}(x)) - \widehat{f}(x) \right| > \epsilon_0 \sqrt{\frac{\log n}{nh^p}} \right) = o(n^{-1-v})$$

$$\mathbb{P} \left( \left| \mathbb{E}(\widehat{g}(x)) - \widehat{g}(x) \right| > \epsilon_0 \sqrt{\frac{\log n}{nh^p}} \right) = o(n^{-1-v})$$

**Preuve 2.2.1** *Nous ferons les deux preuves simultanément en posant*

$$\widehat{\psi}(x) = \widehat{f}(x) \text{ ou } \widehat{g}(x)$$

et

$$\Delta_i = Y_i^l K \left( \frac{x - X_i}{h} \right) - \mathbb{E} \left( Y_i^l K \left( \frac{x - X_i}{h} \right) \right), \quad l = 0 \text{ ou } l = 1$$

Il vient pour tout  $\epsilon > 0$  et pour  $r > 1$  que

$$\begin{aligned} \mathbb{P} \left( \left| \mathbb{E}(\widehat{\psi}(x)) - \widehat{\psi}(x) \right| > \epsilon \right) &= \mathbb{P} \left( \frac{1}{nh^p} \left| \sum_{i=1}^n \Delta_i \right| > \epsilon \right) \\ &= \mathbb{P} \left( \left| \sum_{i=1}^n \Delta_i \right| > \frac{4(nh^p\epsilon)}{4} \right) \\ &\leq 4 \left( 1 + \frac{\epsilon^2 nh^{2p}}{16rS_n^2} \right)^{\frac{-r}{2}} + 2ncr^{-1} \left( \frac{8r}{\epsilon nh^p} \right)^{a+1} \end{aligned}$$

On commence d'abord, par la quantité

$$\begin{aligned} S_n^2 &= \sum_{i,j}^n |\text{cov}(\Delta_i, \Delta_j)| \\ &= \sum_{i=1}^n \sum_{i \neq j} |\text{cov}(\Delta_i, \Delta_j)| + n \text{Var}(\Delta_i) \\ &= S_n^{2*} + n \text{Var}(\Delta_i) \end{aligned}$$

En effet, par définition de  $\Delta_i$  on a

$$\text{cov}(\Delta_i, \Delta_j) = \mathbb{E}(\Delta_i \Delta_j).$$

D'une part,

$$\begin{aligned} \mathbb{E}(\Delta_i \Delta_j) &= E \left[ \left( Y_i^l K \left( \frac{x-X_i}{h} \right) - \mathbb{E} Y_i^l K \left( \frac{x-X_i}{h} \right) \right) \left( Y_j^l K \left( \frac{x-X_j}{h} \right) - \mathbb{E} Y_j^l K \left( \frac{x-X_j}{h} \right) \right) \right] \\ &= \mathbb{E} \left[ \left( Y_i^l K \left( \frac{x-X_i}{h} \right) \right) \left( Y_j^l K \left( \frac{x-X_j}{h} \right) \right) \right] \\ &\quad - \mathbb{E} \left[ Y_i^l K \left( \frac{x-X_i}{h} \right) \right] \mathbb{E} \left[ Y_j^l K \left( \frac{x-X_j}{h} \right) \right] \\ &= c \mathbb{E} \left[ K \left( \frac{x-X_i}{h} \right) K \left( \frac{x-X_j}{h} \right) \right] \\ &\quad - c \mathbb{E} \left[ K \left( \frac{x-X_i}{h} \right) \right] E \left[ K \left( \frac{x-X_j}{h} \right) \right] \\ &= c \int \int K \left( \frac{x-u}{h} \right) K \left( \frac{x-v}{h} \right) f_{ij}(u, v) dudv \\ &\quad - c \int \int K \left( \frac{x-u}{h} \right) K \left( \frac{x-v}{h} \right) f(u) f(v) dudv \\ &= c \int \int K \left( \frac{x-u}{h} \right) K \left( \frac{x-v}{h} \right) [f_{ij}(u, v) - f(u) f(v)] dudv. \end{aligned}$$

On prend le changement des variables usuel  $\frac{x-u}{h} = z, \frac{x-v}{h} = t$ , d'où,

$$|\mathbb{E}(\Delta_i \Delta_j)| \leq ch^{2p} \int \int |K(z)K(t) [f_{ij}(x-hz, x-ht) - f(x-zt)f(x-ht)]| dzdt.$$

Pour  $n$  assez grand, on peut trouver une constante  $C$  telle que

$$|\mathbb{E}(\Delta_i \Delta_j)| \leq ch^{2p} \int \int |K(z)K(t) [f_{ij}(x, x) - f^2(x)]| dzdt.$$

Puisque les fonction  $K$ ,  $f$  et  $f_{ij}$  sont bornées, alors

$$|\mathbb{E}(\Delta_i \Delta_j)| = O(h^{2p}). \quad (2.7)$$

D'autre part, on peut majorer cette covariance en utilisant l'inégalité du lemme (1.2.5, [3])

$$|\text{cov}(\Delta_i \Delta_j)| \leq 4 \|\Delta_i\|_\infty \cdot \|\Delta_j\|_\infty \alpha(|i-j|). \quad (2.8)$$

Par la suite, on fait appelle aux techniques de Masry ([7],1986), en considérant une suite  $u_n$  des entiers naturels et on montre que à l'aide des majorations précédentes

$$S_n^{2*} \leq c \left[ \sum_{0 \leq |i-j| \leq u_n} \sum h^{2p} + \sum_{|i-j| u_n} \alpha(|i-j|) \right] = O(h^{2p} n u_n + n^2 \alpha(u_n)). \quad (2.9)$$

Pour la suite  $u_n$ , on prend  $u_n = \left[ \frac{1}{h^p \log n} \right]$  ceci implique

$$\begin{aligned} S_n^{2*} &= O\left(h^{2p} n \cdot \frac{1}{h^p \log n} + n^2 \alpha(h^p \log n)^{-1}\right) \\ &= O\left(\frac{nh^p}{\log n}\right) + O\left(n^2 \alpha(h^p \log n)^{-1}\right). \end{aligned}$$

Il est évident que  $\lim_{n \rightarrow \infty} \frac{\frac{nh^p}{\log n}}{nh^p} = 0$  c'est à dire

$$S_n^{2*} = o(nh^p) + O\left(n^2 \alpha(h^p \log n)^{-1}\right).$$

D'après la deuxième partie de la condition  $(H_3)$  on a

$$\begin{aligned} h^p \leq c_1 n^{\frac{1}{1-a}-\theta} &\Rightarrow h^{p(a-1)} \leq c_1^{(a-1)} n^{(\frac{1}{1-a}-\theta)(a-1)} \\ &\Rightarrow h^{p(a-1)} \leq c_1^{(a-1)} n^{-1-\theta(a-1)}. \end{aligned}$$

Posant  $\varepsilon = \theta(a-1)$  alors  $h^{p(a-1)} = O(n^{-1-\varepsilon})$ . Par conséquent, en vertu du lemme (1.2.4, [3]),

$$S_n^{2*} = o(nh^p).$$

Donc,

$$S_n^2 = o(nh^p)$$

Pour  $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{nh_n}}$

$$\begin{aligned} \mathbb{P} \left[ |\mathbb{E}\hat{\psi}(x) - \hat{\psi}(x)| > \epsilon_0 \sqrt{\frac{\log n}{nh^p}} \right] \\ \leq 4 \left( 1 + \frac{\epsilon_0^2 (\log n) n^2 h^{2p}}{n^2 h^{2p} 16r} \right)^{\frac{-r}{2}} + 2ncr^{-1} \left( \frac{8r(nh^p)^{\frac{1}{2}}}{\epsilon_0 nh^p (\log n)^{\frac{-1}{2}}} \right) \end{aligned}$$

$$\leq 4 \left( 1 + \frac{\epsilon_0^2 \log n}{16r} \right)^{\frac{-r}{2}} + 2ncr^{-1} \left( \frac{8r^{a+1}}{\epsilon_0} \right) (nh^p \log n)^{-\left(\frac{a+1}{2}\right)}.$$

Pour un choix de  $r = C (\log(n))^2$ , le premier terme est de type

$$A_1 = \left( 1 + \frac{\epsilon_0^2 \log n}{16r} \right)^{\frac{-r}{2}} \leq c \exp \frac{-\epsilon_0^2 \log n}{32} \quad (2.10)$$

et le deuxième terme

$$A_2 \leq c \varepsilon^{-(a+1)} n^{-(\frac{a+1}{2})+1+ab} h^{-p(\frac{a+1}{2})}$$

sous la condition  $(H_3)$  sur le paramètre de lissage  $h_n$  on a

$$A_2 \leq cn^{-1-\frac{a}{2}[-2b+p\theta]} \quad (2.11)$$

Les équations (2.10) et (2.11) implique que

$$\sum_{n=1} \mathbb{P} \left[ |\mathbb{E} \hat{\psi}(x) - \hat{\psi}(x)| > \varepsilon \sqrt{\frac{\log n}{nh^p}} \right] < \infty.$$

## 2.3 Convergence presque complète uniforme

L'obtention de résultats analogues uniformes sur un compact  $S$  de  $\mathbb{R}^p$  est subordonnée à des hypothèses supplémentaires, en considérons les deux cas d'observations indépendant au fortement mélangeants.

### 2.3.1 Cas i.i.d

Pour établir la convergence presque complète, on a besoin des hypothèses suivants :

### 2.3.2 Hypothèses

$H_1$   $r$  et  $f$  sont  $k$  fois continument différentiables au voisinage de  $x$

$H_2$  la densité  $f$  est strictement positive sur le compact  $S$  vérifient

$$\exists \theta > 0, \quad \inf_{x \in S} f(x) > \theta$$

$H_3$  sur le noyau

$$\exists \beta > 0, \exists C < \infty, \forall x \in S, \forall y \in S, |K(x) - K(y)| \leq C|x - y|$$

$H_4$  le paramètre de lissage est tel que :

$$\alpha > 0, \lim_{n \rightarrow \infty} n^\alpha h^{(p+2)} = \infty$$

$H_5$  le noyau  $K$  est intégrable, borné et à support compact

$H_6$  sur la variable de réponse

$$|Y| < M < \infty$$

**Théorème 2.3.1** *Si les conditions  $(H_1) \dots (H_6)$  sont réalisées on a*

$$\sup_{x \in S} |\hat{r}(x) - r(x)| = o(h^k) + o\left(\sqrt{\frac{\log(n)}{nh^p}}\right), \quad p.co$$

**Démonstration** la démonstration de ce théorème est basé sur la décomposition suivante :

$$\begin{aligned} \sup_{x \in S} |\hat{r}(x) - r(x)| &\leq \frac{\sup_{x \in S} |\hat{g}(x) - \mathbb{E}(\hat{g}(x))|}{\inf_{x \in S} |\hat{f}(x)|} + \frac{\sup_{x \in S} |\mathbb{E}(\hat{g}(x) - g(x))|}{\inf_{x \in S} |\hat{f}(x)|} \\ &+ \frac{\sup_{x \in S} |r(x)|}{\inf_{x \in S} |\hat{f}(x)|} \left\{ \sup_{x \in S} |\hat{f}(x) - \mathbb{E}(\hat{f}(x))| + \sup_{x \in S} |\mathbb{E}(\hat{f}(x) - f(x))| \right\} \end{aligned}$$

la preuve se repose les lemmes suivants :

**Lemme 2.3.1** *Sous les hypothèses  $(H_1) - (H_3), (H_4), (H_6)$*

$$\sup_{x \in S} |\mathbb{E}(\hat{g}(x)) - g(x)| = o(h^k) \quad (2.12)$$

$$\sup_{x \in S} |\mathbb{E}(\hat{f}(x)) - f(x)| = o(h^k) \quad (2.13)$$

**Lemme 2.3.2** *Sous les hypothèses  $(H_1) - (H_6)$*

$$\sup_{x \in S} |\mathbb{E}(\widehat{g}(x)) - \widehat{g}(x)| = o\left(\sqrt{\frac{\log n}{nh^p}}\right) \quad (2.14)$$

$$\sup_{x \in S} |\mathbb{E}(\widehat{f}(x)) - \widehat{f}(x)| = o\left(\sqrt{\frac{\log n}{nh^p}}\right) \quad (2.15)$$

et

$$\exists \delta > 0, \quad \sum_{i=1}^{\infty} \mathbb{P}(\inf_{x \in S} \|\widehat{f}(x)\| \leq \delta) < \infty \quad (2.16)$$

### Démonstration du lemme (2.3.1)

Les observations sont indépendant identiquement distribuées,

$$\begin{aligned} \mathbb{E}\widehat{g}(x) &= \mathbb{E}\left[\frac{1}{nh^p} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) Y_i\right] \\ &= \frac{1}{h^p} \mathbb{E}\left[K\left(\frac{x - X_1}{h}\right) Y_1\right] \\ &= \frac{1}{h^p} \mathbb{E}\left[K\left(\frac{x - X_1}{h}\right) \mathbb{E}(Y/X_1)\right] \\ &= \frac{1}{h^p} \mathbb{E}\left[K\left(\frac{x - X_1}{h}\right) r(X_1)\right] \\ &= \frac{1}{h^p} \int K\left(\frac{x - u}{h}\right) r(u) f(u) du. \end{aligned}$$

On considère le changement des variables suivant :  $\frac{x - u}{h} = y$ , alors,

$$\mathbb{E}\widehat{g}(x) = \int K(y) g(x - hy) dy$$

La fonction  $g$  étant de  $C^k$ , alors, en utilisant le développement de Taylor pour obtenir

$$\begin{aligned}
g(x - hy) - g(x) &= -h \sum_{j_1=1}^p \frac{\partial g(x)}{\partial x_{j_1}} y_{j_1} + \frac{h^2}{2} \sum_{j_2=1}^p \sum_{j_1=1}^p \frac{\partial^2 g(x)}{\partial x_{j_1} \partial x_{j_2}} y_{j_1} y_{j_2} + \dots \\
&\quad + \frac{(-h)^k}{k!} \sum_{j_k=1}^p \dots \sum_{j_1=1}^p \frac{\partial^k g(x)}{\partial x_{j_1} \dots \partial x_{j_k}} y_{j_1} \dots y_{j_k} + o(h^k) \\
&= -h \sum_{i_1+i_2+\dots+i_p=j=1} \frac{\partial^j g(x)}{\partial x_1^{i_1} \partial x_2^{i_2} \dots \partial x_p^{i_p}} y_1^{i_1}, \dots, y_p^{i_p} \\
&\quad + \frac{h^2}{2} \sum_{j=i_1+\dots+i_p=2} \frac{\partial^j g(x)}{\partial x_1^{i_1} \partial x_2^{i_2} \dots \partial x_p^{i_p}} y_1^{i_1}, \dots, y_p^{i_p} + \dots \\
&\quad + \frac{(-h)^k}{k!} \sum_{j=i_1+\dots+i_p=k} \frac{\partial^j g(x)}{\partial x_1^{i_1} \dots \partial x_p^{i_p}} y_1^{i_1}, \dots, y_p^{i_p} + o(h^k) \\
&= \sum_{j=1}^k \frac{(-h)^j}{j!} \sum_{j=i_1+\dots+i_p} \frac{\partial^j g(x)}{\partial x_1^{i_1} \dots \partial x_p^{i_p}} y_1^{i_1}, \dots, y_p^{i_p} + o(h^k), \quad \forall x \in S
\end{aligned}$$

La fonction  $\partial^k g$  est bornée sur le compact  $S$ , alors, on peut écrire

$$\begin{aligned}
&\int K(y_1 \dots y_p) g(x - hy) dy - \int g(x) K(y_1 \dots y_p) dy = \\
&\quad \sum_{j=1}^k \frac{(-h)^j}{j!} \sum_{j=i_1+\dots+i_p} \frac{\partial^j g(x)}{\partial x_1^{i_1} \dots \partial x_p^{i_p}} \int K(y_1 \dots y_p) y_1^{i_1}, \dots, y_p^{i_p} dy + o(h^k). \\
\mathbb{E} \widehat{g}(x) - g(x) &= \sum_{j=1}^k \frac{(-h)^j}{j!} \left[ \sum_{i=1}^p \frac{\partial^j g(x)}{\partial x_i^j} \right] \int (y_1^{i_1}, \dots, y_p^{i_p}) K(y_1 \dots y_p) dz + o(h^k) \\
&= \frac{(-h)^k}{k!} \left[ \frac{\partial^k g(x)}{\partial x_1^k} \int y_1^k K(y_1 \dots y_p) dy + \frac{\partial^k g(x)}{\partial x_2^k} \int y_2^k K(y_1 \dots y_p) dy \right] \\
&\quad + \dots \frac{(-h)^k}{k!} \left[ \frac{\partial^k g(x)}{\partial x_p^k} \int y_p^k K(y_1 \dots y_p) dz \right] + o(h^k) \\
&= \frac{(-h)^k}{k!} \sum_{j=1}^p \left[ \frac{\partial^k g(x)}{\partial x_j^k} \right] T_K(j) + o(h^k).
\end{aligned}$$

Sous l'hypothèse  $(H_3)$  on obtient,

$$\sup_{x \in S} |\mathbb{E} \widehat{g}(x) - g(x)| = O(h^k).$$

D'une manière analogie on démontre que,

$$\sup_{x \in S} \left| \mathbb{E} \widehat{f}(x) - f(x) \right| = o(h^k).$$

### Démonstration du lemme (2.3.2)

Sans perdre de généralité, on démontre seulement le cas  $\widehat{g}(x)$  et par analogie, on obtient le cas de l'estimateur  $\widehat{f}(x)$ . Comme  $S$  est un compact de  $\mathbb{R}$ , alors, on peut considérer le recouvrement suivant  $S \subset \bigcup B_k(t, l)$  où  $B_k$  est la boule centré en  $t_k \in S$  de rayon  $l_k = n^{-(\alpha+1)/2}$ . De ce recouvrement, on peut extraire un recouvrement fini, dont le nombre des boules est  $\tau_n \leq \frac{1}{l_n}$  et on écrit

$$\begin{aligned} \sup_{x \in S} |\widehat{g}(x) - \mathbb{E} \widehat{g}(x)| &= \sup_{x \in S} |\widehat{g}(x) - \widehat{g}(t_k(x)) + \widehat{g}(t_k(x)) - \mathbb{E} \widehat{g}(t_k(x)) + \mathbb{E} \widehat{g}(t_k(x)) - \mathbb{E} \widehat{g}(x)| \\ &\leq \sup_{x \in S} |\widehat{g}(x) - \widehat{g}(t_k(x))| + \sup_{x \in S} |\widehat{g}(t_k(x)) - \mathbb{E} \widehat{g}(t_k(x))| \\ &\quad + \sup_{x \in S} |\mathbb{E} \widehat{g}(t_k(x)) - \mathbb{E} \widehat{g}(x)| \end{aligned} \tag{2.17}$$

avec  $t_k(x)$  est le centre de la boule la plus proche de  $x$ . Ainsi,

$$\begin{aligned} P \left( \sup_{x \in S} |\widehat{g}(x) - \mathbb{E} \widehat{g}(x)| > \frac{\varepsilon}{2} \right) &\leq P \left( \sup_{x \in S} |\widehat{g}(t_k(x)) - \mathbb{E} \widehat{g}(t_k(x))| > \frac{\varepsilon}{6} \right) \\ &\quad + P \left( \sup_{x \in S} |\mathbb{E} \widehat{g}(x) - \mathbb{E} \widehat{g}(t_k(x))| > \frac{\varepsilon}{6} \right) \\ &\quad + P \left( \sup_{x \in S} |\widehat{g}(x) + \widehat{g}(t_k(x))| > \frac{\varepsilon}{6} \right). \end{aligned}$$



Pour le premier terme, on a

$$\begin{aligned}
|\widehat{g}(x) - \widehat{g}(t_k(x))| &= \left| \frac{1}{nh^p} \sum_{i=1}^n \left( K \left( \frac{x - X_i}{h} \right) Y_i \right) - \frac{1}{nh^p} \sum_{i=1}^n \left( K \left( \frac{t_k(x) - X_i}{h} \right) Y_i \right) \right| \\
&= \left| \frac{1}{nh^p} \sum_{i=1}^n K \left( \frac{x - X_i}{h} \right) Y_i - K \left( \frac{t_k(x) - X_i}{h} \right) Y_i \right| \\
&\leq \frac{C}{h^p} \left| \frac{x - X_i}{h} - \frac{t_k(x) - X_i}{h} \right| \\
&\leq \frac{C}{h^p} |x - t_k(x)| \\
&\leq C \frac{l_n}{h^p}.
\end{aligned}$$

Comme  $l = n^{-(\alpha+1)/2}$ ,

$$\begin{aligned}
\sup_{x \in S} |\widehat{g}(x) - \widehat{g}(t_k(x))| &\leq \frac{Cn^{-(\alpha+1/2)}}{h^p} \leq \frac{C}{h^p n^{(\alpha+1)/2}} \\
&\leq \frac{C}{n^{(\alpha+1)/2} h^p}.
\end{aligned}$$

L'hypothèse  $(H_4)$  implique qu'il existe

$$\frac{C}{n^{(\alpha+1)/2} h^p} = o \left( \sqrt{\frac{\log n}{nh_n^p}} \right).$$

Par conséquent,

$$\sup_{x \in S} |\widehat{g}(x) - \widehat{g}(t_k(x))| = o \left( \sqrt{\frac{\log n}{nh_n^p}} \right).$$

Pour le deuxième terme,

$$\begin{aligned}
\mathbb{P} \left( \sup_{x \in S} |\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x))| \right) &\leq P \left( \max_{k=1, \dots, \tau_n} |\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x))| \right) \\
&\leq \sum_{k=1}^{\tau_n} P (|\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x))|) \\
&\leq \tau_n \max_{k=1, \dots, \tau_n} P (|\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x))|).
\end{aligned}$$

Posons,

$$\Delta_i = \frac{1}{h^p} \left[ Y_i K \left( \frac{t_k(x) - X_i}{h} \right) - \mathbb{E} Y_i K \left( \frac{t_k(x) - X_i}{h} \right) \right].$$

D'où, on peut écrire

$$\begin{aligned} |\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x))| &= \left| \frac{1}{nh^p} \sum_{k=1}^n K \left( \frac{t_k(x) - X_i}{h} \right) Y_i - \mathbb{E} \frac{1}{nh^p} \sum_{k=1}^n K \left( \frac{t_k(x) - X_i}{h} \right) Y_i \right| \\ &= \left| \frac{1}{n} \sum_{k=1}^n \left[ \frac{1}{h^p} \left[ K \left( \frac{t_k(x) - X_i}{h} \right) Y_i - \mathbb{E} K \left( \frac{t_k(x) - X_i}{h} \right) Y_i \right] \right] \right| \\ &\leq \frac{1}{n} \sum_{k=1}^n \left| \frac{1}{h^p} \left[ K \left( \frac{t_k(x) - X_i}{h} \right) Y_i - \mathbb{E} \left( K \left( \frac{t_k(x) - X_i}{h} \right) Y_i \right) \right] \right| \\ &= \frac{1}{n} \sum_{i=1}^n \Delta_i \end{aligned}$$

On applique, maintenant, l'inégalité d'Hoeffding (1.2.1). Du faite que le noyau  $K$  et le variable aléatoire  $Y$  sont borné, alors, on peut dire,

$$|\Delta_i| \leq \frac{C}{h^p}.$$

Il suffit évaluer

$$\text{Var} \Delta_i \leq \mathbb{E}(\Gamma_i^2)$$

avec  $\Gamma_i = \frac{1}{h^p} \left[ Y_i K \left( \frac{t_k(x) - X_i}{h} \right) \right]$ , alors,

$$\begin{aligned} \mathbb{E}(\Gamma_i^2) &= \mathbb{E} \left( \frac{1}{h^{2p}} Y^2 K^2 \left( \frac{t_k(x) - X_i}{h} \right) \right) \\ &= \mathbb{E} \left( \frac{1}{h^{2p}} \mathbb{E}(Y^2/X_1) K^2 \left( \frac{t_k(x) - X_i}{h} \right) \right) \\ &= \frac{1}{h^{2p}} \int \phi(u) f(u) K^2 \left( \frac{t_k(x) - u}{h} \right) du \end{aligned}$$

où  $\phi(u) = \mathbb{E}(Y^2/X = u)$ .

On considère le changement des variables usuelle  $z = \frac{t_k(x) - u}{h}$  pour démontrer que,

$$\mathbb{E}(\Gamma_i^2) = \frac{1}{h^p} \int \phi(t_k(x) - zh) f(t_k(x) - zh) K^2(z) dz.$$

Puisque la fonction  $\phi$  est borné, la densité  $f$  est continue et le noyau  $K$  est à support compact, alors, il existe une constante  $C$  telle que :

$$\mathbb{E}(\Gamma_i^2) \leq \frac{C}{h^p}.$$

Par ailleurs, on applique l'inégalité de d'Hoeffding, pour laquelle on a,

$$\mathbb{P} \left[ \frac{1}{n} \left| \sum_{i=1}^n \Delta_i \right| > \varepsilon \right] \leq 2 \exp \left( \frac{-n\varepsilon^2}{4\delta_2} \right). \quad (2.18)$$

En choisissant,

$$\varepsilon = \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}}$$

pour tout  $\varepsilon_0 > 0$  à,

$$\begin{aligned} \mathbb{P} \left[ |\mathbb{E}\hat{g}(t_k(x)) - \hat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] &\leq 2 \exp \left( \frac{-n\varepsilon_0^2 h^p \log n}{4nh^p C} \right) \\ &\leq 2 \exp \left( \frac{\varepsilon_0^2 \log n}{4C} \right) \\ &\leq 2n^{-\frac{\varepsilon_0^2}{4C}}. \end{aligned}$$

Donc,

$$\max_{t_k} \mathbb{P} \left[ |\mathbb{E}\hat{g}(t_k(x)) - \hat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \leq 2n^{-\frac{\varepsilon_0^2}{4C}}.$$

ceci implique,

$$\tau_n \max_{t_k} \mathbb{P} \left[ |\mathbb{E}\hat{g}(t_k(x)) - \hat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \leq 2\tau_n n^{-\frac{\varepsilon_0^2}{4C}}.$$

D'où

$$\sum_{n=1}^{\infty} \tau_n \max_{t_k} \mathbb{P} \left[ |\mathbb{E}\hat{g}(t_k(x)) - \hat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \leq 2 \sum_{n=1}^{\infty} n^{((\alpha+1)/2 - \frac{\varepsilon_0^2}{4C})}$$

Un choix convenable de  $\varepsilon_0$  permet de conclure que

$$\sup_{x \in S} |\widehat{g}(x) - \widehat{g}(t_k(x))| = O\left(\sqrt{\frac{\log n}{nh^p}}\right). \quad P.co$$

En ce qui concerne le dernier terme du,

$$\begin{aligned} |\mathbb{E}\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(x)| &= \left| \mathbb{E} \frac{1}{nh^p} \sum_{i=1}^n K\left(\frac{t_k - X_i}{h}\right) Y_i - \mathbb{E} \frac{1}{nh^p} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) Y_i \right| \\ &= \left| \frac{1}{h^p} \left( \mathbb{E} \left( K\left(\frac{t_k(x) - X_i}{h}\right) Y_i \right) - \mathbb{E} \left( K\left(\frac{x - X_i}{h}\right) Y_i \right) \right) \right| \\ &\leq \frac{1}{h^p} \mathbb{E} \left| K\left(\frac{t_k(x) - X_i}{h}\right) Y_i - K\left(\frac{x - X_i}{h}\right) Y_i \right| \\ &\leq \frac{C}{h^p} \mathbb{E} \left| K\left(\frac{t_k(x) - X_i}{h}\right) - K\left(\frac{x - X_i}{h}\right) \right| \\ &\leq \frac{C}{h^{p+\beta}} \mathbb{E} |t_k(x) - x| \\ &\leq \frac{Cl_n}{h^p} \end{aligned}$$

De même, par définition de la suite, on obtient,

$$\begin{aligned} \sup_{x \in S} |\mathbb{E}\widehat{g}(x) - \mathbb{E}\widehat{g}(t_k(x))| &\leq \frac{Cn^{-(\alpha+1)/2}}{h^p} \leq \frac{C}{h^p n^{-(1+\alpha)/2}} \\ &\leq \frac{C}{(n^{(1+\alpha)/2} h)^p} \end{aligned}$$

Sous  $(H_4)$  on montre que on peut trouver,

$$\sup_{x \in S} |\mathbb{E}\widehat{g}(x) - \mathbb{E}\widehat{g}(t_k(x))| = O\left(\sqrt{\frac{\log n}{nh^p}}\right), \quad P.co$$

### Démonstration de (2.16)

Les résultats des lemmes (2.3.1) et (2.3.2) entraînent en particulier la convergence presque complète uniforme sur un compact  $S$  de  $\widehat{f}(x)$  vers  $f(x)$ . On a aussi,

$$\mathbb{P}\left(\inf_{x \in S} |\widehat{f}(x)| \leq \frac{\theta}{2}\right) \leq \mathbb{P}\left(\sup_{x \in S} |\widehat{f}(x) - f(x)| > \frac{\theta}{2}\right)$$

où  $\theta = \inf_{x \in S} f(x)$ . Ceci implique,

$$\sum_{i=1}^{\infty} \mathbb{P} \left( \inf_{x \in S} |\hat{f}(x)| \leq \frac{\theta}{2} \right) \leq \sum_{i=1}^{\infty} \mathbb{P} \left( \sup_{x \in S} |\hat{f}(x) - f(x)| > \frac{\theta}{2} \right) < \infty$$

D'ou le résultat du lemme avec  $\delta = \frac{\theta}{2}$

### 2.3.3 Cas $\alpha$ -mélange

On garde les mêmes hypothèses, ainsi les mêmes notations et on ajoute les deux hypothèses suivantes

(H7) Les fonctions  $r$  et  $f$  sont  $k$ -fois continûment dérivables au voisinage de  $x$ .

(H8) Le couple des variables aléatoires  $(X_i, X_j)$  admet une densité notée  $f_{ij}$ , pour tout  $i \neq j$

(H9) Les observations  $(X_i, Y_i), i = 1, 2 \dots n$  sont  $\alpha$  mélangeante et de coefficient de mélange il existe deux constantes  $C \in \mathbb{R}^{*+}$  et  $a \in \mathbb{R}^{*+}$  telles que

$$\alpha(n) \leq Cn^{-a}.$$

### 2.3.4 Propriétés asymptotiques

**Théorème 2.3.2** *Sous les hypothèses  $(H_1) - (H_6)$  et  $(H7) - (H9)$  on a*

$$\sup_{x \in S} |\hat{r}(x) - r(x)| = O(h^k) + O \left( \sqrt{\frac{\log n}{nh^p}} \right) \quad P.co. \quad (2.19)$$

**Démonstration** De même que le cas i.i.d, ce théorème est une conséquence directe des lemmes suivants :

**Lemme 2.3.3** *Sous les hypothèses du théorème, on a,*

$$\sup_{x \in S} |\hat{g}(x) - \mathbb{E}\hat{g}(x)| = O \left( \sqrt{\frac{\log n}{nh^p}} \right), \quad (2.20)$$

$$\sup_{x \in S} |\hat{f}(x) - \mathbb{E}\hat{f}(x)| = O \left( \sqrt{\frac{\log n}{nh^p}} \right). \quad (2.21)$$

**Démonstration du lemme (2.3.3)**

On garde les mêmes arguments utilisé dans la démonstration du du lemme (2.3.1) (resp. du lemme (2.3.2)), ainsi que le même recouvrement du compact  $S$  et on démontre seulement l'équation (2.20) et par analogie on obtient (2.21). Il est claire avec le choix de  $\alpha = \frac{(p+2)(a-4)}{pa+2p+2}$  nous permet de montre que le premier et le troisième de la décomposition (2.17) terme sont de type  $o\left(\sqrt{\frac{\log n}{nh_n^p}}\right)$ . En effet, pour le deuxième terme

$$\begin{aligned}\widehat{g}(t_k(x)) - \mathbb{E}\widehat{g}(t_k(x)) &= \frac{1}{nh^p} \sum_{i=1}^n Y_i K\left(\frac{t_k(x) - X_i}{h}\right) - \mathbb{E}\left(\frac{1}{nh^p} \sum_{i=1}^n Y_i K\left(\frac{t_k - X_i(x)}{h}\right)\right) \\ &= \frac{1}{nh^p} \sum_{i=1}^n \left(Y_i K\left(\frac{t_k(x) - X_i}{h}\right) - \mathbb{E}\left(Y_i K\left(\frac{t_k(x) - X_i}{h}\right)\right)\right) \\ &= \frac{1}{nh^p} \sum_{i=1}^n \Delta_i(t_k(x))\end{aligned}$$

En utilisant l'inégalité exponentielle de Fuk-Nagaev (lemme (1.2.2)), on a pour tout  $\varepsilon > 0$  et pour tout  $r > 1$ ,

$$\mathbb{P}\left[\left|\sum_{k=1}^n \Delta_k\right| > 4\varepsilon\right] \leq \left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{\frac{-r}{2}} + 2ncr^{-1} \left(\frac{2r}{\varepsilon}\right)^{a+1} \quad (2.22)$$

où

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\text{cov}(\Delta_i, \Delta_j)|.$$

On a déjà démontré dans le Théorème (2.2.2) que

$$S_n^2 = O(nh^p)$$

alors

$$\begin{aligned}\mathbb{P}\left[|\mathbb{E}\widehat{g}(t_k(x)) - \widehat{g}(t_k(x))| > \varepsilon\right] &= \mathbb{P}\left[\left|\frac{1}{nh_n^p} \sum_{i=1}^n \Delta_i\right| > \varepsilon\right] \\ &= \mathbb{P}\left[\left|\sum_{i=1}^n \Delta_i\right| > \frac{4(\varepsilon nh^p)}{4}\right] \\ &\leq 4 \left(1 + \frac{\varepsilon^2 n^2 h^{2p}}{nh^p 16r}\right)^{\frac{-r}{2}} + 2ncr^{-1} \left(\frac{8r}{\varepsilon nh^p}\right)^{a+1}.\end{aligned}$$

Pour  $\varepsilon = \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh_n^p}}$

$$\begin{aligned} \mathbb{P} \left[ |\mathbb{E}\widehat{g}(t_k(x)) - \widehat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \\ \leq 4 \left( 1 + \frac{\varepsilon_0^2 n^2 h^{2p} \log n}{n^2 h^{2p} 16r} \right)^{-\frac{r}{2}} + 2ncr^{-1} \left( \frac{8r(nh^p)^{\frac{1}{2}}}{\varepsilon_0 nh^p (\log n)^{\frac{1}{2}}} \right)^{a+1} \\ \leq 4 \left( 1 + \frac{\varepsilon_0^2 \log n}{16r} \right)^{-\frac{r}{2}} + 2ncr^{-1} \left( \frac{8r}{\varepsilon_0} \right)^{a+1} (nh^p \log n)^{-\left(\frac{a+1}{2}\right)}. \end{aligned}$$

Pour un choix de  $r = C(\log n)^2$ ,

$$\begin{aligned} \leq 4 \left( \exp \left( -\frac{\varepsilon^2}{32} \log n \right) + nr^a (nh^p \log n)^{-\left(\frac{a+1}{2}\right)} \right) \\ \leq 4 \left( n^{-\frac{\varepsilon^2}{32}} + n^{1-\frac{a+1}{2}} h^{-p\frac{a+1}{2}} (\log n)^{2a-\left(\frac{a+1}{2}\right)} \right) \end{aligned}$$

L'hypothèse  $h^{-p\left(\frac{a+1}{2}\right)} \leq Cn^{-(\frac{3-a}{a}-\theta)\left(\frac{a+1}{2}\right)}$  implique que

$$\begin{aligned} \mathbb{P} \left[ |\mathbb{E}\widehat{g}(t_k(x)) - \widehat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \\ \leq 4 \left( \left( n^{-\frac{\varepsilon^2}{32}} + n^{1-\frac{a+1}{2}-\left(\frac{3-a}{a}-\theta\right)\left(\frac{a+1}{2}\right)} \log n \right)^{2a-\left(\frac{a+1}{2}\right)} \right). \end{aligned}$$

Donc,

$$\begin{aligned} \tau_n \mathbb{P} \left[ |\mathbb{E}\widehat{g}(t_k(x)) - \widehat{g}(t_k(x))| > \frac{\varepsilon_0}{6} \sqrt{\frac{\log n}{nh^p}} \right] \\ \leq Cn^{(\alpha+1)/2} \left( \left( n^{-\frac{\varepsilon^2}{32}} + n^{1-\frac{a+1}{2}-\left(\frac{3-a}{a}-\theta\right)\left(\frac{a+1}{2}\right)} \log n \right)^{2a-\left(\frac{a+1}{2}\right)} \right) \end{aligned}$$

Pour un choix convenable de  $\varepsilon_0 >$  le premier terme est un terme d'une série converge et à l'aide d'un calcul simple, on montre aussi que le deuxième terme fait l'objet d'un terme d'une série converge

$$\sup_{x \in S} |\mathbb{E}\widehat{g}(t_k(x)) - \widehat{g}(t_k(x))| = O \left( \frac{\log n}{nh^p} \right) \quad P.co$$





## Chapitre 3

# Régressoin semi paramétrique fonctionnelle

Le but de ce chapitre est d'introduire le modèle à indice simple de la fonction de la régression de variable de réponse scalaire par rapport à valeurs dans un espace de Hilbert. Sous des conditions générales, nous établissons la convergence presque complete avec la vitesse de convergence dans les deux cas indépendants et dépendant.

### 3.1 Modèle

Nous introduisons  $n$  paires de variables aléatoires  $(X_i, Y_i)$  pour  $i = 1, \dots, n$  du couple  $(X, Y)$  à valeur dans  $\mathcal{H} \times \mathbb{R}$ , tel que  $\mathcal{H}$  est un espace de Hilbert. Cet espace est supposé muni d'un produit scalaire. Le modèle de régression étudié est donnée par l'équation

$$Y = r(\langle \theta_\theta, X \rangle) + \epsilon,$$

Tel que  $\epsilon$  est une variable aléatoire réelle indépendante de  $X$  de loi normale centrée réduite. La fonction  $r_{\theta_\theta}$  à indice fonctionnel simple est définie par

$$r_{\theta_\theta}(x) = E(Y / \langle \theta_\theta, X = x \rangle), \quad \theta_\theta \in \mathcal{H}$$

Etant donnée  $(X_1, Y_1), \dots, (X_n, Y_n)$  des observations de  $(X, Y)$ . L'estimateur de la fonction de régression  $\hat{r}_\theta$  de  $r_\theta$ , est donnée par :

$$\forall x \in \mathcal{H}, \quad \hat{r}_\theta(x, h) = \frac{\sum_{i=1}^n K\left(\frac{\langle \theta_\theta, x - X_i \rangle}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{\langle \theta_\theta, x - X_i \rangle}{h}\right)}$$

Où  $K$  est un noyau,  $h = h_n$  est une suite de nombre réels positifs.

## 3.2 Convergence presque complète ponctuelle

### 3.2.1 Cas i.i.d

Soit  $x$  un point fixe dans  $\mathcal{H}$ , dans notre modèle semi paramétrique on a besoin les conditions suivantes :

1.  $\forall \alpha > 0, \quad G_{\theta_0}(x, h) = \mathbb{P}(|\langle \theta_0, X - x \rangle| < \alpha) > 0$
2. La variable de réponse est telle que :

$$|Y| \leq M < \infty$$

3.  $\exists C < \infty, \exists \beta > 0, \forall (x, y) \in \mathcal{H}_{\theta_0}^2,$

$$|r_{\theta_0}(x) - r_{\theta_0}(y)| \leq C |\langle \theta_0, x - y \rangle|^\beta$$

4. Le noyau  $K$  est lipchitzien d'ordre 1, de support  $[-1, 1]$ , vérifiant

$$\exists(k_1, k_2), \quad \forall t \in [-1, 1], \quad 0 < k_1 < K(t) < k_2 < \infty$$

5. Le paramètre de lissage  $h$  est tel que :

$$\lim_{n \rightarrow +\infty} h = 0, \quad \lim_{n \rightarrow +\infty} \frac{nG_{\theta_0}(x, h)}{\log n} = \infty$$

**Théorème 3.2.1** *Sous les l'hypothèses (1 – 5), on a*

$$\widehat{r}_{\theta, l}(x) - r_{\theta_0}(x) = o(h^k) + o\left(\sqrt{\frac{\log(n)}{nG_{\theta_0}(x, h)}}\right) \quad p.co$$

tel que

$$\widehat{r}_{\theta, l}(x) = \frac{1}{n\mathbb{E}(K_1(x))} \sum_{i=1}^n Y_i^l K_i(x) \quad \text{avec } l = 0, 1$$

où

$$K_i(x) = K\left(\frac{\langle \theta_0, x - X_i \rangle}{h}\right), \quad i = \overline{1, n}$$

la preuve de ce théorème est basé sur la decomposition suivante :

$$\begin{aligned}\widehat{r}_{\theta,l}(x) - r_{\theta_0}(x) &= \frac{1}{\widehat{r}_{\theta,1}(x)} (\widehat{r}_{\theta,1}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x)) + (r_{\theta_0}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) \\ &\quad - \frac{r_{\theta_0}(x)}{\widehat{r}_{\theta,0}(x)} ((\widehat{r}_{\theta,0}(x)) - \mathbb{E}(\widehat{r}_{\theta,0}(x)) - (\mathbb{E}\widehat{r}_{\theta,0}(x) - 1))\end{aligned}$$

Le théorème (3.2.1) va découler des lemmes suivants :

**Lemme 3.2.1** *Sous les hypothèses (1 – 4), on a*

$$r_{\theta_0}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x)) = o(h^k) \quad (3.1)$$

**Lemme 3.2.2** *Sous les hypothèses (1 – 5), on a*

$$\widehat{r}_{\theta,l}(x) - \mathbb{E}(\widehat{r}_{\theta,l}(x)) = o\left(\sqrt{\frac{\log(n)}{nG_{\theta_0}(x,h)}}\right), \quad l = 0, 1 \quad (3.2)$$

**Corollaire 3.2.1** *Sous les hypothèses du lemme (3.2.2), on a*

$$\exists \delta > 0, \quad \sum_{i=1}^{\infty} \mathbb{P}(|\widehat{r}_{\theta,0}(x)| \leq \delta) < \infty \quad (3.3)$$

**Preuve du lemme 3.2.1 :**

On a, par équadistribution des observations,

$$\begin{aligned}|\mathbb{E}(\widehat{r}_{\theta,1}(x)) - r_{\theta_0}(x)| &= |(\mathbb{E}(K_1(x)))^{-1} (\mathbb{E}Y_1 K_1(x)) - r_{\theta_0}(x)| \\ &= (\mathbb{E}(K_1(x))^{-1} |\mathbb{E}(r_{\theta_0}(X_1) K_1(x)) - \mathbb{E}(K_1(x)) r_{\theta_0}(x)| \\ &\leq (\mathbb{E}(K_1(x))^{-1} \mathbb{E}(|r_{\theta_0}(X_1) - r_{\theta_0}(x)| K_1(x)) \\ &= C |< \theta, x - X_1 >|^{\beta} \\ &\leq ch^{\beta}\end{aligned}$$

La dernière majoration est une conséquence directe de l'hypothèse (H3). Par passage à la limite, quand  $n$  tend vers l'infinie, on obtient le résultat.

**Preuve du lemme 3.2.2 :**

L'idée est d'appliquer l'inégalité de Hoeffding (lemme (1.2.2)) aux variables

$$\Delta_i = \frac{1}{\mathbb{E}K_1(x)} (Y_i^l K_i(x) - \mathbb{E}(Y_i^l K_i(x))), \quad l = 0, 1$$

En effet,

$$\hat{r}_{\theta,l}(x) - \mathbb{E}(\hat{r}_{\theta,l}(x)) = \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n \Delta_i$$

Pour cela il est nécessaire de trouver des majorations pour  $|\Delta_i|$  et  $\mathbb{E}(\Delta_i^2)$ , alors d'après  $(H_1)$ ,  $(H_2)$  et  $(H_4)$  on arrive sans difficulté à

$$|\Delta_i| \leq \frac{C}{G_{\theta_0}(x, h)}, \quad \mathbb{E}(\Delta_i^2) \leq C/G_{\theta_0}(x, h)$$

L'inégalité permet d'écrire que pour  $\epsilon_0 > 0$  suffisamment petit on a

$$\begin{aligned} \mathbb{P} \left( |\mathbb{E}(\hat{r}_{\theta,l}(x)) - \hat{r}_{\theta,l}(x)| > \epsilon_0 \sqrt{\frac{\log n}{nG_{\theta_0}(x, h)}} \right) &\leq 2 \exp \left( \frac{-n\epsilon_0^2 \frac{\log n}{nG_{\theta_0}(x, h)}}{4 \frac{C}{G_{\theta_0}(x, h)}} \right) \\ &\leq 2n^{-\frac{\epsilon_0^2}{4C}} \end{aligned}$$

Il suffit de choisir  $\epsilon_0$  de telle sorte que  $\epsilon_0 > 2\sqrt{C}$  pour que la série converge comme une série de Rieman. Par conséquent,

$$\exists \epsilon_0 > 0, \quad \sum_{i=1}^n \mathbb{P} \left( |\mathbb{E}(\hat{r}_{\theta,l}(x)) - \hat{r}_{\theta,l}(x)| > \epsilon_0 \sqrt{\frac{\log n}{nG_{\theta_0}(x, h)}} \right) < +\infty$$

d'où

$$\hat{r}_{\theta,l}(x) - \mathbb{E}(\hat{r}_{\theta,l}(x)) = o \left( \sqrt{\frac{\log n}{nG_{\theta_0}(x, h)}} \right) \quad p.co$$

**Preuve du corrolaire 3.2.1** On a  $\mathbb{E}(\hat{r}_0(x)) = 1$ , alors,

$$\mathbb{P}(|\hat{r}_{\theta,0}(x)| \leq 1/2) \leq \mathbb{P}(|\hat{r}_0(x) - \mathbb{E}(\hat{r}_{\theta,0}(x))| \geq 1/2)$$

Donc,

$$\sum_n \mathbb{P}(|\hat{r}_{\theta,0}(x)| \leq 1/2) \leq \sum_n \mathbb{P}(|\hat{r}_0(x) - \mathbb{E}(\hat{r}_{\theta,0}(x))| \geq 1/2) \leq \infty$$

### 3.2.2 Cas de mélange forte

Pour étendre des résultats de convergence du cadre i.i.d à un cadre de dépendance, la difficulté se traduit le plus souvent par la nécessité d'avoir à inclure des termes de type covariance dans les calculs habituels de variance.

### 3.2.3 Hypothèses

On garde les mêmes hypothèses ainsi les mêmes notations et on ajouté les conditions suivante :

(H1) Pour tout  $i, j$ ;

$$\sup_{i,j} \psi_{ij}(x, h) = o\left(\left(\frac{G_{\theta_0}(x, h)}{n}\right)^{\frac{p}{a(p+2)}}\right)$$

où

$$\psi_{ij}(x, h) = \frac{\mathbb{P}(|\langle \theta_0, X_i - x \rangle| < h, |\langle \theta_0, X_j - x \rangle| < h)}{G_{\theta_0}(x, h)}$$

H2 La variable réponse est telle que : pour tout  $i, j$

$$\mathbb{E}(|Y_i Y_j| / (X_i, X_j)) \leq M < \infty$$

H3 Le coefficient de mélange est tel que

$$\exists c \in \mathbb{R}^{*+} \quad \text{et} \quad \exists a \in \mathbb{R}^{*+}, \quad \alpha(n) \leq cn^{-a}$$

H4 Il existe un  $\eta > 0$  telle que  $\lim_{n \rightarrow \infty} h_n = 0$ ,  $\lim_{n \rightarrow \infty} n^{1 - \frac{4(a+p)}{p(a+1)} - \eta} G_{\theta_0}(x, h) = \infty$

### 3.2.4 Propriétés asymptotiques

**Théorème 3.2.2** *Sous les hypothèses (1 – 5) et (H1) – (H4), on a*

$$\hat{r}_{\theta}(x) - r_{\theta}(x) = o(h^{\beta}) + o\left(\frac{\log n}{nG_{\theta}(x, h)}\right), \quad p.co \quad (3.4)$$

**Démonstration**

La preuve de ce théorème est basé seulement sur la partie de dispersion. Cependant, le terme de covariance n'a aucune influence sur la partie de biais.

**Lemme 3.2.3** *Sous les hypothèses du Théorème, on a :*

$$\hat{r}_{\theta,l}(x) - \mathbb{E}(\hat{r}_{\theta,l}(x)) = o\left(\sqrt{\frac{\log n}{nG_{\theta}(x,h)}}\right), \quad l = 0, 1 \quad (3.5)$$

**Preuve :**

Utilisons la décomposition suivante

$$\hat{r}_{\theta,l}(x) - \mathbb{E}(\hat{r}_{\theta,l}(x)) = \frac{1}{n\mathbb{E}K_i(x)} \sum_{i=1}^n \Delta_i(x)$$

Une application directe de (1.2.3) aux variables

$$\Delta_i(x) = Y_i^l K_i(x) - \mathbb{E}Y_i^l K_i(x), \quad l = 0, 1$$

Amène alors que pour tout  $\lambda > 0$  et pour tout  $r > 1$ ,

$$\mathbb{P}\left[\left|\sum_{k=1}^n \Delta_{i,l}\right| > 4\lambda n\mathbb{E}(k_1(x))\right] \leq 4(A_1 + A_2)$$

Tel que

$$A_1 = \left(1 + \frac{\lambda^2 n^2 (\mathbb{E}K_1(x))^2}{r s_{n,l}^2}\right)^{\frac{-r}{2}} \quad \text{et} \quad A_2 = ncr^{-1} \left(\frac{r}{\lambda n\mathbb{E}(K_1(x))}\right)^{\frac{p(a+1)}{a+p}}$$

On ce qui concerne le terme  $A_1$ , on a

$$\begin{aligned} s_{n,l}^2 &= \sum_{j=1}^n \sum_{i=1}^n |\text{cov}(\Delta_i, \Delta_j)| \\ &= \sum_{i \neq j}^n \sum_{i=1}^n |\text{cov}(\Delta_i, \Delta_j)| + n\text{Var}(\Delta_1) \\ &= S_n^{2*} + n\text{Var}(\Delta_1) \end{aligned}$$

Par ailleurs,

$$\begin{aligned}
|cov(\Delta_i, \Delta_j)| &= |\mathbb{E}(Y_i Y_j K_i(x) K_j(x)) - (\mathbb{E} Y_1 K_1(x))^2| \\
&\leq C \mathbb{P}(\langle \theta_0, X_i - x \rangle | \langle h, | \langle \theta, X_j - x \rangle | \langle h) \\
&\quad + (\mathbb{P}(\langle \theta, X_i - x \rangle | \langle h))^2 \\
&\leq C(\psi_{i,j}(x, h) G_\theta(x, h) + G_\theta(x, h)^2)
\end{aligned}$$

D'après l'hypothèse (H1) on a,

$$\begin{aligned}
|cov(\Delta_{i;l}, \Delta_{j;l})| &\leq C \left( \left( \frac{G_\theta(x, h)}{n} \right)^{\frac{p}{a(p-2)}} G_\theta(x, h) + G_\theta(x, h)^2 \right) \\
&= 0 \left( \left( \frac{G_\theta(x, h)}{n} \right)^{\frac{p}{a(p-2)}} G_\theta(x, h) \right)
\end{aligned}$$

On peut aussi majorer cette covariance directement à partir des techniques de Masry ([7, 1986]) et on partage la somme sur les deux ensembles suivant :

$$\begin{cases} S_1 = \{(i, j) \text{ tel que } 1 \leq j - i \leq u_n\} \\ S_2 = \{(i, j) \text{ tel que } u_n + 1 \leq j - i \leq n - 1\} \end{cases}$$

Alors,

$$\begin{aligned}
S_n^2 &= \sum_{i,j=1}^n |cov(\Delta_i(x) + \Delta_j(x))| \\
&= \sum_{S_1} |cov(\Delta_i(x) + \Delta_j(x))| + \sum_{S_2} |cov(\Delta_i(x) + \Delta_j(x))| \\
&\leq C \left( \left( \frac{G_\theta(x, h)}{n} \right)^{\frac{p}{a(p-2)}} G_\theta(x, h) n u_n + n^2 (\alpha_{|i-j|})^{\frac{p-2}{p}} \right)
\end{aligned}$$

Le choix de  $u_n = \left( \frac{G_\theta(x, h)}{n} \right)^{\frac{-p}{a(p-2)}}$  implique

$$S_n^{2*} = o(n G_\theta(x, h)) \quad (3.6)$$

Ainsi le terme de variance, d'après les hypothèses (H1, H2, H4) on a

$$\text{Var}(\Delta_i(x)) = o(nG_\theta(x, h)) \quad (3.7)$$

À l'aide de (3.6) et (3.7) on a,

$$s_{n,l}^2 = o(nG_\theta(x, h))$$

On prend  $r = C(\log n)^2$  et  $\lambda = \lambda_0(n\mathbb{E}K_1(x))^{-1}\sqrt{nG_\theta(x, h) \log n}$  pour obtenir

$$\exists \nu > 0, \quad A_1 \leq Cn^{-1-\nu}$$

et

$$A_2 \leq Cn^{-1-\nu}$$

Ce qui implique que,

$$\exists \nu > 0, \quad \mathbb{P} \left( |\hat{r}_{\theta,l}(x) - \mathbb{E}(\hat{r}_{\theta,l}(x))| > \lambda_0 \sqrt{\frac{\log n}{nG_\theta(x, h)}} \right) \leq Cn^{-1-\nu}$$

Ceci achève la démonstration du lemme.



# Chapitre 4

## Simulation

L'objectif de ce chapitre est de montrer à l'aide de données simulées, comment on peut implémenter facilement et rapidement la fonction de régression par la méthode du noyau et que le choix du paramètre de lissage joue un rôle crucial dans le comportement asymptotique de l'estimateur en insistant sur une méthode de sélection du paramètre de lissage.

### 4.1 Algorithme

On considère le modèle de régression suivant :

$$Y_i = f(X_i) + \epsilon_i$$

Où  $f(x, y) = \exp(0.9x + 0, 1y)$ ,  $X$  est de dimension 2 et les deux composantes sont de loi normale centrée réduite et indépendante de  $\epsilon$  qui suit aussi la loi  $\mathcal{N}(0, 1)$  et on simule par un noyau exponentiel. On procède par l'algorithme suivant :

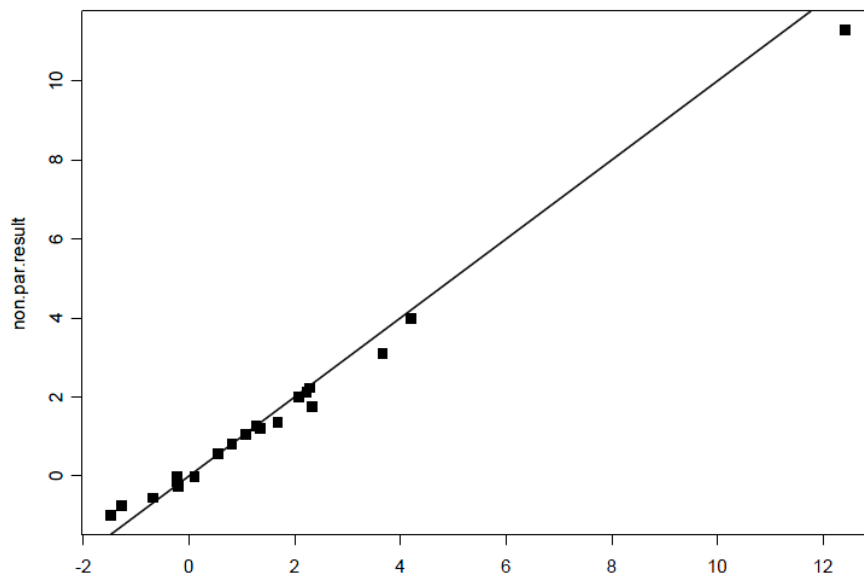
1. On divise nos observations en deux paquets : d'apprentissage  $(X_i, Y_i)_{i=\overline{1,80}}$  et du test  $(X_i, Y_i)_{i=\overline{81,100}}$ .
2. Déterminer les trois prédicteurs : Linéaire (par la méthode des moindres carrés), non paramétrique (resp. semi paramétrique) par la méthode du noyau.
3. On trace la droite de régression entre les deux quantités (les vraies valeurs et les valeurs prédites) pour les trois méthodes.
4. On calcul la moyenne d'erreurs entre les vraies valeurs et les valeurs prédites pour tester les méthodes.

### 4.1.1 Résultats

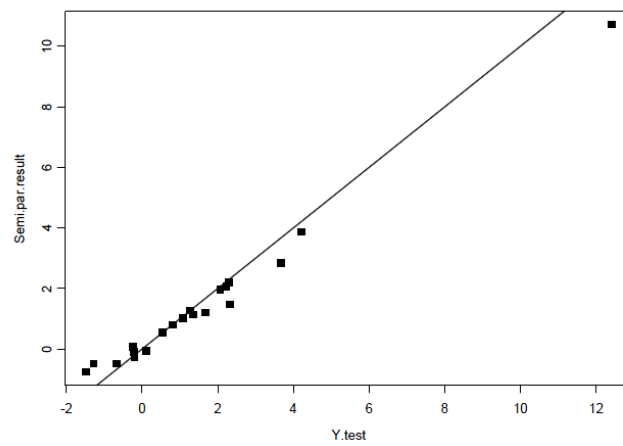
Il est clair que dans cet exemple la méthode non-paramétrique est la plus appropriée et que la méthode semi-paramétrique prévoit mieux que la méthode linéaire. Le tableau suivant donne les paramètres optimaux de la méthode non-paramétrique. Notons que, on donne juste la première composante pour l'indice de modèle, parce que la deuxième est obtenue en utilisant le fait que cet indice est normalisé

theta.opt	h.opt
0.1372678	1.313573682
0.1372678	0.130342340
-0.0118034	0.516645127
0.8378870	0.130117703
0.1621130	0.129000271
0.0875774	0.057019245

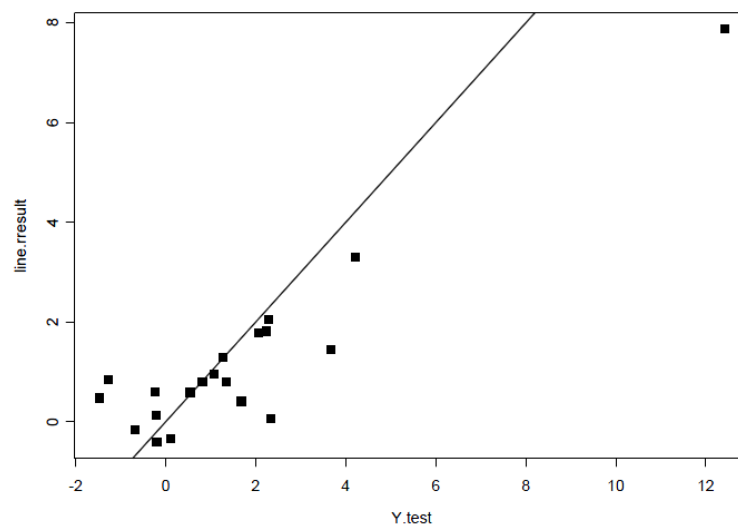
Les résultats sont donnés par le graphe suivant :



La prévision via la régression non-paramétrique, erreur=0.236547



La prevision par la régression semi-paramétrique, erreur=0.523669



La prévision via la régression linéaire, err=1.127252



# Conclusion

On peut trouvé deux remarques concernant l'ensemble des résultats obtenus :

1. L'augmentation de la dimension à une importance capitale dans la dégradation de la vitesse de convergence. Cependant, la dégradation est liée à la faible concentration de la mesure de probabilité de la variable explicative fonctionnelle.
2. Le choix du paramètre du paramètre de lissage à une importance dans l'estimation par la méthode du noyau.



# Bibliographie

- [1] Ait Saidi, A., Ferraty, F., Kassa, R., (2005), *Single functional index model for time series*. Rev. Roumaine Math. Pures Appl., in print.
- [2] Bosq, D., (2000), *Linear processes in function spaces. Theory and applications*. Lecture Notes in Statistics. New York.
- [3] Ferraty, F et Vieu, Ph. (2004). Modèle de régression pour variables aléatoires uni, multi et  $\infty$ -dimensionnées. Cours de DEA.
- [4] Ferraty, F., Peuch, A., Vieu, Ph., (2003), *Modèle à indice fonctionnel simple*. C. R., Math., Acad. Sci. Paris **336**, No.12, 1025-1028.
- [5] Ferraty, F et Vieu, Ph. (2005). Nonparametric functional data analysis. Springer-Verlag.
- [6] Ichimura, Hidehiko. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. J. Econometrics 58 (1993), no. 1-2, 71–120
- [7] Masry, E. (1986). Recursive probability density estimation for weakly dependent stationary processes, IEEE Trans. Inform. Theory, IT18, 254-267.
- [8] Ramsay, J. O., Silverman, B. W., (2002), *Applied functional data analysis. Methods and case studies*. Springer Series in Statistics. Springer-Verlag, New York.
- [9] Tukey, J. (1961). Curves as parameters and touch estimation. *Proc. 4th Berkeley Symp. Math. Statist. Probab.* 1 681-684.
- [10] J. M. Bardet. Tests d'autosimilarité des processus gaussiens. Dimension fractale et dimension de corrélation. Thèse 3eme cycle, Paris-Sud, (1997).
- [11] F. Ferraty et P. Vieu Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *Compte Rendus Acad. Sci. Paris*, 330, 139-142. (2000).

- 
- [12] F. Ferraty et P. Vieu Functional Nonparametric Model : a New Tool for Spectrometric Data. soumis pour publication.(2001).
- [13] F. Ferraty et P. Vieu Modèles Non-paramétriques de Régression. Cours de D.E. A. (2002- 2003).
- [14] F. Ferraty et P. Vieu Nonparametric models for function data, with application in regression, time-series prediction and curve discrimination. The International Conference on Recent Trends and Directions in Nonparametric Statistics. J. Nonparametric. Stat. 16 (1-2), 111- 125, (2004).
- [15] F. Ferraty et P. Vieu Nonparametric function data. Springer series statistics. (2006). 62
- [16] F. Ferraty et P. Vieu. Nonparametric modelling for function data. Springer-Verlag, New York, (2006a).
- [17] F. Ferraty et P. Vieu Function nonparametric statistics in action. The art of semiparametrics. Contrib. Statist. Physica-Verlag / Springer, Heidelberg, 112-129, (2006b).
- [18] R. D. Gill and S. Johansen. A survey of product integration with a view toward application in survival analysis. The Annals of Statistics. Vol 18, No 4, 1501-1555, 1990.
- [19] Z. Guessoum and E. Ould Said. On the nonparametric estimation of the regression function under censorship model. Statist.Decisions. 26, 159-177, (2009).