

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université A.MIRA- Béjaia  
Faculté des Sciences Exactes  
**Département de Mathématiques/M.I.**



جامعة بجاية  
Tasdawit n'Bgayet  
Université de Béjaïa

## MEMOIRE DE MAGISTER

Intitulé

**Convergence presque complète d'estimateurs non  
paramétriques pour des variables fonctionnelles**

Spécialité : Mathématiques  
Option : Analyse et Probabilités

Présenté par  
**Rachid AGOUNE**

Devant le jury composé de :

<i>Président</i>	OURBIH	Megdouda	Professeur, U.A.M Béjaia
<i>Rapporteur</i>	AIT SAIDI	Ahmed	Professeur, U.A.M Béjaia
<i>Examineur</i>	DAHMANI	Abdelnasser	Professeur, U.A.M Béjaia
<i>Examineur</i>	KASSA	Rabah	MCA, U.A.M Béjaia

## Remerciement

*Mes vifs remerciements et toute ma reconnaissance*

*Au Pr Ait Saidi Ahmed mon encadreur qui, par ses encouragements et son aide précieuse, a permis à ce que ce travail aboutisse.*

*Aux Pr Ourbih Megdoua, Pr Dahmani Abdchasser et Drs Kassa Rabah, de l'honneur qu'ils me feront en acceptant d'être membres de jury.*

*A ma famille et à tous mes amis qui m'ont soutenu moralement.*

*A tous ce qui de loin ou de près ont contribué à la réalisation de ce modeste travail.*

*Merci à tous ceux qui m'ont aidé sans ménager ni leurs temps, ni leurs encouragements, ni leurs savoirs.*

*Merci à tous les chercheurs que j'ai pu rencontrer et qui se sont intéressés à mes travaux*



## 0.1 Introduction

La statistique fonctionnelle a connu un très important développement ces dernières années. Cette branche de statistique vise à étudier des données qui, de part leurs structures et le fait qu'elles soient collectées sur des grilles très fines, comme la fonction du temps ou de l'espace par exemple. Le besoin de considérer ce type de données, maintenant couramment rencontré sous le nom de données fonctionnelles dans la littérature, est avant tout un besoin pratique. Compte tenu des capacités actuelles des appareils de mesure et de stockage informatique, les situations pouvant fournir de telles données sont multiples et issues de domaines variés.

Cependant au-delà de cet aspect pratique, il est nécessaire de donner un cadre théorique pour l'étude de ces données. Bien que la statistique fonctionnelle ait les mêmes objectifs que les autres branches de la statistique (analyse des données, inférence,...), les données qui ont cette particularité prennent leurs valeurs dans des espaces fonctionnelles, et les méthodes usuelles de la statistique multivariée sont ici mises en défaut. En effet, la principale source de difficultés, tant d'un point de vue théorique que pratique, provient du fait que les observations de ce type de variables fonctionnelles sont supposées appartenir à un espace de dimension infinie. Ainsi, l'intérêt de ce travail réside dans l'apport de solutions à ce problème de dimension infinie, dans les deux cadres indépendant et mélangeant, en mettant en place un cadre théorique suffisamment général. Les outils utilisés pour les développements théoriques sont de natures variées. En effet, ils relèvent de l'analyse fonctionnelle, mais aussi d'outils probabilistes telles que les inégalités exponentielles pour des sommes de variables aléatoires.

Ce travail a pour objet une familiarisation avec les méthodes statistiques qui permettent de tester la validité des théories économiques (et éventuellement celles d'autres sciences sociales). Il vise à rendre les utilisateurs de ces méthodes aptes à choisir les techniques les plus adéquates pour résoudre un problème donné, à interpréter les résultats obtenus lors de leurs applications ainsi qu'à évaluer la validité des hypothèses sur lesquelles leurs propriétés optimales reposent.

Un intérêt particulier sera porté aux exemples concrets d'application. Ces exemples seront choisis dans diverses disciplines de la science économique (macro-économie, micro-économie, économie du travail, économie publique, économie internationale) mais également dans les domaines extérieures à l'économie (criminologie, agronomie, écologie, pédagogie, démographie).

Dans ce mémoire, nous proposons d'apporter une contribution à l'étude des données fonctionnelles dans le contexte où la variable fonctionnelle sert à expliquer un phénomène représenté par une autre variable. Le premier problème qui va nous intéresser est celui d'une fonction de régression dans le cas où la variable explicative est fonctionnelle. C'est un sujet sur lequel la littérature est très conséquente. D'un point de vue très général, ce modèle de

régression fonctionnelle peut s'écrire :

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

- $X$  est une variable aléatoire fonctionnelle explicative supposée à valeurs dans un espace de dimension infini  $\mathcal{F}$ ,
- $Y$  est une variable aléatoire réelle (variable réponse),

-  $\epsilon$  est une variable aléatoire d'erreurs telle que  $E(\epsilon / X) = 0$ .

- Les variables aléatoires considérées sont toutes définies sur un même espace probabilisé  $(\Omega, \mathcal{A}, P)$  :

Le but est alors d'estimer l'opérateur  $r : \mathcal{F} \rightarrow \mathbb{R}$  inconnu, la répartition conditionnelle, la densité conditionnelle, le mode, la médiane et le quantile conditionnel sur la base des données  $(X_i; Y_i)_{i=1, \dots, n}$  dans le cadre indépendant et sous le mélange fort. Le choix d'un tel modèle a été guidé par le fait que l'on reproche au modèle non-paramétrique son manque d'éléments permettant de représenter la relation entre la variable réelle et la variable fonctionnelle.

Afin de définir le cadre de notre étude, il nous a paru logique de présenter les différents thèmes abordés dans un ordre respectant leur chronologie en donnant ce qui nous paraissent les plus importants.

Ainsi, nous avons opté pour un plan de travail scindé en quatre chapitres :

Le premier chapitre est consacré à donner des définitions et mettre en place les moyens nécessaires à l'étude du modèle de régression. Nous avons décidé de présenter brièvement quelques outils mathématiques et de probabilité nécessaires à l'étude asymptotique des estimateurs à noyau de la fonction de régression, la répartition conditionnelle, la densité conditionnelle, le mode, la médiane et le quantile conditionnel, objectif de ce mémoire. Dans ce but nous présenterons les différentes notions de convergence, nous rappellerons certaines inégalités exponentielles pour des sommes de variables aléatoires.

Le second chapitre sera consacré à la prédiction des estimateurs des fonctions citées ci-dessus et nous allons traiter un exemple pour clarifier les concepts : variable fonctionnelle,

donnée fonctionnelle, modélisation des variables fonctionnelles et non-paramétrique.

Dans le troisième, nous présenterons le cadre théorique de notre travail, nous allons traiter la convergence presque complète des estimateurs non-paramétriques des fonctions citées ci-dessus dans le cas d'indépendance et le cas de dépendance sous mélange fort. D'un point de vue théorique nous ferons appel à des outils probabilistes, lesquels, combinés aux hypothèses sur la distribution de la variable fonctionnelle nous permettront l'obtention des résultats asymptotiques.

Le dernier sera consacré au côté pratique en simulant la fonction de régression dans les cas indépendant et fort mélangeant en utilisant le logiciel Matlab.

## Chapitre 1

# les outils probabilistes

On va présenter en bref quelques outils probabilistes. Parmi ces outils, ceux qui sont reformuler dans des nouveaux types afin de les rendre simplement applicable pour les modèles non paramétrique fonctionnels. Ces nouvelles formulations seront également utiles pour toute personne intéressée pour le développement de nouvelles avancées sur l'étude asymptotiques en statistiques fonctionnels non paramétriques.

Le fil conducteur de ce mémoire consiste à présenter des récents développements dans le cas des variables fonctionnelles. Cependant, l'obtention des résultats asymptotiques nécessite l'utilisation des outils de probabilité de base pour variables aléatoires réelles et de nombreux résultats présentés ci-dessous concernent les variables aléatoires réelles.

La première section de ce chapitre traite la notion de convergence presque complète et met l'accent sur le lien entre ce mode de convergence et d'autres modes standards (tels que la convergence presque sûre ou la convergence en probabilité). L'étude des propriétés de la convergence presque complète repose principalement sur certaines inégalités exponentielles

pour des sommes de variables aléatoires. La deuxième section rappelle certaines de ces inégalités ayant une forme adaptée au type de développements théoriques. La dernière section de ce chapitre sera consacrée aux conditions de mélange des variables aléatoires (réelles ou fonctionnelles). Plus précisément, dans la dernière partie de cette section nous présenterons quelques inégalités pour la somme des variables aléatoires réelles mélangeantes. En outre, comme pour la deuxième section, nous avons choisi parmi la large littérature de ces inégalités, celles ayant une forme adaptée au cadre de ce mémoire, certaines de ces inégalités ont été reformulée dans de nouveaux types, afin de rendre leur application plus facile.

Il est impossible, dans ce mémoire, de donner les preuves de tous ces outils probabilistes, et nous référerons principalement à la littérature existante.

Tout au long de ce chapitre,  $(X_n)_{n \in \mathbb{N}}$  et  $(Y_n)_{n \in \mathbb{N}}$  sont deux suites de variables aléatoires réelles,  $(u_n)$  est une suite de nombres réels positifs. On note que  $(\xi_n)_{n \in \mathbb{Z}}$  soit une suite de variables aléatoires (non nécessairement réelles), et  $(T_n)_{n \in \mathbb{Z}}$  soit une suite stationnaires de variables aléatoires réelles. On note aussi que  $(Z_n)_{n \in \mathbb{N}}$  soit une suite de v.a.r. indépendantes et centrées, et  $(W_n)_{n \in \mathbb{N}}$  soit une suite stationnaires de v.a.r. dépendantes et centrées.

## 1.1 La convergence presque complète

Le mode de convergence presque complète implique les autres modes standards de convergence. Par conséquent, en raison de ces deux avantages [C84], il est devenu tout à fait habituel pour les modèles non paramétriques fonctionnels, d'exprimer les résultats asymptotiques par la notion de convergence presque complète. Cette convergence a été



présentée depuis longtemps dans [HR47].

Dans cette section, nous avons décidé de rappeler quelques définitions et propriétés de base au sujet de cette notion sans démonstration (on peut consulter toutes les preuves dans [F]).

**Définition 1** *On dit que la suite  $(X_n)_{n \in \mathbb{N}}$  converge presque complète vers la variable aléatoire réelle  $X$ , si et seulement si*

$$\forall \epsilon > 0, \sum_{n \in \mathbb{N}} P(|X_n - X| > \epsilon) < \infty,$$

et on note la convergence presque complète de  $(X_n)_{n \in \mathbb{N}}$  vers  $X$  par :

$$\lim_{n \rightarrow \infty} X_n = X, p.co.$$

**Définition 2** *On dit que  $(X_n)_{n \in \mathbb{N}}$  converge en probabilité vers une variable aléatoire réelle  $X$ , si et seulement si*

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0,$$

et on écrit :

$$\lim_{n \rightarrow \infty} X_n = X, p$$

**Définition 3** *On dit que  $(X_n)_{n \in \mathbb{N}}$  converge presque sûrement vers une variable aléatoire réelle  $X$ , si et seulement si*

$$\lim_{n \rightarrow \infty} X_n = X, p.s \Leftrightarrow P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1.$$

On peut trouver dans des livres de probabilité élémentaire une présentation plus générale des divers liens entre ces modes de convergence (on peut consulter [BL87] par exemple). Les preuves de ces propriétés peuvent également être trouvées dans [BL87].

**Proposition 4** Si  $\lim_{n \rightarrow \infty} X_n = X, p.co.$ , alors on a :

$$i) \lim_{n \rightarrow \infty} X_n = X, p.$$

$$ii) \lim_{n \rightarrow \infty} X_n = X, p.s.$$

Selon nos connaissances, il n'existe pas de notion équivalente à la notion de convergence complète. Le but de la définition suivante est de préciser cette notion.

**Définition 5** On dit que la vitesse de convergence presque complète de  $(X_n)_{n \in \mathbb{N}}$  vers  $X$  est d'ordre  $u_n$  si et seulement si

$$\exists \epsilon_0 > 0, \sum_{n \in \mathbb{N}} P(|X_n - X| > \epsilon_0 u_n) < \infty$$

et on écrit

$$X_n - X = O(u_n), p.co.$$

Nous pouvons trouver d'autres points de vues au sujet de manière à mesurer un tel type de vitesse de convergence (voir par exemple [HSV98] et [HV02]).

Cette nouvelle définition est de double intérêts. D'une part, elle mérite de donner une définition précise et formelle qui est intéressante d'un point de vue probabiliste puisque elle implique la vitesse de convergence en probabilité  $O_p$  et la vitesse de convergence presque sûre  $O_{p.s}$ , d'autre part, elle est intéressante d'un point de vue statistique puisque elle facilite la preuve de  $O_p$  et de  $O_{p.s}$ .

**Proposition 6** Supposons que  $X_n - X = O(u_n), p.co.$  On a :

$$i) X_n - X = O(u_n), p.$$

$$ii) X_n - X = O(u_n), p.s.$$

Maintenant, on va présenter dans la proposition 7 et 8 quelques règles de calcul élémentaires concernant le mode de convergence presque complète.

**Proposition 7** *Supposons que  $\lim_{n \rightarrow \infty} u_n = 0$ ,  $\lim_{n \rightarrow \infty} X_n = l_x, p.co.$ , et  $\lim_{n \rightarrow \infty} Y_n = l_y, p.co.$ , où  $l_x$  et  $l_y$  sont deux nombres réels déterminés :*

*i) On a :*

$$a) \lim_{n \rightarrow \infty} X_n + Y_n = l_x + l_y, p.co.$$

$$b) \lim_{n \rightarrow \infty} X_n Y_n = l_x l_y, p.co.$$

$$c) \lim_{n \rightarrow \infty} \frac{1}{Y_n} = \frac{1}{l_y}, p.co. \text{ avec } l_y \neq 0$$

*ii) Si  $X_n - l_x = O(u_n), p.co.$  et  $Y_n - l_y = O(u_n), p.co.$  on a :*

$$a) (X_n + Y_n) - (l_x + l_y) = O(u_n), p.co.$$

$$b) X_n Y_n - l_x l_y = O(u_n), p.co.$$

$$c) \frac{1}{Y_n} - \frac{1}{l_y} = O(u_n), p.co. \text{ avec } l_y \neq 0$$

**Proposition 8** *Supposons que  $\lim_{n \rightarrow \infty} u_n = 0$ ,  $X_n = O(u_n), p.co.$  et  $\lim_{n \rightarrow \infty} Y_n = l_y, p.co.$ , où  $l_y$  est un nombre réel déterminé.*

$$i) \text{ On a : } X_n Y_n = O(u_n), p.co.$$

$$ii) \text{ On a : } \frac{X_n}{Y_n} = O(u_n), p.co. \text{ avec } l_y \neq 0.$$

## 1.2 Les inégalités exponentielles pour les v.a.r indépendantes.

Dans cette section, on considère que  $Z_1, Z_2, \dots, Z_n$  sont des variables aléatoires réelles centrées et indépendantes. Comme nous pouvons voir tout au long de ce travail, l'énoncé des propriétés de convergence presque complète a besoin de trouver une borne

supérieure pour la probabilité de la somme des v.a.r tel que :

$$P \left( \left| \sum_{i=1}^n Z_i \right| > \epsilon \right)$$

où  $\epsilon$  est un réel positif qui décroît avec  $n$ .

Dans ce contexte, ils existent des outils probabilistes puissants appelés Inégalités exponentielles et la littérature fait état de plusieurs versions de ces inégalités qui diffèrent selon les différentes hypothèses vérifiées par les variables  $Z_i$ . Nous nous concentrons ici sur les inégalités exponentielles de type Bernstein, ce choix est dû au fait que la forme de l'inégalité de type Bernstein est facile et plus adaptée aux développements théoriques de statistique fonctionnelle qui vont être énoncés tout au long de ce mémoire. D'autres formes de ces inégalités peuvent être trouvées dans [FN71], [N97] et [N98].

**Proposition 9** *Soit  $A_n = a_1^2 + a_2^2 + \dots + a_n^2$ . et supposons que*

$$\forall m \geq 2, |EZ_i^m| \leq (m!/2) (a_i)^2 b^{m-2},$$

*Alors on a :*

$$\forall \epsilon \geq 0, P \left( \left| \sum_{i=1}^n Z_i \right| > \epsilon A_n \right) \leq 2 \exp \left( - \frac{\epsilon^2}{2a^2 \left( 1 + \frac{\epsilon b}{A_n} \right)} \right).$$

Les preuves de la proposition 7 ci dessus sont données dans [U37] et [B46]. Comme nous pouvons également trouver la preuve des résultats plus généraux dans [Y76].

Notons que cette inégalité est énoncée pour les variables aléatoires réelles non identiquement distribuées. Notons également que chaque variable  $Z_i$  peut dépendre de  $n$ . C'est pour ces deux raisons que nous introduisons le corollaire 10 qui est plus utilisé que la proposition générale 7 car il est plus adapté au cadre de notre travail.

**Corollaire 10** *i) Si*

$$\forall m \geq 2, \exists C_m > 0, E |Z_1^m| \leq C_m a^{2(m-1)}, \text{ on a : } \forall \epsilon \geq 0, P \left( \left| \sum_{i=1}^n Z_i \right| > \epsilon n \right) \leq 2 \exp \left\{ -\frac{\epsilon^2 n}{2a^2(1+\epsilon)} \right\}$$

*ii) Supposons que les variables aléatoires sont dépendantes de  $n$  (c.à.d  $Z_i = Z_{i,n}$ ). Si*

$$\forall m \geq 2, \exists C_m > 0, E |Z_1^m| \leq C_m a_n^{2(m-1)} \text{ et si } u_n = \frac{a_n^2 \log n}{n} \text{ vérifie } \lim_{n \rightarrow \infty} u_n = 0,$$

*on a :*

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{u_n}), \text{ p.co.}$$

Notons aussi que toutes les inégalités précédentes sont fournies pour des variables aléatoires non bornées utilisées dans la régression fonctionnelle non paramétrique. Naturellement, elles s'appliquent directement pour les variables bornées, comme c'est le cas pour la densité conditionnelle fonctionnelle. C'est pour cette raison qu'on va présenter une nouvelle version du corollaire précédent, qui est directement adapté aux variables bornées.

**Corollaire 11** *i) Si  $\exists M < \infty, |Z_1| < M$  et en faisant la notation  $\sigma^2 = EZ_1^2$  on a :*

$$\forall \epsilon \geq 0, P \left( \left| \sum_{i=1}^n Z_i \right| > \epsilon n \right) \leq 2 \exp \left\{ -\frac{\epsilon^2 n}{2\sigma^2(1+\epsilon \frac{M}{\sigma^2})} \right\}$$

*ii) Supposons que les variables aléatoires sont dépendantes de  $n$  (c.à.d  $Z_i = Z_{i,n}$ ) telles que*

$$\exists M = M_n < \infty, |Z_1| < M \text{ et en faisant la notation } \sigma_n^2 = EZ_1^2. \text{ Si } u_n = \frac{a_n^2 \log n}{n} \text{ vérifie}$$

$\lim_{n \rightarrow \infty} u_n = 0$ , et si  $M/\sigma_n^2 < C < \infty$ , alors on a :

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{u_n}), \text{ p.co.}$$

### 1.3 Les inégalités exponentielles pour les v.a.r $\alpha$ -mélangeante.

La statistique non-paramétrique pour les processus réels mélangeants a suscité beaucoup d'attention pendant plusieurs décennies, il s'avère que le progrès statistique dans ce domaine a été lié directement au développement des outils probabilistes pour les suites mélangeantes. Principalement, il ya deux types d'outils qui sont utilisés en statistique non-paramétrique : les inégalités exponentielles et les inégalités de covariance. Le but de cette section est de rappeler quelques unes de ces inégalités pour les processus réels  $\alpha$ -mélangeant. Il faut noter qu'au moyen de la proposition 12 ci-dessous, ces inégalités pour les variables réelles seront utilisées pour des variables définies dans des espaces semi-métriques.

Commençons d'abord par rappeler quelques inégalités de covariance. Il faut savoir qu'il ya une littérature large au sujet des inégalités de covariance pour les variables mélangeantes. Ces inégalités diffèrent selon le type de condition de mélange présenté et le type d'hypothèses vérifiées par les variables. La preuve de chaque résultat présenté peut être trouvée dans [RI87] ou [Y92]. Des preuves plus courtes sont également données dans le *chapitre 1* de [Y94] et le *chapitre 1* de [R00]. Ces inégalités s'appellent également les inégalités du moment (voir par exemple [K94] [CK95] ou [EQV02] pour des extensions récentes des inégalités du moment d'ordre plus élevé, utiles dans la statistique non-paramétrique). Nous nous concentrons ici sur des structures de dépendance  $\alpha$ -mélangeante et nous donnons juste dans la proposition ci-dessous deux inégalités de covariance pour des variables aléatoires bornées et non bornées. D'après nos connaissances, le premier résultat ci-dessous a été donné dans [I62] tandis que le second a été énoncé dans [D68].

**Proposition 12** *Soit  $(T_n)_{n \in \mathbb{Z}}$  une suite stationnaire de variables aléatoires réelles. Suppo-*



sons que  $(T_n)_{n \in \mathbb{Z}}$  est  $\alpha$ -mélangeante, pour  $k \in \mathbb{Z}$ , on considère la variable réelle  $T$  (resp.  $T'$ ) qui est  $A_{-\infty}^k$ -mesurable (resp.  $A_{n+k}^{+\infty}$ -mesurable). *i)* Si  $T$  et  $T'$  sont bornées, alors :

$$\exists C, 0 < C < \infty, \text{cov}(T, T') \leq C\alpha(n)$$

*ii)* Si, pour les nombres positifs  $p, q, r$  tel que  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ , on a  $ET^p < \infty$  et  $ET'^q < \infty$ , alors :

$$\exists C > 0, \text{cov}(T, T') \leq C (ET^p)^{\frac{1}{p}} (ET'^q)^{\frac{1}{q}} \alpha(n)^{\frac{1}{r}}$$

Maintenant on va présenter quelques inégalités exponentielles pour les sommes partielles d'une suite  $(W_n)_{n \in \mathbb{N}}$  de variables aléatoires réelles  $\alpha$ -mélangeantes stationnaires et centrées. Dans un certain sens, même si les formes ne sont pas complètement comparables, les résultats présentés ci-dessous, sont des extensions au cas dépendant de ceux présentés dans le cas d'indépendance dans la deuxième section.

Tout au long des vingt dernières années, La littérature sur les inégalités exponentielles pour des suites mélangeantes a été directement liée aux développements théoriques des données dépendantes en statistique fonctionnelles non-paramétriques. Cette connexion a commencé avec les inégalités précédentes de type Bernstein données par [B75] [C84] dans une structure de mélange plus restrictive que celle du  $\alpha$ -mélange. D'après nos connaissances, la première inégalité exponentielle pour les variables  $\alpha$ -mélangeantes est due à [C83]. Ce précédent résultat a été amélioré dans plusieurs autres travaux et nous pouvons trouver dans [B93] et [Ro96] une large discussion de la bibliographie sur ce point. Pour l'objectif de notre travail, on a décidé d'utiliser une version au cas dépendant de l'inégalité de Fukk Nagaev qui a été précédemment présenté par [FN71] pour des variables indépendantes et qui a été simplifiée dans [N97] et [N98]. Un rappel de cette inégalité est donnée dans la pro-

position 13. Pour les raisons discutées le long de la section 1.3, on va présenter un résultat pour les variables bornées et un pour les non bornées et pour simplifier les écritures, nous énoncerons le résultat sans indiquer les expressions exactes des constantes impliquées dans les limites des intervalles.

En faisant maintenant la notation :

$$s_n^2 = \sum_{i=1}^n \sum_{j=1}^n |\text{cov}(W_i, W_j)|.$$

**Proposition 13** *Supposons que  $(W_n)_{n \in \mathbb{N}^*}$  est une suite de variables aléatoires identiquement distribuées et algébriquement  $\alpha$ -mélangeante avec le coefficient de mélange d'ordre  $a > 1$ .*

i) S'il existe  $p > 2$  et  $M > 0$  tel que  $\forall t > M, P(|W_1| > t) \leq t^{-p}$ , alors on a pour tout  $r \geq 1$  et  $\epsilon > 0$  et pour  $C < \infty$  :

$$P\left(\left|\sum_{i=1}^n W_i\right| > \epsilon\right) \leq C \left\{ \left(1 + \frac{\epsilon^2}{rs_n^2}\right)^{-r/2} + nr^{-1} \left(\frac{r}{\epsilon}\right)^{(a+1)p/(a+p)} \right\}.$$

ii) S'il existe  $M < \infty$  tel que  $|W_1| \leq M$ , alors on a pour tout  $r \geq 1$  et pour  $C < \infty$  :

$$P\left(\left|\sum_{i=1}^n W_i\right| > \epsilon\right) \leq C \left\{ \left(1 + \frac{\epsilon^2}{rs_n^2}\right)^{-r/2} + nr^{-1} \left(\frac{r}{\epsilon}\right)^{(a+1)} \right\}$$

Le résultat fourni dans la proposition 13 est énoncé et prouvé dans [R00] dans un cadre plus général que le notre. Pour rendre plus facile l'utilisation de chaque résultat probabiliste, soit pour nous soit pour toute personne intéressé par d'autres développements, nous proposons d'autres formulations de cette inégalité (voir le corollaire 14 et le corollaire 15) qui sont adaptées directement aux applications de la statistique non paramétrique. Ces deux corollaires diffèrent selon le type des coefficients de mélange utilisé : algébrique ou géométrique.

**Corollaire 14** *Supposons que les variables sont dépendantes de  $n$  (c.à.d  $W_i = W_{i,n}$ ), et que  $W_1, \dots, W_n$  sont les termes d'une suite  $\alpha$ -mélangeante avec les coefficients algébriques d'ordre  $a > 1$ . on considère la suite  $u_n = \frac{s_n^2 \log n}{n^2}$ . Supposons aussi que l'une des deux hypothèses suivantes est satisfaite :*

$$\begin{aligned} \exists p > 2, \exists \theta > 2, \exists M = M_n < \infty, \text{ tel que} \\ \forall t > M_n, P(|W_1| > t) \leq t^{-P} \text{ et } s_n^{-\frac{(a+1)p}{(a+p)}} = o(n^{-\theta}), \end{aligned}$$

ou

$$\begin{aligned} \exists M = M_n < \infty, \exists \theta > 2, \text{ tel que} \\ |W_1| \leq M_n \text{ et } s_n^{(a+1)} = o(n^{-\theta}). \end{aligned}$$

Alors on a :

$$\frac{1}{n} \sum_{i=1}^n W_i = O(\sqrt{u_n}), p.co.$$

**Corollaire 15** *Supposons que les variables sont dépendantes de  $n$  (c.à.d  $W_i = W_{i,n}$ ), et que  $W_1, \dots, W_n$  sont les termes d'une suite  $\alpha$ -mélangeante avec les coefficients géométriques. On considère la suite  $u_n = \frac{s_n^2 \log n}{n^2}$ . Supposons aussi que l'une des deux hypothèses suivantes est satisfaite :*

$$\begin{aligned} \exists p > 2, \exists \theta > \frac{2}{p}, \exists M = M_n < \infty, \text{ tel que} \\ \forall t > M_n, P(|W_1| > t) \leq t^{-P} \text{ et } s_n^{-1} = o(n^{-\theta}), \end{aligned}$$

ou

$$\begin{aligned} \exists M = M_n < \infty, \exists \theta > 2, \text{ tel que} \\ |W_1| \leq M_n \text{ et } s_n^{-1} = o(n^{-\theta}). \end{aligned}$$

*Alors on a :*

$$\frac{1}{n} \sum_{i=1}^n W_i = O(\sqrt{u_n}), p.co.$$

## Chapitre 2

# Pondération locale des variables fonctionnelles

Dans le cas de dimension finie, les techniques de pondération locales sont très populaires dans la communauté des nonparamétriciens parce qu'ils sont très bien adaptées à des modèles non-paramétriques. Le but de ce chapitre est d'expliquer comment le concept de lissage local peut être étendu au cas de données fonctionnelles. De toute évidence, les approches locales doivent avoir à main quelques structures topologiques permettant de mesurer la proximité entre les données fonctionnelles, et donc ce chapitre sera directement lié à la modélisation semi-métrique dans le cas de dimension finie. L'une des approches les plus courantes parmi ces méthodes de pondération locales est certainement celle du noyau. Il est impossible de donner une bibliographie exhaustive sur les méthodes non paramétriques en cas dimension finie, mais l'état de l'art dans ce domaine est bien résumé dans [S00] et [AP03] tandis qu'un grand nombre de références peuvent être trouvées dans [SV00] concernant les

méthodes à noyaux en particulier. Nous allons voir dans ce chapitre comment l'idée de lissage par noyau peut être adaptée au cadre fonctionnelle. Le chapitre est organisé comme suit :

Dans la section 4.1 nous donnons une discussion de base sur la méthode du noyau, en expliquant comment (et pourquoi) ce qui est classiquement fait pour la dimension finie peut être adapté au cadre fonctionnelle. Le deuxième objectif (Section 4.2) consiste à voir comment la pondération locale est en relation avec la notion de probabilité de petite boule. Comme nous le verrons, les probabilités de petite boule peuvent être considérées comme un outil pour décrire certains comportements locaux de données fonctionnelles et l'approche du noyau nous permet de prendre en considération le type de ces propriétés locales. La section 4.3 propose quelques aspects théoriques générales concernant la pondération du noyau. Enfin, notons que les sections 4.1 et 4.2 sont d'un grand intérêt pour un grand public, tandis que la section 4.3 est destinée à des statisticiens intéressés par l'aspect théorique.

## **2.1 Pourquoi utiliser des méthodes du noyau pour les données fonctionnelles ?**

Les méthodes du noyau sont bien connues et utilisées intensivement par la communauté des non-paramétriciens parce qu'ils sont un moyen utile de faire la pondération locale. Nous commençons brièvement par un rappel sur la pondération du noyau local dans le cas réel et multivariée avant de l'étendre au contexte fonctionnel.



### 2.1.1 Cas réel

Comme il est bien connu, la pondération du noyau local est basée sur une fonction du noyau ( noté  $K$  ) et d'un paramètre de lissage, et généralement désigné par  $h$ . Si  $x$  est un nombre réel fixe, la pondération du noyau local transforme  $n$  v.a.r  $X_1, X_2, \dots, X_n$  à  $n$  quantités  $\Delta_1, \Delta_2, \dots, \Delta_n$

$$\Delta_i = \Delta_i(x, h, K) = \frac{1}{h} K \left( \frac{d(x, X_i)}{h} \right)$$

L'idée principale de la pondération locale autour de  $x$  est d'attribuer à chaque variable aléatoire  $X_i$  un poids en tenant compte de la distance entre  $x$  et  $X_i$ ; Plus que  $X_i$  est éloignée de  $x$ , la pondération est plus petite. Avant de continuer, rappelons-nous ce qui est une fonction noyau exactement dans cette situation la plus simple. En fait, il existe une grande variété de noyaux. Toute la densité de fonction peut être considérée comme noyau, mais même les fonctions positives peuvent être utilisées comme des noyaux (voir [GM84]). Une vaste littérature existe sur ce domaine (voir [MN89] et [B93] pour des avancées intéressantes et [HVZ] pour une présentation de l'état de l'art). Pour simplifier notre propos, nous considérons à ce stade seulement les noyaux symétriques et positives qui sont les plus classiques. Plusieurs types de noyaux sont couramment utilisés, ils sont définis analytiquement comme suit :

Uniforme :

$$K(u) = \frac{1}{2} \mathbf{1}(|u| \leq 1)$$

Triangle :

$$K(u) = (1 - |u|) \mathbf{1}(|u| \leq 1)$$

Epanechnikov :

$$K(u) = \frac{3}{4} (1 - u^2) 1(|u| \leq 1)$$

Quadratique :

$$K(u) = \frac{15}{16} (1 - u^2)^2 1(|u| \leq 1)$$

Cubique :

$$K(u) = \frac{35}{32} (1 - u^2)^3 1(|u| \leq 1)$$

Gaussien :

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

### 2.1.2 Cas multivarié

Dans des situations de cas multivarié l'observation de  $n$  vecteurs aléatoires  $X_1, X_2, \dots, X_n$  sont à valeurs dans  $\mathbb{R}^p$ . La pondération du noyau locale peut être facilement étendu à cette situation. Enfin, il suffit de considérer un noyau multivarié  $K^*$  qui sera une fonction de  $\mathbb{R}^p$  dans  $\mathbb{R}$  et qui peut être défini comme un produit de  $p$  fonctions du noyau réels  $K_1, \dots, K_p$  :

$$\forall u = {}^t (u_1, \dots, u_p) \in \mathbb{R}^p, K^*(u) = K_1(u_1) \times K_2(u_2) \times \dots \times K_p$$

Maintenant, on va discuter sur l'interprétation de ce cas en termes de pondération locale.

En effet, ce qui arrive est très similaire au cas réel. Soit  $x$  un vecteur fixe de  $\mathbb{R}^p$ . La pondération du noyau multivarié locale consiste à transformer  $n$  vecteurs aléatoires  $X_1, X_2, \dots, X_n$  à  $n$  variables  $\Delta_1, \Delta_2, \dots, \Delta_n$  :

$$\Delta_i = \frac{1}{h^p} K^* \left( \frac{x - X_i}{h} \right)$$

Si l'on considère les noyaux à support compact, il apparaît clairement que  $\Delta_i$  sont des transformations localement pondérées des variables  $X_i$ , donc  $\Delta_i = 0$  si  $X_i$  correspondante est hors de voisinage de  $x$ . Par ailleurs, la normalisation  $1/h^p$  est proportionnel au volume de l'ensemble sur lequel les  $x_i$  sont prises en compte.

### 2.1.3 Cas fonctionnel

L'arrière-plan présenté ci-dessus est suffisante pour introduire la pondération du noyau local dans le cas fonctionnelle. Soit  $X_1, X_2, \dots, X_n$   $n$  variables aléatoires fonctionnelles à valeurs dans  $E$ . Et soit  $x$  un élément fixe de  $E$ . L'idée de l'extension fonctionnelle vient de la pondération naïve du noyau locale multivarié qui sert à transformer  $X_1, X_2, \dots, X_n$  à  $n$  quantités de forme

$$\Delta_i = \frac{1}{V(h)} K \left( \frac{d(x, X_i)}{h} \right), i = 1, \dots, n$$

où  $d$  est une semi métrique de  $E$ ,  $K$  est un noyau (symétrique) et  $V(h)$  est le volume de la boule

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

Cependant cette approche appelle à la définition de  $V(h)$ , autrement dit ceci a besoin d'avoir une mesure dans  $E$ . Et c'est ici que la différence entre les deux cas réel et multivarié apparait où la mesure de Lebesgue est implicitement introduite, alors que dans l'espace fonctionnel on ne peut utiliser aucune mesure. C'est pour cette raison qu'on choisit une mesure particulière construite par la normalisation en utilisant la distribution de probabilité

de v.a.f. Les variables fonctionnelles du poids local du noyau sont définies par :

$$\Delta_i = \frac{K\left(\frac{d(x, X_i)}{h}\right)}{E\left(K\left(\frac{d(x, X_i)}{h}\right)\right)}$$

**Remarque 16** En cas multivarié, sous les conditions standards de  $K$  on a :

$$E\left(K\left(\frac{d(x, X_i)}{h}\right)\right) = f(x) h^p$$

où  $f$  est la densité de  $X_i$  avec la mesure de Lebesgue.

Notons que les noyaux utilisés sont : le noyau en boîte symétrique, le noyau triangulaire symétrique, le noyau quadratique symétrique et le noyau gaussien symétrique. Pour faciliter leurs utilisations, on considère deux types de noyaux pour les variables fonctionnelles :

i) La fonction  $K$  de  $\mathbb{R}$  dans  $\mathbb{R}^+$  tel que  $\int K = 1$  est appelée noyau de type *I* s'il existe deux constantes réelles  $0 < C_1 < C_2$  tel que :

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

ii) La fonction  $K$  de  $\mathbb{R}$  dans  $\mathbb{R}^+$  tel que  $\int K = 1$  est appelée noyau de type *II* si elle est de support  $[0, 1]$  et si sa dérivée  $K'$  existe sur  $[0, 1]$  et vérifie pour deux constantes  $-\infty < C_2 < C_1 < 0$  :

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

Le noyau de type *I* représente le noyau discontinu usuel (c'est la boîte symétrique), alors que le noyau de type *II* représente le noyau continu symétrique standard comme le triangulaire, le quadratique et le gaussien. En fin pour le poids local des variables aléatoires réelles le noyau est de type 0 :

**Définition 17** La fonction  $K$  de  $\mathbb{R}$  dans  $\mathbb{R}^+$  tel que  $\int K = 1$  avec le support compact  $[-1, 1]$  et tel que  $\forall u \in (0, 1), K(u) > 0$  est appelée noyau de type 0.

## 2.2 Le lien entre le poids local et les petites boules de probabilités

On considère un noyau de type  $I$  qui est le noyau en boîte symétrique. Soit  $X$  la v.a.f à valeur dans  $E$  et soit  $x$  un point fixe de  $E$ . On peut écrire :

$$E \left( 1_{[0,1]} \left( \frac{d(x, X)}{h} \right) \right) = E (1_{B(x,h)}(X)) = P(X \in B(x, h))$$

En s'inspirant de

$$\Delta_i = \frac{K \left( \frac{d(x, X_i)}{h} \right)}{E \left( K \left( \frac{d(x, X_i)}{h} \right) \right)}$$

la probabilité des boules  $B(x, h)$  apparait clairement dans la normalisation. On va voir que le paramètre de lissage  $h$  (qu'on appelle aussi la largeur de la fenêtre) décroît avec la taille de l'échantillon des variables fonctionnelles ( $h \rightarrow 0$ , lorsque  $n \rightarrow \infty$ ). Donc, lorsque  $n$  est suffisamment grand,  $h$  est proche de 0. C'est pour cette raison,  $B(x, h)$  est considérée comme petite boule.

Maintenant pour tout  $x$  dans  $E$  et pour tout réel positif  $h$ , on utilise la notation suivante :

$$\varphi_x(h) = P(X \in B(x, h))$$

Cette notion de probabilité des petites boules joue un rôle important dans les deux formes théorique et pratique. Comme la notion des boules est liée uniquement à la semi métrique  $d$ , donc le choix de la semi métrique est important. Et on va voir que la vitesse de convergence

des estimateurs fonctionnels non-paramétriques est en relation avec  $d$  et du comportement de la probabilité des petites boules autour de 0.

### 2.3 Quelques rappels 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 18** *Si  $K$  est un noyau de type I, alors il existe deux nombres réels finis positifs  $C$  et  $C'$*

*tel que :*

$$C\varphi_x(h) \leq EK\left(\frac{d(x, X)}{h}\right) \leq C'\varphi_x(h)$$

**Lemme 19** *Si  $K$  est un noyau de type II et si  $\varphi_x(\cdot)$  vérifie :*

$$\exists C_3 > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > C_3\epsilon\varphi_x(\epsilon),$$

*alors il existe deux nombres réels finis positifs  $C$  et  $C'$  tel que pour  $h$  suffisamment petit on a :*

$$C\varphi_x(h) \leq EK\left(\frac{d(x, X)}{h}\right) \leq C'\varphi_x(h)$$



## Chapitre 3

# Méthodologies de prévision non paramétrique fonctionnelle

Ce chapitre décrit plusieurs approches relatives à la prédiction non-paramétrique de certaines réponses scalaires. Le cadre fonctionnel apparaît à travers la variable explicative fonctionnelle. Nous nous concentrons sur trois méthodes de prévision complémentaires, à savoir l'espérance conditionnelle, la médiane conditionnelle et le mode conditionnel. L'espérance conditionnelle est liée à la régression alors que la médiane conditionnelle et le mode conditionnel sont fortement liés à l'estimation de la distribution conditionnelle. Après l'introduction de certains aspects bibliographiques (section 3.1), nous présentons dans la section 3.2 les trois méthodes de prédiction fonctionnelle non- paramétrique. La section 3.3 présente les modèles non paramétriques associés à ces problèmes de prédiction, tandis que Section 3.4 met l'accent sur la construction des estimateurs.

Il existe de nombreuses situations dans lesquelles on peut étudier le lien entre

deux variables, avec l'objectif principal d'être en mesure de prédire de nouvelles valeurs de l'un des données de l'autre. Ce problème de prédiction a été largement étudié dans la littérature lorsque les deux variables sont dans le cas de dimension finie. Bien sûr, le même problème peut se produire lorsque certaines variables sont fonctionnelles. Notre souhait est d'étudier ce problème lorsque la variable explicative est fonctionnelle et la variable réponse est toujours réelle. Afin de fixer les idées et mettre l'accent sur le grand intérêt et l'utilité de ce problème dans de nombreux domaines des sciences appliquées, on donne un exemple sur les données chimométriques.

Soit un échantillon statistique (de taille  $n = 215$ ) composé de courbes de spectrométrie  $X_1, \dots, X_n$  (Ce sont des données fonctionnelles) correspondant aux spectres observés pour 215 pièces fines de viande hachée. En outre, par un procédé chimique analytique, nous avons mesuré le contenu de graisse de chaque pièce  $Y_1, \dots, Y_n$  (Ce sont les réponses scalaires).

Ainsi, nous recueillons les observations d'une réponse scalaire (la teneur en graisse) et une variable explicative fonctionnelle (spectres). Une question qui peut se poser est la suivante : étant observé quelques spectres d'un morceau de viande, peut-on prédire sur son contenu correspondant à la graisse ? Ceci est typiquement un problème de prédiction fonctionnelle. Pour répondre à cette question, il convient d'estimer le lien entre la teneur en graisses et les spectres. Malheureusement, il n'y a ni un moyen d'afficher cette relation ni une information sur la structure de ce sujet. Par conséquent, il devient naturel d'introduire les modèles non-paramétriques afin de déterminer la forme du lien. L'aspect fonctionnel du problème est aussi très important. En utilisant les courbes spectrométriques, la continuité

et d'autres caractéristiques fonctionnelles des spectres doivent être prises en compte. Il est donc nécessaire d'avoir besoin de développer des méthodes combinant les deux concepts non-paramétriques et modélisation des variables fonctionnelles. Bien sûr, il existe une littérature conforme à la fois autour de prédiction non-paramétrique et des données fonctionnelles. Mais, jusqu'à présent, les variables fonctionnelles ont été étudiées essentiellement dans un cadre paramétrique. Cela a été popularisé par [RS97] (principalement pour les points de vue pratiques) et les développements théoriques peuvent être trouvés dans [B00] dans le contexte spécifique des variables fonctionnelles dépendantes. Les récentes avancées concrètes peuvent être trouvées par exemple dans [CGS04] alors que certaines études asymptotiques sont détaillées dans [CFS03] et [CFF02].

Dans une autre direction de la recherche statistique, les problèmes de prévision non paramétrique ont été étudiés de manière intensive à la fois dans des cas réel et multivarié.

Il est impossible de donner une description exhaustive de la bibliographie liée, mais pour fixer les idées, le lecteur pourrait regarder le précurseur des œuvres de [W64] et [N64], lors de l'enquête intermédiaire par [C85] et à [S00] ou [AP03] pour une description de l'état de l'art.

Le but de ce travail est de marier les avantages du libre-modélisation avec une méthodologie entièrement fonctionnelle afin de répondre au problème de prédiction fonctionnelle tel que celui de spectrométrie. Nous présentons trois approches statistiques fonctionnelles et non paramétriques pour le problème de prédiction.

L'intéressé doit prendre en considération le modèle à étudier et le fait que dans la désignation de méthode fonctionnelle non-paramétrique de prévision, le mot fonctionnelle se

réfère à la notion de variable fonctionnelle (implicitement : "Nous devons prendre en compte la caractéristique fonctionnelle de la variable ") tandis que le mot non paramétrique signifie que nous utilisons une modélisation de libre-paramètre pour les opérateurs non-linéaires d'être estimer. En outre, il est important de noter que notre méthodologie est également basée sur le libre-modélisation de la distribution puisque aucune hypothèse paramétrique est nécessaire pour la distribution des variables aléatoires.

### 3.1 Diverses approches au problème de prédiction

On considère les hypothèses suivantes :

Soit l'échantillon  $(X_i, Y_i)_{i=1, \dots, n}$  identiquement distribués et indépendantes, à valeurs dans  $\mathcal{F} \times \mathbb{R}$  où  $(\mathcal{F}, d)$  est un espace semi-métrique. Soit  $x$  (*resp*  $y$ ) un élément fixé de  $\mathcal{F}$  (*resp*  $\mathbb{R}$ ), soit  $S$  un sous espace de  $\mathbb{R}$  fixé. Etant donné  $x$ , on appelle  $\hat{y}$  la valeur de prédiction du scalaire réponse.

La régression  $r$  de  $Y$  en  $x$  est définie par :

$$r(x) = E(Y/X = x) \quad (3.1)$$

La fonction de répartition conditionnelle de  $Y$  en  $x$  est définie par :

$$\forall y \in \mathbb{R}, F_Y^X(x, y) = P(Y \leq y | X = x) \quad (3.2)$$

En outre, supposons que  $F_Y^X(x, \cdot)$  est différentiable en  $y$  donc on peut écrire :

$$\forall y \in \mathbb{R}, f_Y^X(x, y) = \frac{\partial}{\partial y} F_Y^X(x, y) \quad (3.3)$$

$f_Y^X(x, y)$  est la valeur de la fonction de densité en  $(x, y)$ .

Il est clair que chaque opérateur non-linéaire cité donne une information sur la relation entre  $X$  et  $Y$ .

Le premier nous permet de poser :

$$\hat{y} = \hat{r}(x) \quad (3.4)$$

où  $\hat{r}(x)$  est l'estimateur de  $r(x)$

Le deuxième sert à considérer la médiane conditionnelle  $m(x)$  de  $F_Y^X$  :

$$m(x) = \inf \left\{ y \in \mathbb{R}, F_Y^X(x, y) \geq \frac{1}{2} \right\} \quad (3.5)$$

et d'utiliser le prédicteur :

$$\hat{y} = \hat{m}(x) \quad (3.6)$$

où  $\hat{m}(x)$  est l'estimateur de  $m(x)$ .

Le mode conditionnel  $\theta(x)$  de la fonction de densité conditionnelle de  $Y$  sachant  $X$  est basé sur le troisième opérateur :

$$\theta(x) = \arg \sup_{y \in \mathcal{S}} f_Y^X(x, y) \quad (3.7)$$

Le prédicteur est :

$$\hat{y} = \hat{\theta}(x) \quad (3.8)$$

où  $\hat{\theta}(x)$  est l'estimateur de  $\theta(x)$ .

Les quantiles sont définés pour  $\alpha \in ]0, 1[$  par :

$$t_\alpha(x) = \inf \{ y \in \mathbb{R}, F_Y^X(x, y) \geq \alpha \} \quad (3.9)$$

donc, l'estimateur  $\hat{t}_\alpha(x)$  de  $t_\alpha(x)$ , se construit de l'intervalle :

$$[\hat{t}_\alpha(x), \hat{t}_{1-\alpha}(x)] \quad (3.10)$$

pour  $\alpha \in ]0, 1/2[$ .

### 3.2 Présentation des modèles non paramétriques pour la prévision.

Dans tout ce qui suit, on considère deux types de modèles : modèles sous condition de continuité et modèles de type Lipschitz.

Pour l'opérateur de régression (opérateur non-linéaire de  $\mathcal{F}$  dans  $\mathbb{R}$ ) on considère les modèles suivants :

$$r \in C_{\mathcal{F}}^0 \quad (3.11)$$

où

$$C_{\mathcal{F}}^0 = \left\{ f : \mathcal{F} \rightarrow \mathbb{R}, \lim_{d(x,x') \rightarrow 0} f(x') = f(x) \right\}$$

et  $\exists \beta > 0$  tel que :

$$r \in Lip_{\mathcal{F},\beta} \quad (3.12)$$

avec

$$Lip_{\mathcal{F},\beta} = \left\{ f : \mathcal{F} \rightarrow \mathbb{R}, \exists C \in \mathbb{R}_+^*, \forall x' \in E, |f(x') - f(x)| < C d(x, x')^\beta \right\}.$$

Pour l'opérateur de la médiane conditionnelle, qui est basé sur l'opérateur  $F_Y^X$  (opérateur non-linéaire de  $\mathcal{F} \times \mathbb{R}$  dans  $\mathbb{R}$ ), on suppose qu'il est lié à l'ensemble suivant :

$$S_{cdf}^x = \{f : \mathcal{F} \times \mathbb{R} \rightarrow \mathbb{R}, f(x, \cdot) \text{ est strictement croissante}\} \quad (3.13)$$

En effet  $F_Y^X \in S_{cdf}^x$  assure l'existence et l'unicité de la médiane conditionnelle

définit comme suit :

$$m(x) = F_Y^{x-1} \left( \frac{1}{2} \right)$$

où

$$\begin{aligned} F_Y^x & : \mathbb{R} \rightarrow [0, 1] \\ y & \rightarrow F_Y^x(y) = F_Y^X(x, y) \end{aligned} \quad (3.14)$$

On considère les ensembles suivants :

$$C_{\mathcal{F} \times \mathbb{R}}^0 = \left\{ \begin{array}{l} f : \mathcal{F} \times \mathbb{R} \rightarrow \mathbb{R}, \forall x' \in N_x, \\ \lim_{d(x, x') \rightarrow 0} f(x', y) = f(x, y) \\ \text{et } \forall y' \in \mathbb{R}, \lim_{|y - y'| \rightarrow 0} f(x', y) = f(x, y) \end{array} \right\} \quad (3.15)$$

et

$$Lip_{\mathcal{F} \times \mathbb{R}, \beta} = \left\{ \begin{array}{l} f : \mathcal{F} \times \mathbb{R} \rightarrow \mathbb{R}, \\ \forall (x_1, x_2) \in N_x^2, \forall (y_1, y_2) \in S^2, \\ |f(x_1, y_1) - f(x_2, y_2)| \leq C \left( d(x_1, x_2)^\beta + |y_1 - y_2|^\beta \right) \end{array} \right\} \quad (3.16)$$

Nous pouvons donc définir deux modèles non-paramétriques fonctionnels de l'opérateur de la fonction de distribution fonctionnelle comme suit :

$$F_Y^X \in C_{\mathcal{F} \times \mathbb{R}}^0 \cap S_{cdf}^x \quad (3.17)$$

et  $\exists \beta > 0$  tel que :

$$F_Y^X \in Lip_{\mathcal{F} \times \mathbb{R}, \beta} \cap S_{cdf}^x \quad (3.18)$$

La prédiction concernant le mode conditionel repose sur l'opérateur de la fonction de densité conditionnelle  $f_Y^X$  (opérateur non linéaire de  $\mathcal{F} \times \mathbb{R}$  dans  $\mathbb{R}$ ). On introduit l'ensemble

suivant :

$$S_{dens}^X = \left\{ \begin{array}{l} f : \mathcal{F} \times \mathbb{R} \rightarrow \mathbb{R}, \\ \exists \xi > 0, \exists y_0 \in S, f(x, \cdot) \text{ est strictement} \\ \text{croissante sur } (y_0 - \xi, y_0) \\ \text{et strictement croissante sur } (y_0 - \xi, y_0) \end{array} \right\} \quad (3.19)$$

alors, si  $f_Y^X \in S_{dens}^X$  alors selon le problème de maximisation de  $f_Y^X(x, y)$  sur  $S$ , elle admet une solution unique  $y_0$ . Donc le mode conditionnel peut être défini de la façon suivante :

$$\theta(x) = \arg \sup_{y \in S} f_Y^X(x, y), \quad (3.20)$$

$f_Y^X \in S_{dens}^X$  assure l'unicité de  $\theta(x)$  et nous avons deux modèles non-paramétriques pour la prédiction du mode conditionnel. Soit

$$f_Y^X \in C_{\mathcal{F} \times \mathbb{R}}^0 \cap S_{dens}^X \quad (3.21)$$

Ou

$$f_Y^X \in Lip_{\mathcal{F} \times \mathbb{R}, \beta} \cap S_{dens}^X \quad (3.22)$$

**Remarque 20** Les modèles (2.11), (2.17), (2.21) s'appellent modèles non-paramétriques fonctionnels de type continuité, ils nous permettent d'étudier la convergence presque complète alors que (2.12), (2.18), (2.22) sont de type Lipschitz et nous permettent d'étudier la vitesse de convergence presque complète.

### 3.3 Présentation des estimateurs à noyau

Une fois le modèle non paramétrique est introduit, on va estimer les divers objets mathématiques exposés dans les modèles ci-dessus, c'est à dire les opérateurs  $r$ ,  $F_Y^X$ , et  $f_Y^X$ .



### 3.3.1 L'estimateur de régression.

Nous proposons pour l'opérateur non linéaire  $r$  l'estimateur à noyau suivant :

$$\hat{r}(x) = \frac{\sum_{i=1}^n Y_i K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))} \quad (3.23)$$

où  $K$  est un noyau non symétrique et  $h$  un réel strictement positif dépendant de  $n$ , c'est ce qu'on appelle le paramètre de lissage ou la largeur de la fenêtre. Cet estimateur est l'extension de l'estimateur de Nadaraya-Watson développé par Ferraty et Vieu (pour plus de discussion voir[FV06]) .  $d$  est la semi-métrique qui mesure la proximité entre les données fonctionnelles. Si on considère les quantités suivantes :

$$w_{i,h}(x) = \frac{K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}$$

ainsi

$$\hat{r}(x) = \sum_{i=1}^n w_{i,h}(x) Y_i \quad (3.24)$$

qui est vraiment la moyenne pondérée car :

$$\sum_{i=1}^n w_{i,h}(x) = 1$$

### 3.3.2 Estimateur de la fonction de répartition conditionnelle.

Pour l'estimateur  $\hat{F}_Y^X$ , on s'appuie sur l'idée utilisée pour la construction du noyau de l'opérateur de régression.

On a

$$F_Y^X(x, y) = P(Y \leq y / X = x)$$

qui s'écrit :

$$F_Y^X(x, y) = E(1_{]-\infty, y]}(Y) / X = x)$$

d'où :

$$F_Y^X(x, y) = \frac{\sum_{i=1}^n 1_{]-\infty, y]}(Y_i) K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}$$

Il suffit de changer la fonction indicatrice. Soit  $K_0$  un noyau symétrique usuel, soit

$H$  défini comme suit :

$$\forall u \in \mathbb{R}, H(u) = \int_{-\infty}^u K_0(v) dv \quad (3.25)$$

On définit l'estimateur à noyau de la fonction de répartition conditionnelle comme

suit :

$$\hat{F}_Y^X = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i)) H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))} \quad (3.26)$$

où  $g$  est un réel strictement positif dépendant de  $n$ .

A partir de (2.26) on définit l'estimateur de la médiane conditionnelle comme suit :

$$\hat{m}(x) = \inf \left\{ y \in \mathbb{R}, \hat{F}_Y^X(x, y) \geq \frac{1}{2} \right\} \quad (3.27)$$

et d'une manière générale l'estimateur des quantiles conditionnels sont donnés pour tout

$\alpha \in [0, \frac{1}{2}]$  par :

$$\hat{t}_\alpha(x) = \inf \left\{ y \in \mathbb{R}, \hat{F}_Y^X(x, y) \geq \alpha \right\} \quad (3.28)$$

### 3.3.3 Estimateur de la densité conditionnelle.

On peut écrire :

$$\hat{f}_Y^X(x, y) = \frac{\partial}{\partial y} F_Y^X(x, y) \quad (3.29)$$

en tenant compte de la différentiabilité de  $H$ , on obtient :

$$\frac{\partial}{\partial y} F_Y^X(x, y) = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i)) \frac{\partial}{\partial y} H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}$$

et l'estimateur du mode conditionnelle peut être défini comme suit :

$$\hat{\theta}(x) = \arg \sup_{x \in S} \hat{f}_Y^X(x, y) \tag{3.30}$$

## Chapitre 4

# Etude Asymptotique

## Première partie

# Cas d'indépendance

## 4.1 La convergence presque complète.

### 4.1.1 Estimation de la régression

On considère le modèle suivant :

$$Y = r(X) + \epsilon \text{ avec } E(\epsilon | X) = 0 \quad (4.31)$$

On considère les hypothèses suivantes :

$$\forall \epsilon > 0, P(X \in B(x, \epsilon)) = \varphi_x(\epsilon) > 0 \quad (4.32)$$

$$\left\{ \begin{array}{l} h \text{ est une suite positive telle que : } \lim_{n \rightarrow \infty} h = 0 \text{ et } \lim_{n \rightarrow \infty} \frac{\log n}{n\varphi_x(h)} = 0, \\ K \text{ est un noyau de type I (ou de type II) et} \\ \exists C > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) d(u) > C\epsilon\varphi_x(\epsilon) \end{array} \right\} \quad (4.33)$$

Pour la variable réponse  $Y$ , on considère

$$\text{Pour tout } m \geq 2 : E(|Y^m| / X = x) < \sigma_m(x) < \infty \text{ avec } \sigma_m(\cdot) \text{ continue en } x \quad (4.34)$$

**Théorème 21** *Sous les conditions (3.11) et (4.32) – (4.34) on a :*

$$\lim_{n \rightarrow \infty} \hat{r}(x) = r(x), p.co \quad (4.35)$$

**Preuve.** Posant : pour  $i = 1, \dots, n$

$$\Delta_i = \frac{K(h^{-1}d(x, X_i))}{EK(h^{-1}d(x, X_i))}$$

En utilisant les lemmes 16, 17 et (2.32) on arrive à :  $EK(h^{-1}d(x, X_i)) > 0$ .

Soit

$$\hat{r}_1(x) = \frac{1}{n} \sum_{i=1}^n \Delta_i \text{ et } \hat{r}_2(x) = \frac{1}{n} \sum_{i=1}^n Y_i \Delta_i \quad (4.36)$$

On a :

$$\hat{r}(x) = \frac{\hat{r}_2(x)}{\hat{r}_1(x)}$$

En faisant la décomposition suivante :

$$\begin{aligned} \hat{r}(x) - r(x) &= \frac{\hat{r}_2(x)}{\hat{r}_1(x)} - r(x) \\ &= \frac{1}{\hat{r}_1(x)} [\hat{r}_2(x) - \hat{r}_1(x)r(x)] \\ &= \frac{1}{\hat{r}_1(x)} [(\hat{r}_2(x) - E\hat{r}_2(x)) - (r(x) - E\hat{r}_2(x))] - \frac{r(x)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \end{aligned}$$

d'où

$$\hat{r}(x) - r(x) = \frac{1}{\hat{r}_1(x)} [(\hat{r}_2(x) - E\hat{r}_2(x)) - (r(x) - E\hat{r}_2(x))] - \frac{r(x)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \quad (4.37)$$

■

**Lemme 22** *Sous les hypothèses (3.11) et (4.33) on a :*

$$\lim_{n \rightarrow \infty} E\hat{r}_2(x) = r(x)$$

**Preuve.** On a :

$$\begin{aligned} r(x) - E\hat{r}_2(x) &= r(x) - E(Y_1\Delta_1) \\ &= r(x) - E(E(Y_1\Delta_1/X_1)) \\ &= r(x) - E(\Delta_1 r(X_1)) \\ &= E((r(x) - r(X_1))\Delta_1) \end{aligned}$$

Puisque le support de  $K$  est  $[0, 1]$  on a :

$$(r(x) - r(X_1))\Delta_1 \leq \sup_{x' \in B(x, h)} (r(x) - r(x'))\Delta_1$$

Puisque  $r$  est continue on aura :

$$\lim_{n \rightarrow \infty} E\hat{r}_2(x) = r(x)$$

■

**Lemme 23** *i) Sous les conditions (4.32), (4.33) et (4.34) on a :*

$$\hat{r}_2(x) - E\hat{r}_2(x) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

*ii) Si les conditions (4.32) et (4.33) sont vérifiées on a :*

$$\hat{r}_1(x) - 1 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

**Preuve.** On pose :  $K_i = K(h^{-1}d(x, X_i))$  pour  $i = 1, \dots, n$ . La démonstration de ce résultat est basée sur l'utilisation d'une inégalité exponentielle de type Bernstein. En effet

$$P(|\hat{r}_2(x) - E\hat{r}_2(x)| > \varepsilon) = P\left(\frac{1}{n} \left| \sum_{i=1}^n (Y_i \Delta_i - E(Y_i \Delta_i)) \right| > \varepsilon\right)$$

En montrant qu'il existe  $\varepsilon_0 > 0$  tel que :

$$\sum_{n \geq 1} P\left(\frac{1}{n} \left| \sum_{i=1}^n (Y_i \Delta_i - E(Y_i \Delta_i)) \right| > \varepsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}\right) < \infty$$

En appliquant l'inégalité exponentielle donnée par le corollaire 8.ii à la variable

$$Z_i = Y_i \Delta_i - E(Y_i \Delta_i)$$

Premièrement on va montrer :

$$\exists C > 0, \forall m \geq 2 : |E(Y_i \Delta_i - E(Y_i \Delta_i))^m| \leq C\varphi_x(h)^{-m+1} \quad (4.38)$$



Montrons d'abords que pour  $m \geq 2$  :

$$E(|Y_1|^m \Delta_1^m) = O(\varphi_x(h)^{-m+1}) \quad (4.39)$$

En effet puisque  $\Delta_1^m = \frac{K_1^m}{(EK_1)^m}$  on aura :

$$\begin{aligned} E(|Y_1|^m \Delta_1^m) &= \frac{1}{(EK_1)^m} [E|Y_1|^m K_1^m] \\ &= \frac{1}{(EK_1)^m} [E(E(|Y_1|^m | X) K_1^m)] \\ &\leq \frac{1}{(EK_1)^m} [E \sigma_m(X) K_1^m] \\ &\leq \frac{1}{(EK_1)^m} [E((\sigma_m(X) - \sigma_m(x)) K_1^m) + \sigma_m(x) EK_1^m] \\ &\leq E((\sigma_m(X) - \sigma_m(x)) \Delta_1^m) + \sigma_m(x) E\Delta_1^m \end{aligned}$$

Ceci implique :

$$\begin{aligned} |E|Y_1|^m \Delta_1^m| &\leq E|\sigma_m(X) - \sigma_m(x)| \Delta_1^m + \sigma_m(x) E\Delta_1^m \\ &\leq \left( \sup_{x' \in B(x, h)} |\sigma_m(X) - \sigma_m(x)| \right) E\Delta_1^m + \sigma_m(x) E\Delta_1^m. \end{aligned}$$

Puisque  $0 < \int K^m < \infty$  donc si  $K$  est de type  $I$  (resp.  $II$ ) alors  $K^m / \int K^m$  est aussi de type  $I$  (resp.  $II$ ). Par suite en appliquant le lemme 16 ou le lemme 17 on arrive à :

$$C_1 \varphi_x(h) \leq EK_1^m \leq C_2 \varphi_x(h). \quad (4.40)$$

En utilisant (4.40) et le lemme 16 ou 17 on peut écrire pour  $m \geq 2$  :

$$\frac{C_1}{\varphi_x(h)^{m-1}} \leq E\Delta_1^m \leq \frac{C_2}{\varphi_x(h)^{m-1}}. \quad (4.41)$$

En conséquence on a :

$$E(|Y_1|^m \Delta_1^m) = O(\varphi_x(h)^{-m+1}).$$

De plus on a :

$$(Y_1\Delta_1 - EY_1\Delta_1)^m = \sum_{k=0}^m c_{m,k} (Y_1\Delta_1)^k (EY_1\Delta_1)^{m-k} (-1)^{m-k},$$

avec  $c_{m,k} = m!/(k!(m-k)!)$  ce qui implique :

$$\begin{aligned} E|Y_1\Delta_1 - EY_1\Delta_1|^m &\leq \sum_{k=0}^m c_{m,k} E|Y_1\Delta_1|^k (E|Y_1\Delta_1|)^{m-k} \\ &\leq \sum_{k=0}^m c_{m,k} E|Y_1\Delta_1|^k (E|E(Y_1\Delta_1/X = x)|)^{m-k} \\ &\leq \sum_{k=0}^m c_{m,k} E|Y_1\Delta_1|^k (E|\Delta_1 r(x)|)^{m-k} \\ &\leq \sum_{k=0}^m c_{m,k} E|Y_1\Delta_1|^k (|r(x)|)^{m-k} \\ &\leq C \max_{k=0,1,\dots,m} E|Y_1\Delta_1|^k \\ &\leq C \max_{k=0,1,\dots,m} \varphi_x(h)^{-k+1} \end{aligned}$$

cette dernière inégalité est obtenu en faisant (4.39) pour  $k \geq 2$ , alors que pour  $k = 1$  on montre que

$$E|Y_1| \Delta_1 = O(1)$$

en effet

$$\begin{aligned} E(|Y_1\Delta_1|) &= E(|E(Y_1\Delta_1 | X_1 = x)|) \\ &= E(|\Delta_1 r(X_1 = x)|) \\ &= |r(x)| \\ &\leq C \end{aligned}$$

Puisque  $\varphi_x(h) \rightarrow 0$ , quand  $n \rightarrow \infty$ , alors :

$$E|Y_1\Delta_1 - EY_1\Delta_1|^m = O\left((\varphi_x(h))^{-m+1}\right).$$

En appliquant le corollaire 8 – ii, en posant  $a^2 = \varphi_x(h)^{-1}$  on aura :

$$u_n = (a^2 \log n) / n = \log n / (n\varphi_x(h))$$

qui tend vers 0 quand  $n \rightarrow \infty$  (en utilisant (4.33)) alors

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n Z_i &= O(\sqrt{u_n}) \text{ p.co} \\ &= O\left(\sqrt{\log n / n\varphi_x(h)}\right), \text{ p.co} \end{aligned}$$

Ce qui prouve que :

$$\hat{r}_2(x) - E\hat{r}_2(x) = O\left(\sqrt{\log n / n\varphi_x(h)}\right), \text{ p.co}$$

ii) En reprenant le même raisonnement ci-dessus, en faisant  $Y_i = 1$  on arrive à :

$$\hat{r}_1(x) - 1 = O\left(\sqrt{\log n / n\varphi_x(h)}\right) \text{ p.co.}$$

■

### Estimation de la médiane conditionnelle.

Puisque les résultats asymptotiques sont donnés pour  $x$  fixé dans  $\mathcal{F}$ , on utilise les notations simplifiées suivantes :

$$\forall x \in \mathcal{F}, F_Y^x(\cdot) \stackrel{\text{def}}{=} F_Y^X(x, \cdot) \text{ et } \hat{F}_Y^x(\cdot) \stackrel{\text{def}}{=} \hat{F}_Y^X(x, \cdot) \quad (4.42)$$

Par convention, on utilise souvent la notation suivante :

$$\forall x \in \mathcal{F}, F_Y^X(\cdot) \stackrel{\text{def}}{=} F_Y^X(X, \cdot) \quad (4.43)$$

On note que la partie fonctionnelle de  $\hat{F}_Y^X$  est identique à celle de la régression, donc les hypothèses nécessaires pour le traitement de cette partie sont celles considérés pour

la régression. Pour la partie scalaire de  $\hat{F}_Y^X$  on considère la condition suivante sur le noyau  $K_0 = H'$  et la largeur de la fenêtre  $g$  :

$$\left\{ \begin{array}{l} g \text{ est une suite positive tel que : } \lim_{n \rightarrow \infty} g = 0, \\ K_0 \text{ est de type 0.} \end{array} \right\} \quad (4.44)$$

Notant que cette condition sur le noyau assure avec la définition 15 que la fonction  $H$  est continue et strictement croissante sur l'ensemble  $\{u/0 < K(u) < 1\}$ . Ceci nous permet de définir l'estimateur de la médiane conditionnelle comme solution unique de l'équation :

$$\hat{m}(x) = \hat{F}_Y^x^{-1}\left(\frac{1}{2}\right) \quad (4.45)$$

**Théorème 24** *Sous le modèle défini par (3.17) et (4.32), et si l'estimateur à noyau vérifie (4.33) et (4.44), on a :*

$$\lim_{n \rightarrow \infty} \hat{m}(x) = m(x) \text{ p.co} \quad (4.46)$$

**Preuve.** La condition (4.44) assure que l'estimateur  $\hat{F}_Y^x(\cdot)$  est continu et strictement croissant. Donc la fonction  $\hat{F}_Y^x^{-1}(\cdot)$  existe et continue. La continuité de  $\hat{F}_Y^x^{-1}(\cdot)$  au point  $\hat{F}_Y^x(m(x))$  nous permet d'écrire :

$$\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0, \forall y, \left| \hat{F}_Y^x(y) - \hat{F}_Y^x(m(x)) \right| \leq \delta(\varepsilon) \implies |y - m(x)| \leq \varepsilon.$$

Dans le cas particulier où  $y = \hat{m}(x)$ , on aura :

$$\begin{aligned} \forall \varepsilon > 0, \exists \delta(\varepsilon) > 0, \forall y, \left| \hat{F}_Y^x(\hat{m}(x)) - \hat{F}_Y^x(m(x)) \right| &\leq \delta(\varepsilon) \\ \implies |\hat{m}(x) - m(x)| &\leq \varepsilon. \end{aligned}$$

$$\begin{aligned} \text{D'où} \quad &: \quad \forall \varepsilon > 0, \exists \delta(\varepsilon) > 0, \forall y, P(|\hat{m}(x) - m(x)| > \varepsilon) \\ &\leq P\left(\left|\hat{F}_Y^x(\hat{m}(x)) - \hat{F}_Y^x(m(x))\right| > \delta(\varepsilon)\right) \\ &= P\left(\left|F_Y^x(m(x)) - \hat{F}_Y^x(m(x))\right| > \delta(\varepsilon)\right) \end{aligned}$$

La dernière égalité est obtenu par :  $F_Y^x(m(x)) = \hat{F}_Y^x(\hat{m}(x)) = 1/2$ .

La convergence *p.co* ponctuelle de  $F_Y^x$  (voir le lemme 23) nous amène à :

$$\forall \varepsilon > 0, \sum_{n \geq 1} P(|\hat{m}(x) - m(x)| > \varepsilon) < \infty.$$

Et maintenant le résultat (4.46) est vérifié. ■

**Lemme 25** *Sous les conditions du théorème 19, on a pour tout nombre réel fixé  $y$  :*

$$\lim_{n \rightarrow \infty} \hat{F}_Y^x(y) = F_Y^x(y), p.co \quad (4.47)$$

**Preuve.** On peut écrire :

$$\begin{aligned} \hat{F}_Y^x(y) - F_Y^x(y) &= \frac{\hat{r}_1(x) \hat{F}_Y^x(y)}{\hat{r}_1(x)} - F_Y^x(y) \\ &= \frac{1}{\hat{r}_1(x)} [\hat{r}_3(x, y) - \hat{r}_1(x) F_Y^x(y)] \\ &= \frac{1}{\hat{r}_1(x)} [(\hat{r}_3(x, y) - E\hat{r}_3(x, y)) - (F_Y^x(y) - E\hat{r}_3(x, y))] \\ &\quad - \frac{F_Y^x(y)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \end{aligned}$$

d'où

$$\begin{aligned} \hat{F}_Y^x(y) - F_Y^x(y) &= \frac{1}{\hat{r}_1(x)} [(\hat{r}_3(x, y) - E\hat{r}_3(x, y)) - (F_Y^x(y) - E\hat{r}_3(x, y))] \\ &\quad - \frac{F_Y^x(y)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \end{aligned} \quad (4.48)$$

Où  $\hat{r}_1$  est définie par (4.36) et

$$\hat{r}_3(x, y) = \hat{r}_1(x) \hat{F}_Y^x(y) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Gamma_i(y)$$

avec

$$\Gamma_i(y) = H(g^{-1}(y - Y_i))$$

On a :

$$\hat{r}_1(x) \rightarrow 1 \text{ quand } n \rightarrow \infty. \frac{F_Y^x(y)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \rightarrow 0 \text{ quand } n \rightarrow \infty.$$

Il nous reste à montrer que :

$$\lim_{n \rightarrow \infty} E\hat{r}_3(x, y) = F_Y^x(y) \quad (4.49)$$

et

$$\lim_{n \rightarrow \infty} \hat{r}_3(x, y) - E\hat{r}_3(x, y) = 0p.co. \quad (4.50)$$

On a :

$$\begin{aligned} E\hat{r}_3(x, y) - F_Y^x(y) &= E\Delta_1 \Gamma_1(y) - F_Y^x(y) \\ &= E(\Delta_1(E(\Gamma_1(y)/X))) - E\Delta_1 F_Y^x(y) \\ &= E(\Delta_1(E(\Gamma_1(y)/X) - F_Y^x(y))) \\ &= E(\Delta_1 1_{B(x, h)}(X) (E(\Gamma_1(y)/X) - F_Y^x(y))) \end{aligned}$$

d'où

$$E\hat{r}_3(x, y) - F_Y^x(y) = E(\Delta_1 1_{B(x, h)}(X) (E(\Gamma_1(y)/X) - F_Y^x(y))) \quad (4.51)$$

En utilisant le théorème de Fubini et le fait que  $H \nu = K_0$ , pour calculer (2.51). En effet :

$$\begin{aligned}
E(\Gamma_1(y)/X) &= \int_{\mathbb{R}} H(g^{-1}(y-u)) dP(u/X) \\
&= \int_{\mathbb{R}} \int_{-\infty}^{\frac{y-u}{g}} K_0(v) dv dP(u/X) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} K_0(v) 1_{[v, +\infty[}(g^{-1}(y-u)) dv dP(u/X) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} K_0(v) 1_{[v, +\infty[}(g^{-1}(y-u)) dP(u/X) dv \\
&= \int_{\mathbb{R}} K_0(v) F_Y^X(y - vg) dv.
\end{aligned}$$

d'où

$$E(\Gamma_1(y)/X) = \int_{\mathbb{R}} K_0(v) F_Y^X(y - vg) dv. \quad (4.52)$$

Puisque  $\int K_0 = 1$ , on peut écrire :

$$E(\Gamma_1(y)/X) - F_Y^x(y) = \int_{\mathbb{R}} K_0(v) (F_Y^X(y - vg) - F_Y^x(y)) dv \quad (2.53)$$

Ainsi, on a :

$$\begin{aligned}
|F_Y^X(y - vg) - F_Y^x(y)| &\leq |F_Y^X(y - vg) - F_Y^x(y - vg)| \\
&\quad + |F_Y^x(y - vg) - F_Y^x(y)|
\end{aligned}$$

Puisque  $K_0$  est de support  $[-1, 1]$ ,  $g$  et  $h$  tendent vers 0, la continuité de  $F_Y^X(y)$

nous permet d'écrire :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, 1]} 1_{B(x, h)}(X) |F_Y^X(y - vg) - F_Y^x(y - vg)| = 0,$$

et

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, 1]} |F_Y^x(y - vg) - F_Y^x(y)| = 0$$

Donc

$$\lim_{n \rightarrow \infty} |E(\Gamma_1(y)/X) - F_Y^x(y)| = 0.$$

Par conséquent :

$$\lim_{n \rightarrow \infty} 1_{B(x,h)}(X) |E(\Gamma_1(y)/X) - F_Y^x(y)| = 0 \quad (4.54)$$

(4.51) est équivalent à

$$E(\Delta_1 (1_{B(x,h)}(X) E(\Gamma_1(y)/X) - F_Y^x(y)))$$

En utilisant (4.54) et  $E\Delta_1 = 1$ , on arrive à prouver (4.49)

Il nous reste à prouver (2.50). Pour cette raison, on pose :

$$\hat{r}_3(x, y) - E\hat{r}_3(x, y) = \frac{1}{n} \sum_{i=1}^n (T_i - ET_i) \quad (4.55)$$

avec

$$T_i = \Delta_i \Gamma_i$$

En utilisant le lemme 16 (ou le lemme 17) avec  $K$  est de type  $I$  (ou de type  $II$ ). Puisque  $T_i$  est borné on a :

$$T_i \leq C/\varphi_x(h) \quad (4.56)$$

D'autre part , pour  $m = 2$  le résultat (4.39) s'écrit :

$$E\left(|Y_1|^2 \Delta_1^2\right) = O\left(\varphi_x(h)^{-1}\right)$$

et on a :

$$ET_i^2 \leq CE(\Delta_1^2) \leq C/\varphi_x(h) \quad (4.57)$$



puisque les variables  $T_i$  sont bornées, les résultats (4.56) et (4.57) nous permet d'appliquer une inégalité exponentielle de type Bernstein donnée par le corollaire 9-ii, et on a :

$$\hat{r}_3(x, y) - E\hat{r}_3(x, y) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right), p.co. \quad (4.58)$$

Ce résultat prouve (4.50), et la démonstration du théorème est achevée. ■

#### 4.1.2 L'estimation du mode conditionnelle.

Puisqu'on s'intéresse à la convergence presque complète ponctuelle pour  $x$  fixé dans  $\mathcal{F}$ , on considère les notations simplifiées suivantes :

$$\forall x \in E, f_Y^x(\cdot) \stackrel{def}{=} f_Y^X(x, \cdot) \text{ et } \hat{f}_Y^x(\cdot) \stackrel{def}{=} \hat{f}_Y^X(x, \cdot) \quad (4.59)$$

Par convention, on utilise souvent la notation suivante :

$$\forall x \in E, f_Y^X(\cdot) \stackrel{def}{=} f_Y^X(X, \cdot) \quad (4.60)$$

Les hypothèses dont on a besoin pour  $\hat{f}_Y^X$  défini par (3.29) sont ceux introduites dans la section pour l'estimation de  $\hat{F}_Y^X$  en ajoutant l'hypothèse suivante :

$$\exists C < \infty, \forall (x, x') \in \mathbb{R} \times \mathbb{R}, |K_0(x) - K_0(x')| \leq C |x - x'|, \quad (4.61)$$

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

**Théorème 26** *Sous le modèle défini par (3.21) et (4.32), et si l'estimateur à noyau vérifie (4.33), (4.34) et (4.61) alors on a :*

$$\lim_{n \rightarrow \infty} \hat{\theta}(x) = \theta(x) p.co. \quad (4.62)$$

**Preuve.** La condition (3.21) assure que  $f_Y^x(\cdot)$  est continue et strictement croissante sur  $]\theta(x) - \zeta, \theta(x)[$ . donc  $f_Y^{x^{-1}}(\cdot)$  existe et continue. La continuité de  $f_Y^{x^{-1}}(\cdot)$  au point  $f_Y^x(\theta(x))$  nous permet d'écrire pour tout  $\epsilon > 0$

$$\exists \delta_1(\epsilon) > 0, \forall y \in ]\theta(x) - \zeta, \theta(x)[,$$

$$|f_Y^x(y) - f_Y^x(\theta(x))| \leq \delta_1(\epsilon) \implies |y - \theta(x)| \leq \epsilon.$$

De la même manière que précédemment, puisque  $f_Y^x(\cdot)$  est continue et strictement décroissante sur  $]\theta(x), \theta(x) + \zeta[$  alors pour tout  $\epsilon > 0$

$$\exists \delta_2(\epsilon) > 0, \forall y \in ]\theta(x), \theta(x) + \zeta[,$$

$$|f_Y^x(y) - f_Y^x(\theta(x))| \leq \delta_2(\epsilon) \implies |y - \theta(x)| \leq \epsilon.$$

En combinant ces deux derniers résultats on arrive à :

$$\exists \delta(\epsilon) > 0, \forall y \in ]\theta(x) - \zeta, \theta(x) + \zeta[,$$

$$|f_Y^x(y) - f_Y^x(\theta(x))| \leq \delta(\epsilon) \implies |y - \theta(x)| \leq \epsilon.$$

Puisque, par construction  $\hat{\theta}(x) \in ]\theta(x) - \zeta, \theta(x) + \zeta[$ , on a :

$$\exists \delta(\epsilon) > 0, \left| f_Y^x(\hat{\theta}(x)) - f_Y^x(\theta(x)) \right| \leq \delta(\epsilon) \implies \left| \hat{\theta}(x) - \theta(x) \right| \leq \epsilon,$$

D'après ce dernier résultat, on peut écrire :

$$\exists \delta(\epsilon) > 0, P\left(\left|\hat{\theta}(x) - \theta(x)\right| > \epsilon\right) \leq P\left(\left|f_Y^x(\hat{\theta}(x)) - f_Y^x(\theta(x))\right| > \delta(\epsilon)\right).$$

D'autre part, il vient des définitions de  $\theta(x)$  et  $\hat{\theta}(x)$  que :

$$\begin{aligned}
\left| f_Y^x(\hat{\theta}(x)) - f_Y^x(\theta(x)) \right| &= f_Y^x(\theta(x)) - f_Y^x(\hat{\theta}(x)) \\
&= \left( f_Y^x(\theta(x)) - \hat{f}_Y^x(\theta(x)) \right) + \left( \hat{f}_Y^x(\theta(x)) - f_Y^x(\hat{\theta}(x)) \right) \\
&\leq \left( f_Y^x(\theta(x)) - \hat{f}_Y^x(\theta(x)) \right) + \left( \hat{f}_Y^x(\hat{\theta}(x)) - f_Y^x(\hat{\theta}(x)) \right) \\
&\leq 2 \sup_{y \in ]\theta(x) - \zeta, \theta(x) + \zeta[} \left| f_Y^x(y) - \hat{f}_Y^x(y) \right|
\end{aligned}$$

d'où

$$\left| f_Y^x(\hat{\theta}(x)) - f_Y^x(\theta(x)) \right| \leq 2 \sup_{y \in ]\theta(x) - \zeta, \theta(x) + \zeta[} \left| f_Y^x(y) - \hat{f}_Y^x(y) \right| \quad (4.63)$$

On peut utiliser la convergence presque complète uniforme de l'estimateur de densité conditionnelle sur le compact  $[\theta(x) - \zeta, \theta(x) + \zeta]$  (voir le lemme 25). Ce qui nous amène directement avec les deux dernières inégalités à :

$$\forall \epsilon > 0, \sum_{n \geq 1} P \left( \left| \hat{\theta}(x) - \theta(x) \right| > \epsilon \right) < \infty$$

■

**Lemme 27** *Sous les conditions du théorème 24, on a pour tout compact  $S$  un sous ensemble de  $\mathbb{R}$  :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} \left| f_Y^x(y) - \hat{f}_Y^x(y) \right| = 0, p.co. \quad (4.64)$$

**Preuve.** La démonstration est basée sur la décomposition suivante :

$$\begin{aligned}
\hat{f}_Y^x(y) - f_Y^x(y) &= \frac{\hat{r}_1(x)f_Y^x(y)}{\hat{r}_1(x)} - f_Y^x(y) \\
&= \frac{1}{\hat{r}_1(x)} [\hat{r}_4(x, y) - \hat{r}_1(x)f_Y^x(y)] \\
&= \frac{1}{\hat{r}_1(x)} [(\hat{r}_4(x, y) - E \hat{r}_4(x, y)) - (f_Y^x(y) - E \hat{r}_4(x, y))] \\
&\quad - \frac{f_Y^x(y)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1]
\end{aligned}$$

d'où

$$\begin{aligned}
\hat{f}_Y^x(y) - f_Y^x(y) &= \frac{1}{\hat{r}_1(x)} [(\hat{r}_4(x, y) - E \hat{r}_4(x, y)) - (f_Y^x(y) - E \hat{r}_4(x, y))] \\
&\quad - \frac{f_Y^x(y)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1]
\end{aligned} \quad (4.65)$$

où  $\hat{r}_1$  est définie par (3.6) et

$$\hat{r}_4(x, y) = \hat{r}_1(x)f_Y^x(y) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Omega_i(y)$$

avec

$$\Omega_i(y) = g^{-1}K_0(g^{-1}(y - Y_i))$$

Le dénominateur du membre de droite de l'égalité de (4.65) est traité directement en utilisant le lemme 21-ii avec la proposition 6-i. Son dernier terme aussi se traite en faisant le lemme 21 - ii et  $f_Y^x(\cdot)$  est uniformément bornée pour  $y \in S$  (car  $f_Y^x(\cdot)$  est continue sur le compact  $S$ ). Donc le résultat de (4.64) sera une conséquence directe des deux résultats suivants :

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |f_Y^x(y) - E \hat{r}_4(x, y)| = 0, p.co. \quad (4.66)$$

et

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - E \hat{r}_4(x, y)| = 0, p.co. \quad (4.67)$$

Montrons le résultat (4.66).

Puisque  $E\Delta_1 = 1$  et  $\int K_0 = 1$ , on a :

$$\begin{aligned} E\hat{r}_4(x, y) - f_Y^x(y) &= E\Delta_1\Omega_1(y) - f_Y^x(y) \\ &= E(\Delta_1(E(\Omega_1(y) | X) - f_Y^x(y))) \\ &= E\left(\Delta_1\left(\int_{\mathbb{R}} g^{-1}K_0(g^{-1}(y-u))f_Y^X(u)du - f_Y^x(y)\right)\right) \\ &= E\left(\Delta_1\int_{\mathbb{R}} g^{-1}K_0(g^{-1}(y-u))(f_Y^X(u) - f_Y^x(y))du\right) \\ &= E\left(\Delta_1\int_{\mathbb{R}} K_0(v)(f_Y^X(y-vg) - f_Y^x(y))dv\right) \\ &= E\left(1_{B(x,h)}(X)\Delta_1\int_{\mathbb{R}} K_0(v)(f_Y^X(y-vg) - f_Y^x(y))dv\right) \end{aligned}$$

d'où

$$E\hat{r}_4(x, y) - f_Y^x(y) = E\left(1_{B(x,h)}(X)\Delta_1\int_{\mathbb{R}} K_0(v)(f_Y^X(y-vg) - f_Y^x(y))dv\right) \quad (4.68)$$

En plus , on peut écrire :

$$|f_Y^X(y-vg) - f_Y^x(y)| \leq |f_Y^X(y-vg) - f_Y^x(y-vg)| + |f_Y^x(y-vg) - f_Y^x(y)|.$$

Puisque  $S$  est compact, la fonction  $f_Y^x(y)$  est uniformément continue pour  $y \in S$ . De plus  $K_0$  est à support compact ( $K_0$  est de type 0), et puisque le modèle (3.21) est de type continue, nous avons :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1,1]} \sup_{y \in S} 1_{B(x,h)}(X) |f_Y^X(y-vg) - f_Y^x(y-vg)| = 0,$$

et

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1,1]} \sup_{y \in S} 1_{B(x,h)}(X) |f_Y^X(y-vg) - f_Y^x(y)| = 0,$$

ce qui implique :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1,1]} \sup_{y \in S} 1_{B(x,h)}(X) |f_Y^X(y - vg) - f_Y^x(y)| = 0 \quad (4.69)$$

La combinaison de (4.68) et (4.69) avec  $\Delta_1$  et  $K_0$  sont positifs nous donnent :

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |f_Y^x(y) - E\hat{r}_4(x, y)| = 0, \quad (4.70)$$

le résultat (4.70) avec le lemme 21 – ii et la proposition 6 – i suffit de prouver (4.66), et le traitement du terme biais est achevé.

Il nous reste à montrer que (4.67) est vrai. En effet puisque  $S$  est compacte, on peut écrire que  $S \subset \cup_{k=1}^{z_n} S_k$  où  $S_k = [t_k - l_n, t_k + l_n]$  et  $l_n$  et  $z_n$  vérifient :

$$l_n = Cz_n^{-1} \sim Cn^{-2\zeta}. \quad (4.71)$$

On prend  $t_y = \arg \min_{t \in \{t_1, \dots, t_{z_n}\}} |y - t|$ , on a :

$$\begin{aligned} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - E\hat{r}_4(x, y)| &\leq \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - \hat{r}_4(x, t_y)|}_{A_1} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, t_y) - E\hat{r}_4(x, t_y)|}_{A_2} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |E\hat{r}_4(x, t_y) - E\hat{r}_4(x, y)|}_{A_3} \end{aligned} \quad (4.72)$$

La condition de continuité de Hölder (4.61) nous permet d'écrire :

$$\begin{aligned} |\hat{r}_4(x, y) - \hat{r}_4(x, t_y)| &= \frac{1}{n} \sum_{i=1}^n \Delta_i |\Omega_i(y) - \Omega_i(t_y)| \\ &\leq \frac{1}{ng} \sum_{i=1}^n \Delta_i |K_0(g^{-1}(y - Y_i)) - K_0(g^{-1}(t_y - Y_i))| \\ &\leq \frac{C}{ng} \sum_{i=1}^n \Delta_i |g^{-1}(y - t_y)| \\ &\leq C\hat{r}_1(x)l_n g^{-2} \end{aligned}$$

En utilisant (4.71) on arrive :

$$A_1 \leq C/(gn^\zeta)^2$$

En utilisant (4.61) on obtient :

$$\lim_{n \rightarrow \infty} A_1 = 0 \quad (4.73)$$

De la même manière que précédemment :

$$\begin{aligned} E |\hat{r}_4(x, y) - \hat{r}_4(x, t_y)| &\leq \sum_{i=1}^n E \Delta_i |\Omega_i(y) - \Omega_i(t_y)| \\ &\leq \frac{1}{g} \sum_{i=1}^n E \Delta_i |K_0(g^{-1}(y - Y_i)) - K_0(g^{-1}(t_y - Y_i))| \\ &\leq \frac{1}{g} C \sum_{i=1}^n E \Delta_i |g^{-1}(y - t_y)| \\ &\leq nCl_n g^{-2} \end{aligned}$$

D'où :

$$\begin{aligned} A_3 &\leq \frac{nCl_n g^{-2}}{\hat{r}_1(x)} \quad (4.74) \\ &\leq \frac{nCn^{-2\zeta} g^{-2}}{\hat{r}_1(x)} \\ &\leq \frac{nC}{\hat{r}_1(x) (gn^\zeta)^2} \end{aligned}$$

avec le lemme 21-ii et la proposition 6-i on arrive à :

$$\lim_{n \rightarrow \infty} A_3 = 0 \text{ p.co.} \quad (4.75)$$

Pour le terme  $A_2$ , on peut écrire pour tout  $\epsilon > 0$  :

$$\begin{aligned} P \left( \sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon \right) &= P \left( \max_{j=1, \dots, z_n} |\hat{r}_4(x, t_j) - E \hat{r}_4(x, t_j)| > \epsilon \right) \quad (4.76) \\ &\leq z_n \max_{j=1, \dots, z_n} P (|\hat{r}_4(x, t_j) - E \hat{r}_4(x, t_j)| > \epsilon) \\ &\leq z_n \max_{j=1, \dots, z_n} P \left( \left| \frac{1}{n} \sum_{i=1}^n (U_i - EU_i) \right| > \epsilon \right) \end{aligned}$$

où

$$U_i = \Delta_i \Omega_i(t_j)$$

En utilisant le lemme 16 (ou le lemme 17) avec  $K$  soit de type  $I$  (ou de type  $II$ ), et puisque

$\Omega_i(y) \leq C/g$ , on a :

$$|U_i| \leq C/(g \varphi_x(h)) \quad (4.77)$$

D'autre part on a :

$$\begin{aligned} EU_i^2 &= E(\Delta_i^2 E(\Omega_i^2(t_j)/X = x)) \\ &= E\left(\Delta_i^2 \int_{\mathbb{R}} g^{-2} K_0^2(g^{-1}(t_j - u)) f_Y^x(u) du\right) \\ &= E\left(\Delta_i^2 g^{-1} \int_{\mathbb{R}} K_0^2(z) f_Y^x(t_j + zg) dz\right) \\ &\leq CE\left(\Delta_i^2 g^{-1} \int_{\mathbb{R}} K_0^2(z) f_Y^x(t_j) dz\right) \end{aligned}$$

Puisque  $f_Y^x$  est bornée (car elle est continue sur le compact  $S$ ) donc en appliquant le résultat

(4.39) pour  $m = 2$ , on obtient :

$$EU_i^2 \leq C/(g \varphi_x(h)) \quad (4.78)$$

Puisque les variables  $U_i$  sont bornées, on peut appliquer une inégalité exponentielle de type

Bernstein, cette inégalité avec (4.76), (4.77) et (4.78) nous donne :

$$P\left(\sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon\right) \leq z_n \exp(-Cn\epsilon^2 g \varphi_x(h)).$$

En utilisant (4.71) on obtient :

$$P\left(\sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon\right) \leq Cn^{2\zeta} \exp(-Cn\epsilon^2 g \varphi_x(h)).$$



puisque  $\frac{\log n}{n\varphi_x(h)}$  tend vers 0, on obtient :

$$\forall \epsilon > 0, \sum_{n \geq 1} P \left( \sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon \right) < \infty \quad (4.79)$$

Le dénominateur de  $A_2$  est traité en utilisant à la fois le lemme 21-ii et la proposition 6-i.

Par conséquent on obtient

$$\lim_{n \rightarrow \infty} A_2 = 0, p.co. \quad (4.80)$$

Finalement le résultat (4.67) est prouvé par (4.72), (4.73), (4.75) et (4.80).

La démonstration du lemme 25 est achevée par la combinaison de (4.65), (4.66) et (4.67). ■

### 4.1.3 Estimation du quantile conditionnelle.

Cette section est consacrée à la généralisation des résultats donnés dans la section concernant l'estimation de la médiane à l'estimation des quantiles conditionnels.

Le quantile d'ordre  $\alpha$ , noté par  $t_\alpha(x)$ , et son estimateur à noyau  $\hat{t}_\alpha(x)$  sont respectivement définis par (3.9) et (3.28). Sous la condition (4.44), l'estimateur de quantile conditionnel peut être défini par l'unique solution de l'équation

$$\hat{t}_\alpha(x) = \hat{F}_Y^{x-1}(\alpha) \quad (4.81)$$

**Théorème 28** *Soit  $\alpha \in [0, 1]$ . Sous le modèle défini par (3.17) et (4.32) et si l'estimateur à noyau vérifie (4.33) et (4.44) on a :*

$$\lim_{n \rightarrow \infty} \hat{t}_\alpha(x) = t_\alpha(x), p.co. \quad (4.82)$$

**Preuve.** Soit  $\varepsilon > 0$  fixé. La continuité de  $\hat{F}_Y^{x-1}$  au point  $\hat{F}_Y^x(t_\alpha(x))$  nous permet d'écrire :

$$\exists \delta(\varepsilon) > 0, \forall y, \left| \hat{F}_Y^x(y) - \hat{F}_Y^x(t_\alpha(x)) \right| \leq \delta(\varepsilon) \rightarrow |y - t_\alpha(x)| \leq \varepsilon.$$

En prenant  $y = \hat{t}_\alpha(x)$ , on arrive finalement au résultat suivant :

$$\begin{aligned} \exists \delta(\varepsilon) &> 0, P\left(|\hat{t}_\alpha(x) - t_\alpha(x)| > \varepsilon\right) \\ &\leq P\left(\left|\hat{F}_Y^x(\hat{t}_\alpha(x)) - \hat{F}_Y^x(t_\alpha(x))\right| > \delta(\varepsilon)\right) \\ &= P\left(\left|\hat{F}_Y^x(\hat{t}_\alpha(x)) - \hat{F}_Y^x(t_\alpha(x))\right| > \delta(\varepsilon)\right), \end{aligned}$$

En utilisant la convergence *p.co.* de l'estimateur de la fonction de répartition (voir le lemme 23) on obtient (4.82). ■

#### 4.1.4 Complément sur l'estimation de la distribution conditionnelle

Les résultats déjà indiqués pour l'estimation du mode conditionnel et les quantiles, sont basés sur des résultats auxiliaires de l'estimation de la distribution de probabilité conditionnelle de  $Y$  sachant  $X$ . Cette section est consacrée à la présentation générale de quelques résultats et de leurs régularités. La première proposition indique la convergence presque complète ponctuelle et uniforme de l'estimateur de la fonction de répartition conditionnelle. La deuxième indique la convergence presque complète ponctuelle et uniforme de l'estimateur de la fonction de densité conditionnelle.

**Proposition 29** *i) Sous les conditions (3.17), (4.32), (4.33) et (4.34), on a pour tout réel fixé  $y$  :*

$$\lim_{n \rightarrow \infty} \hat{F}_Y^x(y) = F_Y^x(y), \text{ p.co.} \quad (4.83)$$

ii) Si de plus la largeur de la fenêtre  $g$  vérifie pour tout  $\zeta > 0$  la condition  $\lim_{n \rightarrow \infty} gn^\zeta = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$  on a :

$$\lim_{n \rightarrow \infty} \sup_{y \in S} \left| F_Y^x(y) - \hat{F}_Y^x(y) \right| = 0, p.co \quad (4.84)$$

**Preuve.** Pour la démonstration de cette proposition, on utilise les mêmes notations que le lemme 23 et puisque le résultat i) est prouvé dans le lemme 23, donc il nous reste à montrer ii). En faisant aussi (4.48) et puisque le terme  $\hat{r}_1(x)$  ne dépend pas de  $y$  dans  $S$ , il est donc traité dans le lemme 23. Enfin il nous reste à montrer les deux résultats suivants :

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |E\hat{r}_3(x, y) - F_Y^x(y)| = 0 \quad (4.85)$$

et

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |E\hat{r}_3(x, y) - \hat{r}_3(x, y)| = 0, p.co. \quad (4.86)$$

Puisque  $K_0$  est de support  $[-1, 1]$  et  $g$  tend vers zéro, la continuité uniforme de  $F_Y^x$  sur le compact  $S$ , nous permet d'écrire :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, 1]} \sup_{y \in S} 1_{B(x, h)}(X) |F_Y^X(y - vg) - F_Y^x(y)| = 0.$$

Ce dernier résultat avec (4.51) et (4.53) sont suffisants pour prouver (4.85). Maintenant il nous reste à vérifier (4.86). La compacité de  $S$ , nous permet d'écrire que  $S \subset \bigcup_{k=1}^{z_n} S_k$  où  $S_k = [t_k - l_n, t_k + l_n]$  et  $l_n$  et  $z_n$  sont choisis de sorte que :

$$l_n = Cz_n^{-1} \sim Cn^{-\zeta} \quad (4.87)$$

En prenant  $t_y = \arg \min_{t \in \{t_1, \dots, t_{z_n}\}} |y - t|$ , on a :

$$\begin{aligned} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - E \hat{r}_3(x, y)| &\leq \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - \hat{r}_3(x, t_y)|}_{D_1} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, t_y) - E \hat{r}_3(x, t_y)|}_{D_2} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |E \hat{r}_3(x, t_y) - E \hat{r}_3(x, y)|}_{D_3} \end{aligned} \quad (4.88)$$

Puisque  $K_0 = H^{(1)}$  est supposé un noyau de type 0, donc  $H$  est différentiable et de dérivée bornée. Donc  $H$  vérifie la condition de continuité de Hölder d'ordre 1, alors on peut écrire :

$$\begin{aligned} |\hat{r}_3(x, y) - \hat{r}_3(x, t_y)| &= \frac{1}{n} \sum_{i=1}^n \Delta_i |\Gamma_i(y) - \Gamma_i(t_y)| \\ &= \frac{1}{n} \sum_{i=1}^n \Delta_i |H(g^{-1}(y - Y_i)) - H(g^{-1}(t_y - Y_i))| \\ &\leq \frac{C}{n} \sum_{i=1}^n \Delta_i |g^{-1}(y - t_y)| \\ &\leq C \hat{r}_1(x) l_n g^{-1} \\ &\leq C \hat{r}_1(x) C n^{-\zeta} g^{-1} \end{aligned} \quad (4.89)$$

Pour  $n$  suffisamment grand on obtient :

$$\lim_{n \rightarrow \infty} D_1 = 0 \quad (4.90)$$

En reprenant les mêmes arguments on arrive à :

$$|E \hat{r}_3(x, y) - E \hat{r}_3(x, t_y)| \leq C l_n g^{-1}$$

En utilisant successivement le lemme 21-ii et la proposition 6-i on obtient :

$$\lim_{n \rightarrow \infty} D_3 = 0 \text{ p.co.} \quad (4.91)$$

Pour le terme  $D_2$  on peut écrire pour tout  $\epsilon > 0$  :

$$\begin{aligned} P\left(\sup_{y \in S} |\hat{r}_3(x, t_y) - E\hat{r}_4(x, t_y)| > \epsilon\right) &= P\left(\max_{j=1, \dots, z_n} |\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)| > \epsilon\right) \\ &\leq z_n \max_{j=1, \dots, z_n} P(|\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)| > \epsilon). \end{aligned}$$

En utilisant (4.58) on obtient pour tout  $y \in S$  :

$$P\left(|\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)| > \epsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}\right) = O(z_n n^{-C\epsilon_0}),$$

En utilisant (4.87) avec  $\epsilon_0$  assez grand on arrive à :

$$\sum_{n \geq 1} P\left(\sup_{y \in S} |\hat{r}_3(x, t_y) - E\hat{r}_4(x, t_y)| > \epsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}\right) < \infty.$$

Le dénominateur est traité en appliquant le lemme 21-ii et la proposition 6-i, par conséquent

$$D_2 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co. \quad (4.92)$$

Enfin, en utilisant (4.88) (4.90) (4.91) et (4.92) on arrive à :

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - E\hat{r}_3(x, y)| = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co,$$

ce qui prouve (4.86). ■

La proposition suivante indique la convergence presque complète ponctuelle et uniforme sur quelques ensembles compacts de l'estimateur à noyau de la densité conditionnelle.

**Proposition 30** *i) Sous les conditions (3.21), (4.32), (4.33) et (4.44), on a pour tout réel fixé  $y$  :*

$$\lim_{n \rightarrow \infty} f_Y^x(y) = \hat{f}_Y^x(y), p.co. \quad (4.93)$$

*ii) Si de plus la largeur de la fenêtre  $g$  vérifie pour tout  $\zeta > 0$  la condition*

*$\lim_{n \rightarrow \infty} gn^\zeta = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |f_Y^x(y) - \hat{f}_Y^x(y)| = 0, p.co.$$

**Preuve.** Il suffit de noter  $f_Y^x = F_Y^{x(1)}$  et de voir que le résultat (4.93) n'est qu'un cas particulier du lemme 23 qui sera étudié plus tard (en appliquant le (4.110) pour  $l=1$ ), d'autre part, le résultat (4.94) est exactement celui qui est donné par le lemme 25. ■

## 4.2 Les vitesses de convergence

### 4.2.1 L'estimation de la régression

Pour compléter la propriété de consistance de l'estimateur à noyau de la régression fonctionnelle  $\hat{r}(x)$  (défini en (3.23)) donné par le théorème 19, cette section sert à étudier la vitesse de convergence presque complète ponctuelle. En effet on considère le modèle de type Lipschitz pour  $r$ , et on va voir que ce modèle nous permet de préciser le comportement du biais et ensuite de tirer la vitesse de convergence.

**Théorème 31** *Sous le modèle (3.12) avec la condition de probabilité (4.32), si l'estimateur vérifie (4.33) et si la variable réponse  $Y$  vérifie (4.34), alors on a :*

$$\hat{r}(x) - r(x) = O(h^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co. \quad (4.95)$$

**Preuve.** *La démonstration de ce théorème est similaire à celle du théorème 2. En effet : en faisant la décomposition suivante :*

$$\begin{aligned} \hat{r}(x) - r(x) &= \frac{1}{\hat{r}_1(x)} [(\hat{r}_2(x) - E\hat{r}_2(x)) - (r(x) - E\hat{r}_2(x))] \\ &\quad - \frac{r(x)}{\hat{r}_1(x)} [\hat{r}_1(x) - 1] \end{aligned}$$

■

Selon le lemme 21 on a :

$$\hat{r}_2(x) - E\hat{r}_2(x) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

$$\hat{r}_1(x) - 1 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

selon le lemme 30 ci-dessous on a :

$$r(x) - E\hat{r}_2(x) = O(h^\beta)$$

d'où en utilisant la proposition 6.

$$\hat{r}(x) - r(x) = O(h^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co$$

**Lemme 32** *Sous les conditions (3.12) et (4.33) on a :*

$$r(x) - E\hat{r}_2(x) = O(h^\beta).$$

**Preuve.** Selon le début de la preuve du lemme 20 on a :

$$r(x) - E\hat{r}_2(x) = E((r(x) - r(X_1)) \Delta_1)$$

puisque  $r$  est Lipschitzien, on a :

$$|r(x) - E\hat{r}_2(x)| \leq CE(d(x, X_1)^\beta \Delta_1)$$

maintenant, puisque le support du noyau  $K$  est  $[0, 1]$  et  $E\Delta_1 = 1$ , on a :

$$|r(x) - E\hat{r}_2(x)| \leq Ch^\beta.$$

■

### 4.2.2 L'estimation de la médiane conditionnelle

Le principe de cette section est de montrer, en ajoutant l'hypothèse de lissage sur la fonction  $\hat{F}_Y^x$ , comment ceci nous permet de préciser la vitesse de convergence. On va voir dans le lemme 32 ci-dessous que le modèle de type-Lipschitz est suffisant de préciser la vitesse de convergence de l'estimateur  $\hat{F}_Y^x$ . Supposons que  $F_Y^x(\cdot)$  est  $j$ -fois continûment différentiable autour de  $m(x)$  avec :

$$F_Y^{x^{(l)}}(m(x)) = 0, \forall l = 1, \dots, j-1$$

et

$$F_Y^{x^{(l)}}(m(x)) > 0 \text{ pour } l = j \quad (4.96)$$

Notons que tout au long de cette section, toutes les dérivées sont prises pour la variable réelle  $y$ . Puisque cette approximation a besoin d'avoir des dérivées de grand ordre de  $\hat{F}_Y^x(\cdot)$  (voir le lemme (15)), l'hypothèse suivante sur les composantes de l'estimateur est nécessaire :

$$\lim_{n \rightarrow \infty} \frac{\log n}{ng^{2j-1}\varphi_x(h)} = 0 \text{ et } H \text{ est } j\text{-fois continûment différentiable} \quad (4.97)$$

**Théorème 33** *Sous le modèle défini par (3.18) et (4.96), si la variable fonctionnelle  $X$  vérifie (4.32), et si l'estimateur à noyau vérifie (4.33), (4.44) et (4.97), alors on a :*

$$\hat{m}(x) - m(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right), p.co. \quad (4.98)$$

$$\hat{m}(x) - m(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right), p.co. \quad (4.99)$$



**Preuve.** Le développement de Taylor de la fonction  $\hat{F}_Y^x$  nous conduit à l'existence de  $m^*$  entre  $\hat{m}(x)$  et  $m(x)$  tel que :

$$\begin{aligned} \hat{F}_Y^x(m(x)) - \hat{F}_Y^x(\hat{m}(x)) &= \sum_{l=1}^{j-1} \frac{(m(x) - \hat{m}(x))^l}{l!} \hat{F}_Y^{x(l)}(m(x)) \\ &\quad + \frac{(m(x) - \hat{m}(x))^j}{j!} \hat{F}_Y^{x(j)}(m^*). \end{aligned}$$

Selon (4.96), cette égalité s'écrit :

$$\begin{aligned} \hat{F}_Y^x(m(x)) - \hat{F}_Y^x(\hat{m}(x)) &= \sum_{l=1}^{j-1} \frac{(m(x) - \hat{m}(x))^l}{l!} \left( \hat{F}_Y^{x(l)}(m(x)) - F_Y^{x(l)}(m(x)) \right) \\ &\quad + \frac{(m(x) - \hat{m}(x))^j}{j!} \hat{F}_Y^{x(j)}(m^*). \end{aligned}$$

Puisque  $\hat{F}_Y^x(\hat{m}(x)) = F_Y^x(m(x)) = \frac{1}{2}$ , on a :

$$\begin{aligned} (m(x) - \hat{m}(x))^j \hat{F}_Y^{x(j)}(m^*) &= \\ &O\left(\hat{F}_Y^x(m(x)) - F_Y^x(m(x))\right) + \\ &O\left(\sum_{l=1}^{j-1} (m(x) - \hat{m}(x))^l \left(\hat{F}_Y^{x(l)}(m(x)) - F_Y^{x(l)}(m(x))\right)\right) \end{aligned}$$

La combinaison du résultat du lemme 33 et du théorème 22, avec  $m^*$  compris entre  $\hat{m}(x)$  et  $m(x)$ , nous amène à :

$$\lim_{n \rightarrow \infty} \hat{F}_Y^{x(j)}(m^*) = F_Y^{x(j)}(m(x)), p.co.$$

Puisque la deuxième partie de l'hypothèse (4.96) assure que cette limite est différente de 0, alors en utilisant la proposition 6-ii on arrive à :

$$\begin{aligned} (m(x) - \hat{m}(x))^j &= O\left(\hat{F}_Y^x(m(x)) - F_Y^x(m(x))\right), p.co + \\ &O\left(\sum_{l=1}^{j-1} (m(x) - \hat{m}(x))^l \left(\hat{F}_Y^{x(l)}(m(x)) - F_Y^{x(l)}(m(x))\right)\right), p.co. \end{aligned}$$

Selon (4.96), et pour tout  $l \in \{0, \dots, j\}$  et pour tout  $y$  dans le voisinage de  $m(x)$ , il existe  $m^*$  compris entre  $y$  et  $m(x)$  tel que :

$$\hat{F}_Y^{x(l)}(y) - F_Y^{x(l)}(m(x)) = \frac{(y - m(x))^{j-l}}{(j-l)!} \hat{F}_Y^{x(j)}(m^*).$$

Ceci implique que  $F_Y^{x(l)}$  est Lipschitzienne autour de  $m(x)$  à l'ordre  $j-l$ . Donc, en utilisant le lemme 32 et le lemme 33 avec l'ordre de Lipschitz adapté, on obtient :

$$\begin{aligned} (m(x) - \hat{m}(x))^j &= O(h^\beta + g^\beta) p.co + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) \\ &+ O\left(\sum_{l=1}^{j-1} A_{n,l}\right) p.co + O\left(\sum_{l=1}^{j-1} B_{n,l}\right) p.co, \end{aligned} \quad (4.99)$$

où

$$A_{n,l} = (m(x) - \hat{m}(x))^l \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}$$

et

$$B_{n,l} = (m(x) - \hat{m}(x))^l g^{j-l}.$$

En faisant la comparaison de  $A_{n,l}$  ou  $B_{n,l}$  avec  $(m(x) - \hat{m}(x))^j$ . Si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que  $(m(x) - \hat{m}(x))^j = O(A_{n,l})$ , on peut écrire :

$$\begin{aligned} |m(x) - \hat{m}(x)|^j &\leq C |m(x) - \hat{m}(x)|^l \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}, \\ \implies |m(x) - \hat{m}(x)|^{j-l} &\leq C \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}, \\ \implies |m(x) - \hat{m}(x)|^j &\leq C \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{j}{2(j-l)}}. \end{aligned}$$

donc selon (4.97). il existe  $l$  pour lequel :  $|m(x) - \hat{m}(x)|^j = O(A_{n,l})$ , alors on a :

$$|m(x) - \hat{m}(x)|^j = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) \quad (4.100)$$

De la même manière, si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que  $(m(x) - \hat{m}(x))^j = O(B_{n,l})$ , on peut écrire :

$$\begin{aligned} |m(x) - \hat{m}(x)|^j &\leq C |m(x) - \hat{m}(x)|^l g^{j-l} \\ \implies |m(x) - \hat{m}(x)|^j &\leq C g^j. \end{aligned}$$

Donc il existe  $l$  tel que  $(m(x) - \hat{m}(x))^j = O(B_{n,l})$ , alors on a :

$$(m(x) - \hat{m}(x))^j = O(g^\beta) \quad (4.101)$$

Finalement la combinaison de (4.99), (4.100) et (4.101) nous donne :

$$(m(x) - \hat{m}(x))^j = O(h^\beta + g^\beta) p.co + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co \quad (4.102)$$

■

Le théorème est prouvé, il nous reste à montrer les lemmes suivants :

**Lemme 34** *Sous les conditions du théorème 31, on a :*

$$F_Y^x - \hat{F}_Y^x = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co \quad (4.103)$$

**Preuve.** La démonstration de ce lemme est identique à celle du lemme 23, en gardant les mêmes notations utilisées. La preuve est basée sur la décomposition (4.48). Le dénominateur dans cette décomposition est traité en appliquant le lemme 21-ii avec la proposition 6-ii. Donc le résultat demandé est obtenu dès que les trois résultats suivants sont vérifiés :

$$\hat{r}_1(x) - 1 = O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right) p.co, \quad (4.104)$$

$$E\hat{r}_3(x, y) - F_Y^x(y) = O(g^\beta) + O(h^\beta), \quad (4.105)$$

et

$$\hat{r}_3(x, y) - E\hat{r}_3(x, y) = O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right) p.co. \quad (4.106)$$

En effet, en prenant  $\varepsilon = \varepsilon_0 \left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}$  dans (4.58) on obtient :

$$\sum_{n \geq 1} P\left(\hat{r}_3(x, y) - E\hat{r}_3(x, y) > \varepsilon_0 \left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right) \leq n^{-C\varepsilon_0^2}, \quad (4.107)$$

par suite pour  $\varepsilon_0$  suffisamment grand :

$$\sum_{n \geq 1} P\left(\hat{r}_3(x, y) - E\hat{r}_3(x, y) > \varepsilon_0 \left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right) < \infty. \quad (4.108)$$

Ceci prouve que (4.106) est vérifié. Il nous reste à montrer (4.105). La condition de Lipschitz (3.18) nous permet d'écrire :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, 1]} 1_{B(x, h)}(X) |F_Y^X(y - vg) - F_Y^x(y)| = O(g^\beta) + O(h^\beta),$$

la combinaison de cette dernière égalité avec (4.53) nous amène à :

$$\lim_{n \rightarrow \infty} 1_{B(x, h)}(X) |E(\Gamma_1(y) | X) - F_Y^x(y)| = O(g^\beta) + O(h^\beta). \quad (4.109)$$

D'où (4.109) avec (4.51) et  $E\Delta_1 = 1$  suffit pour la preuve de (4.105). La preuve de ce lemme est ainsi achevée, puisque le résultat (4.104) est déjà étudié lors de l'étude de la régression (voir le lemme 21-ii). ■

**Lemme 35** Soit  $l$  un entier appartenant à l'ensemble  $\{1, \dots, j\}$ . Sous les conditions du théorème 31, on a :

$$\lim_{n \rightarrow \infty} \hat{F}_Y^{x(l)}(y) = F_Y^{x(l)}(y), p.co. \quad (4.110)$$

Si de plus la fonction  $F_Y^{x^{(l)}}(\cdot)$  est continûment Lipschitzienne d'ordre  $\beta_0$ , c.à.d si

$$\exists C \in ]0, \infty[, \forall (y, y') \in \mathbb{R}^2, \left| F_Y^{x^{(l)}}(y) - F_Y^{x^{(l)}}(y') \right| \leq C |y - y'|^{\beta_0} \quad (4.111)$$

alors on a :

$$F_Y^{x^{(l)}}(y) - \hat{F}_Y^{x^{(l)}}(y) = O\left(h^\beta + h^{\beta_0}\right) + O\left(\left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}\right). \quad (4.112)$$

**Preuve.** On considère la décomposition suivante :

$$\begin{aligned} \hat{F}_Y^{x^{(l)}}(y) - F_Y^{x^{(l)}}(y) &= \frac{1}{\hat{r}_1(x)} \left( \hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) \right) \\ &\quad - \frac{1}{\hat{r}_1(x)} \left( F_Y^{x^{(l)}}(y) - E\hat{r}_3^{(l)}(x, y) \right) \\ &\quad + \frac{F_Y^{x^{(l)}}(y)}{\hat{r}_1(x)} (1 - \hat{r}_1(x)) \end{aligned} \quad (4.113)$$

Comme pour le lemme 23, pour la démonstration de ce lemme il suffit de prouver les résultats suivantes :

$$\lim_{n \rightarrow \infty} \left| \hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) \right| = 0, p.co. \quad (4.114)$$

et aussi

$$\lim_{n \rightarrow \infty} E\hat{r}_3^{(l)}(x, y) = F_Y^{x^{(l)}}(y), \quad (4.115)$$

où  $F_Y^{x^{(l)}}$  est continue, et

$$E\hat{r}_3^{(l)}(x, y) - F_Y^{x^{(l)}}(y) = O\left(g^{\beta_0}\right) + O\left(h^\beta\right), \quad (4.116)$$

où  $F_Y^{x^{(l)}}$  est continûment Lipschitzienne. La preuve de (4.114) est similaire à celle de (4.50),

On peut écrire :

$$\hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) = \frac{1}{n} \sum_{i=1}^n (S_i - ES_i), \quad (4.117)$$

où  $S_i$  sont des variables aléatoires centrées et vérifiant :

$$|S_i| \leq C / \left( g^l \varphi_x(h) \right). \quad (4.118)$$

Le moment d'ordre deux des variables  $S_i$  peut être calculé par :

$$\begin{aligned} ES_i^2 &= \frac{1}{g^{2l}} E \left( \Delta_i^2 E \Gamma_i^{(l)2} / X = x \right) \\ &= \frac{1}{g^{2l}} E \left( \Delta_i^2 \int_{\mathbb{R}} H^{(l)2} \left( \frac{y-u}{g} \right) f_Y^X(u) du \right) \\ &\leq \frac{1}{g^{2l-1}} E \left( \Delta_i^2 \int_{\mathbb{R}} H^{(l)2}(v) f_Y^X(y-gv) dv \right) \\ &\leq C \frac{1}{g^{2l-1}} E(\Delta_i^2) \\ &\leq C \frac{1}{g^{2l-1} \varphi_x(h)} \end{aligned} \quad (4.119)$$

Cette dernière inégalité est obtenu grâce à (4.57). Notons que l'existence de  $f_Y^x$  est assurée par l'hypothèse de différentiabilité fait sur  $F_Y^x$ . Appliquons maintenant l'égalité exponentielle de type Bernstein donnée par le corollaire 9-ii, ceci nous amène à :

$$\hat{r}_3^{(l)}(x, y) - E \hat{r}_3^{(l)}(x, y) = O \left( \frac{\log n}{n g^{2l-1} \varphi_x(h)} \right)^{\frac{1}{2}} p.co. \quad (4.120)$$

D'où le résultat (4.114) est satisfait. La preuve de (4.115) et (4.116) sont très proches de celles de (4.49) et (4.105). En faisant l'intégration par partie  $(l-1)$  fois et en faisant le changement de variable, on arrive à :

$$\begin{aligned} E \left( \Gamma_1^{(l)}(y) / X = x \right) &= \frac{1}{g^l} \int_{\mathbb{R}} H^{(l)} \left( \frac{y-u}{g} \right) f_Y^X(u) du \\ &= \frac{1}{g} \int_{\mathbb{R}} H^{(l)} \left( \frac{y-u}{g} \right) F_Y^{X^{(l)}}(u) du \\ &= \int_{\mathbb{R}} K_0(v) F_Y^{X^{(l)}}(y-gv) dv. \end{aligned}$$

Puisque les noyaux sont à support compact alors on peut écrire :

$$F_Y^X(y) = F_Y^x(y) + O(h^\beta)$$

avec  $O(h^\beta)$  est indépendant de  $y$ . Donc on a :

$$\begin{aligned}
& E\hat{r}_3^{(l)}(x, y) - F_Y^{x(l)}(y) \\
&= E\left(\Delta_1\left(\int_{\mathbb{R}} K_0(v)\left(F_Y^{X(l)}(y - gv) - F_Y^{x(l)}(y)\right)dv\right)\right) \\
&= E\left(\Delta_1\left(\int_{\mathbb{R}} K_0(v)\left(F_Y^{x(l)}(y - gv) - F_Y^{x(l)}(y)\right)dv\right)\right) \\
&\quad + O(h^\beta)
\end{aligned} \tag{4.121}$$

Comme  $K$  et  $K_0$  sont de support compact et comme  $g$  et  $h$  tendent vers 0 le modèle de type continu nous permet d'écrire :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, 1]} \left| F_Y^{x(l)}(y - gv) - F_Y^{x(l)}(y) \right| = 0 \tag{4.122}$$

la combinaison de (4.121) , (4.122) et  $E(\Delta_1) = 1$  donne directement le résultat (4.105).

De la même manière, le résultat(4.116) se déduit du fait que le modèle de  $F_Y^{x(l)}$  est de type Lipschitz et qui nous permet d'écrire :

$$\sup_{v \in [-1, 1]} \left| F_Y^{x(l)}(y - gv) - F_Y^{x(l)}(y) \right| = O(g^{\beta_0}) + O(h^\beta). \tag{4.123}$$

Ainsi la preuve du lemme (15) est achevée. ■

### 4.2.3 L'estimation du mode conditionnel

Le principe de cette section est de montrer comment , en ajoutant l'hypothèse de lissage sur la fonction  $\hat{f}_Y^x$  , nous permet de préciser la vitesse de convergence. Dans l'étude de la régression (voir la section ci-dessus) on a vu dans le lemme 32 ci-dessus que le modèle de type-Lipschitz est suffisant de préciser la vitesse de convergence de l'estimateur  $\hat{f}_Y^x$ . Supposons que  $\hat{f}_Y^x(\cdot)$  est  $j$ -fois continûment différentiable autour de  $\theta(x)$  avec :

$$f_Y^{x(l)}(\theta(x)) = 0, \forall l = 1, \dots, j - 1$$

et

$$f_Y^{x^{(l)}}(\theta(x)) \neq 0, \text{ si } l = j \quad (4.124)$$

Notons que, tout au long de cette section, toutes les dérivées sont prises pour la variable réelle  $y$ .

**Théorème 36** *Sous le modèle défini par (3.22) et (4.124), si la variable fonctionnelle  $X$  vérifie (4.32), et si l'estimateur à noyau vérifie (4.33), (4.44), (4.61) et (4.97), alors on a :*

$$\hat{\theta}(x) - \theta(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right), p.co. \quad (4.125)$$

**Preuve.** Le développement de Taylor d'ordre  $j$  de la fonction  $f_Y^x$  au point  $\theta(x)$ , en tenant compte de la condition (4.124), nous permet d'affirmer l'assertion suivante : il existe  $\theta^*$  compris entre  $\theta(x)$  et  $\hat{\theta}(x)$  tel que :

$$f_Y^x(\hat{\theta}(x)) = f_Y^x(\theta(x)) + \frac{1}{j!} f_Y^{x(j)}(\theta^*) (\theta(x) - \hat{\theta}(x))^j,$$

la combinaison de ce résultat avec (4.63) nous permet d'écrire :

$$f_Y^{x(j)}(\theta^*) (\theta(x) - \hat{\theta}(x))^j = O\left(\sup_{s \in (\theta(x) - \xi, \theta(x) + \xi)} \left| \hat{f}_Y^x(s) - f_Y^x(s) \right|\right) \quad (4.126)$$

D'autre part comme  $f_Y^{x(j)}$  est continue alors :

$$\forall \epsilon > 0, \exists \delta_\epsilon > 0, |\theta(x) - \theta^*| \leq \delta_\epsilon \implies \left| f_Y^{x(j)}(\theta(x)) - f_Y^{x(j)}(\theta^*) \right| \leq \epsilon$$

d'où :

$$\forall \epsilon > 0, \exists \delta_\epsilon > 0, P\left(\left| f_Y^{x(j)}(\theta(x)) - f_Y^{x(j)}(\theta^*) \right| > \epsilon\right) \leq P(|\theta(x) - \theta^*| > \delta_\epsilon).$$



Puisque le résultat du théorème 24 assure que  $\theta^*$  tend vers  $\theta(x)$  presque complète on obtient :

$$\lim_{n \rightarrow \infty} f_Y^{x(j)}(\theta^*) = f_Y^{x(j)}(\theta(x)) \neq 0 \quad (4.127)$$

En utilisant (4.126), (4.127) avec la proposition 6-ii, on arrive à :

$$|\theta(x) - \theta^*| = O \left( \sup_{s \in (\theta(x) - \xi, \theta(x) + \xi)} \left| \hat{f}_Y^x(s) - f_Y^x(s) \right| \right) p.co, \quad (4.128)$$

en appliquant le lemme 35 ci dessous avec  $l = 1$  et  $S = (\theta(x) - \xi, \theta(x) + \xi)$  on obtient :

(4.125). ■

**Lemme 37** *Sous les conditions du théorème 34, on a pour tout compact  $S \subset \mathbb{R}$  et pour tout entier  $l = 1, \dots, j$  :*

$$\begin{aligned} \sup_{y \in S} \left| F_Y^{x(l)}(y) - \hat{F}_Y^{x(l)}(y) \right| &= O \left( h^\beta + g^{\beta_0} \right) \\ &+ O \left( \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) p.co \end{aligned} \quad (4.129)$$

**Preuve.** En gardant les mêmes notations utilisées pour la preuve du lemme 33. En faisant la décomposition (4.113), en reprenant les mêmes arguments de la preuve du lemme 33, donc la démonstration de ce lemme revient à démontrer les deux résultats suivants :

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) \right| = O \left( \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) p.co \quad (4.130)$$

et

$$\sup_{y \in S} \left| E\hat{r}_3^{(l)}(x, y) - F_Y^{x(l)}(y) \right| = O \left( h^\beta + g^{\beta_0} \right) \quad (4.131)$$

Pour la preuve de (4.131), la limite qui apparait dans (4.123) ne dépend pas de  $y$  (puisque  $S$  est un compact), donc on peut écrire :

$$\sup_{v \in [-1, 1]} \sup_{y \in S} \left| F_Y^{x(l)}(y - gv) - F_Y^{x(l)}(y) \right| = O \left( g^{\beta_0} \right) + O \left( h^\beta \right), \quad (4.132)$$

maintenant le résultat(4.131) se déduit de la combinaison de (4.121), (4.132) et  $E\Delta_1 = 1$ .

Il reste à vérifier (4.130). En effet puisque  $S$  est un compact alors on peut écrire que  $S \subset \cup_{k=1}^{z_n} S_k$  où  $S_k = (t_k - l_n, t_k + l_n)$  et  $l_n$  et  $z_n$  vérifient :

$$l_n = C / z_n \sim C n^{-(l+1)\zeta - \frac{1}{2}} \quad (4.133)$$

En prenant  $t_y = \arg \min_{t \in \{t_1, \dots, t_{z_n}\}} |y - t|$ , on a :

$$\begin{aligned} & \frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - E \hat{r}_3^{(l)}(x, y) \right| \\ & \leq \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - \hat{r}_3^{(l)}(x, t_y) \right|}_{B_1} \\ & \quad + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right|}_{B_2} \\ & \quad + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| E \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, y) \right|}_{B_3} \end{aligned} \quad (4.134)$$

La condition de continuité de Hölder sur le noyau  $H$  nous permet d'écrire :

$$\begin{aligned} & \left| \hat{r}_3^{(l)}(x, y) - \hat{r}_3^{(l)}(x, t_y) \right| \\ & = \frac{1}{ng^l} \sum_{i=1}^n \Delta_i \left| H^{(l)}(g^{-1}(y - Y_i)) - H^{(l)}(g^{-1}(t_y - Y_i)) \right| \\ & \leq \frac{C}{ng^l} \sum_{i=1}^n \Delta_i \frac{|y - t_y|}{g} \\ & \leq C \hat{r}_1(x) l_n g^{-l-1} \end{aligned} \quad (4.135)$$

La condition (4.133) avec la deuxième partie de (4.61) nous amène à :

$$\begin{aligned} B_1 & \leq C n^{-(l+1)\zeta - \frac{1}{2}} g^{-l-1} \\ B_1 & = O \left( \sqrt{\frac{\log n}{ng^{2l-1} \varphi_x(h)}} \right) \end{aligned} \quad (4.136)$$

Pour le terme  $B_3$  on a :

$$\begin{aligned} & \left| E \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, y) \right| \\ &= O \left( \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) \end{aligned} \quad (4.137)$$

Pour le terme  $B_2$ , on peut écrire que pour tout  $\epsilon > 0$  :

$$\begin{aligned} & P \left( \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right| > \epsilon \right) \\ &= P \left( \max_{j=1, \dots, z_n} \left| \hat{r}_3^{(l)}(x, t_j) - E \hat{r}_3^{(l)}(x, t_j) \right| > \epsilon \right) \\ &\leq z_n \max_{j=1, \dots, z_n} P \left( \left| \hat{r}_3^{(l)}(x, t_j) - E \hat{r}_3^{(l)}(x, t_j) \right| > \epsilon \right). \end{aligned}$$

En utilisant l'inégalité de Bernstein pour les variables bornées avec les conditions (4.118)

et (4.119) on obtient pour tout  $t_j$  :

$$\begin{aligned} P \left( \left| \hat{r}_3^{(l)}(x, t_j) - E \hat{r}_3^{(l)}(x, t_j) \right| > \epsilon_0 \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) &\leq C \exp \left( -C \frac{\epsilon_0^2 \log n}{(1+g^{-1})} \right) \\ &= O \left( n^{-C\epsilon_0^2} \right) \end{aligned}$$

pour  $\epsilon_0$  suffisamment grand et en faisant (4.133) on obtient :

$$\sum_{n \geq 1} P \left( \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right| > \epsilon_0 \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) < \infty,$$

puisque le dénominateur est traité en utilisant le lemme 21-ii et la proposition 6-ii alors

$$B_2 = O \left( \sqrt{\frac{\log n}{ng^{2l-1}\varphi_x(h)}} \right) \quad (4.138)$$

■

Finalement (4.134), (4.136), (4.137) et (4.138) prouvent (4.130), et la preuve du lemme est achevée.

#### 4.2.4 L'estimateur du quantile conditionnel

Cette section est consacrée à la généralisation des résultats donnés dans la section concernant l'estimation de la médiane à l'estimation des quantiles conditionnels. En d'autres termes cette section précise la vitesse de convergence apparait dans les résultats de la section 3.2.4. Le quantile d'ordre  $\alpha$ , noté par  $t_\alpha(x)$ , et son estimateur à noyau  $\hat{t}_\alpha(x)$  sont respectivement définis par (3.9) et (3.28). Sous la condition (4.44), l'estimateur de quantile conditionnel peut être défini par l'unique solution de l'équation

$$\hat{t}_\alpha(x) = \hat{F}_Y^{x-1}(\alpha) \quad (4.139)$$

comme pour l'estimateur de la médiane, on a besoin de l'hypothèse suivante, qui sert à contrôler la monotonie de la fonction de répartition conditionnelle autour du quantile qu'on veut estimer.

$$F_Y^{x(l)}(t_\alpha(x)) = 0, \forall l = 1, \dots, j-1$$

et

$$F_Y^{x(l)}(t_\alpha(x)) > 0 \text{ pour } l = j. \quad (4.140)$$

**Théorème 38** Soit  $\alpha \in (0, 1)$ . Sous le modèle défini par (3.17) et (4.140), Si la variable fonctionnelle  $X$  vérifie (4.32), et si l'estimateur à noyau vérifie (4.33), (4.44), (4.97) et (4.140), alors on a :

$$\hat{t}_\alpha(x) - t_\alpha(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right), p.co \quad (4.141)$$

**Preuve.** La preuve de ce théorème est similaire à celle du théorème 31. C'est pour cette raison qu'on va la présenter brièvement. Le développement de Taylor de la fonction

$\hat{F}_Y^x$  nous conduit à l'existence de  $t^*$  compris entre  $\hat{t}_\alpha(x)$  et  $t_\alpha(x)$  tel que :

$$\begin{aligned} \hat{F}_Y^x(t_\alpha(x)) - \hat{F}_Y^x(\hat{t}_\alpha(x)) &= \sum_{l=1}^{j-1} \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^l}{l!} \hat{F}_Y^{x(l)}(t_\alpha(x)) \\ &\quad + \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^j}{j!} \hat{F}_Y^{x(j)}(t^*). \end{aligned}$$

Selon (4.140), cette égalité s'écrit :

$$\begin{aligned} \hat{F}_Y^x(t_\alpha(x)) - \hat{F}_Y^x(\hat{t}_\alpha(x)) &= \sum_{l=1}^{j-1} \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^l}{l!} \left( \hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x)) \right) \\ &\quad + \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^j}{j!} \hat{F}_Y^{x(j)}(t^*). \end{aligned}$$

Puisque  $\hat{F}_Y^x(\hat{t}_\alpha(x)) = F_Y^x(t_\alpha(x)) = \alpha$ , on a :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j \hat{F}_Y^{x(j)}(t^*) &= O\left(\hat{F}_Y^x(t_\alpha(x)) - F_Y^x(t_\alpha(x))\right) + \\ &\quad O\left(\sum_{l=1}^{j-1} (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x))\right)\right) \end{aligned}$$

La combinaison du résultat du lemme 33 et du théorème 22, avec  $t^*$  compris entre  $\hat{t}_\alpha(x)$  et  $t_\alpha(x)$ , nous amène à :

$$\lim_{n \rightarrow \infty} \hat{F}_Y^{x(j)}(t^*) = F_Y^{x(j)}(t_\alpha(x)), p.co.$$

Puisque la deuxième partie de l'hypothèse (4.140) assure que cette limite est différente de 0, alors en utilisant la proposition 6-ii on arrive à :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j &= O\left(\hat{F}_Y^x(t_\alpha(x)) - F_Y^x(t_\alpha(x))\right), p.co + \\ &\quad O\left(\sum_{l=1}^{j-1} (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x))\right)\right), p.co. \end{aligned}$$

Selon (4.140) et pour tout  $l \in \{0, \dots, j\}$  et pour tout  $y$  dans le voisinage de  $t_\alpha(x)$ , il existe  $t^*$  compris entre  $y$  et  $t_\alpha(x)$  tel que :

$$\hat{F}_Y^{x(l)}(y) - F_Y^{x(l)}(t_\alpha(x)) = \frac{(y - t_\alpha(x))^{j-l}}{(j-l)!} \hat{F}_Y^{x(j)}(t^*).$$

Ceci implique que  $F_Y^{x(l)}$  est Lipschitzienne autour de  $t_\alpha(x)$  à l'ordre  $j-l$ . Donc, en utilisant le lemme 32 et le lemme 33 avec l'ordre de Lipschitz adapté, on obtient :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j &= O(h^\beta + g^\beta) p.co + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) \\ &+ O\left(\sum_{l=1}^{j-1} A_{n,l}^*\right) p.co + O\left(\sum_{l=1}^{j-1} B_{n,l}^*\right) p.co, \end{aligned} \quad (4.142)$$

où

$$A_{n,l}^* = (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}$$

et

$$B_{n,l} = (t_\alpha(x) - \hat{t}_\alpha(x))^l g^{j-l}.$$

En faisant la comparaison de  $A_{n,l}^*$  ou  $B_{n,l}^*$  avec  $(t_\alpha(x) - \hat{t}_\alpha(x))^j$ . Si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(A_{n,l})$ , on peut écrire :

$$\begin{aligned} |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C |t_\alpha(x) - \hat{t}_\alpha(x)|^l \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}, \\ \implies |t_\alpha(x) - \hat{t}_\alpha(x)|^{j-l} &\leq C \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}, \\ \implies |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C \left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{j}{2(j-l)}}. \end{aligned}$$

donc selon (4.97), il existe  $l$  pour lequel :  $|t_\alpha(x) - \hat{t}_\alpha(x)|^j = O(A_{n,l})$ , alors on a:

$$|t_\alpha(x) - \hat{t}_\alpha(x)|^j = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co, \quad (4.143)$$

De la même manière, si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que :  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(B_{n,l})$ , on peut écrire :

$$|t_\alpha(x) - \hat{t}_\alpha(x)|^j \leq C |t_\alpha(x) - \hat{t}_\alpha(x)|^l g^{j-l}$$

$$\implies |t_\alpha(x) - \hat{t}_\alpha(x)|^j \leq Cg^j.$$

Donc il existe  $l$  tel que :  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(B_{n,l})$ , alors on a :

$$(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(g^\beta) \quad (4.144)$$

Finalement la combinaison de (4.142), (4.143) et (4.144) nous donne :

$$(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right), p.co.$$

■

#### 4.2.5 Complément sur l'estimation de la distribution conditionnelle

Cette section est consacrée à la présentation générale de plusieurs résultats concernant la vitesse de convergence de l'estimateur de la fonction de répartition conditionnelle et de la fonction de densité conditionnelle. La plupart de ces résultats sont déjà prouvés au paravant lors de l'étude du mode conditionnel et de l'estimateur du quantile.

**Proposition 39** *i) Sous les conditions (3.18), (4.96), (4.32), (4.33) et (4.44), On a pour tout réel fixé  $y$  :*

$$F_Y^x(y) - \hat{F}_Y^x(y) = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co$$

*ii) Si en plus la largeur de fenêtre  $g$  vérifie pour  $a > 0$  la condition  $\lim_{n \rightarrow \infty} gn^a = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\sup_{y \in S} |F_Y^x(y) - \hat{F}_Y^x(y)| = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

**Preuve.** i) est déjà étudié par le lemme 32. Il reste à prouver ii). En faisant la décomposition (4.48) et puisque  $\hat{r}_1(x)$  ne dépend pas de  $y \in S$  donc son traitement est exactement celle déjà vu dans la preuve du lemme 23. Alors il nous reste à prouver les deux résultats suivants :

$$\sup_{y \in S} |E\hat{r}_3(x) - F_Y^x(y)| = O(h^\beta + g^\beta) \quad (4.145)$$

et

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - E\hat{r}_3(x, y)| = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

En effet ce dernier résultats est déjà vu en (2.92). enfin ce qu'il nous reste à prouver est (4.145). Puisque  $K_0$  est de support  $[-1, 1]$  et  $F_Y^x$  est Lipschitzienne alors :

$$\sup_{y \in S} \sup_{v \in [-1, 1]} 1_{B(x, h)}(X) |F_Y^x(y - vg) - F_Y^x(y)| = O(g^\beta + h^\beta)$$

La combinaison de ce dernier résultat avec (4.51) et (4.53) sont suffisantes pour la preuve de (2.145). ■

**Proposition 40** *i) Sous les conditions (3.22), (4.32), (4.33) et (4.44), pour tout nombre réel fixé  $y$  on a :*

$$f_Y^x(y) - \hat{f}_Y^x(y) = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

*ii) Si en plus la largeur de fenêtre  $g$  vérifie pour  $a > 0$  la condition  $\lim_{n \rightarrow \infty} gn^a = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\sup_{y \in S} |f_Y^x(y) - \hat{f}_Y^x(y)| = O(h^\beta + g^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) p.co.$$

**Preuve.** cette proposition n'est qu'un cas particulier (pour  $l = 1$  et  $\beta = \beta_0$ ) des résultats déjà vus dans la preuve de la deuxième partie du lemme 33 et du lemme 35. ■



## Deuxième partie

### Cas du mélange fort

La condition du mélange est une structure usuelle pour les modèles non paramétriques en cas de dépendance de la suite des variables aléatoires. Avant d'aborder l'étude asymptotique on se souvient de quelques définitions et de quelques notations. On s'intéresse dans ce chapitre à la notion de  $\alpha$ -mixing qui est défini comme suit :

**Définition 41** Soit  $(\xi_i)_{i \in \mathbb{Z}}$  une suite de variables aléatoires. On note  $\mathcal{F}_k = \sigma(X_i, i \leq k)$  et  $\mathcal{L}_l = \sigma(X_i, l \leq i)$ . On définit les coefficients de mélange associés à la suite  $(\xi_i)_{i \in \mathbb{Z}}$  par :

$$\alpha(n) = \sup_{k \in \mathbb{Z}} \sup_{A \in \mathcal{F}_k, B \in \mathcal{L}_{n+k}} |P(A \cap B) - P(A)P(B)|$$

Une suite de variables  $(\xi_i)_{i \in \mathbb{Z}}$  est dite  $\alpha$ -mélangeante si ses coefficients de mélange vérifient  $\alpha(n) \rightarrow 0$ , quand  $n \rightarrow \infty$ .

Parmi les différents types de décroissance on distingue deux types : arithmétique et géométrique.

**Définition 42** Les coefficients de  $\alpha$ -mélange sont dits algébriques d'ordre  $a > 0$  s'il existe une constante positive  $C$  telle que :  $\alpha(n) \leq Cn^{-a}$ .

Les coefficients de  $\alpha$ -mélange sont dits géométriques s'il existe une constante positive  $C$  et  $0 \leq \rho < 1$  tels que :  $\alpha(n) \leq C\rho^n$ .

En s'appuyant sur la méthode du noyau on veut prédire le scalaire réponse  $Y$  du prédicteur fonctionnel  $X$ , en utilisant l'opérateur de régression non linéaire  $r$  défini par :

$$r(x) = E(Y/X = x) \tag{4.145}$$

Soit  $(X_i, Y_i)_{i=1, \dots, n}$   $n$  couples identiquement distribués du couple  $(X, Y)$  vérifiant une condition de mélange.

L'estimateur à noyau de  $r$  est :

$$\hat{r}(x) = \frac{\sum_{i=1}^n Y_i K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))} \quad (4.146)$$

où  $K$  est un noyau symétrique et  $h$  qui dépend de  $n$  est un réel strictement positif.

### 4.3 La convergence presque complète de la régression

Comme dans le cas d'indépendance, il ya une décomposition : biais/variance, dans ce cas de dépendance sauf les termes concernant le calcul de variance qui changent et la structure de mélange sera contrôler par les quantités suivantes :

$$s_{n,2} = \sum_{i=1}^n \sum_{j=1}^n cov(Y_i \Delta_i, Y_j \Delta_j), \quad (4.147)$$

$$s_{n,1} = \sum_{i=1}^n \sum_{j=1}^n cov(\Delta_i, \Delta_j), \quad (4.148)$$

où

$$\Delta_i = \frac{K(h^{-1}d(x, X_i))}{EK(h^{-1}d(x, X_i))}$$

Dans ce qui suit  $s_n$  désigne une suite quelconque des entiers positifs. Comme pour les résultats étudié en cas d'indépendance, on a besoin des hypothèses suivantes pour contrôler les covariances affectées :

$(X_i, Y_i)_{i=1, \dots, n}$  sont  $\alpha$ -mélangeantes avec des coefficients algébriques d'ordre  $a >$

1 et

$$\exists \theta > 2, s_n^{-(n+1)} = o(n^{-\theta}) \quad (4.149)$$

où  $(X_i, Y_i)_{i=1, \dots, n}$  sont  $\alpha$ -mélangeantes avec des coefficients géométriques, et

$$\exists \theta > 1, s_n^{-1} = o\left(n^{-\theta}\right) \quad (4.150)$$

**Théorème 43** Posons  $s_n = \max(s_{n,1}, s_{n,2})$  et supposons que (2.149) ou (2.150) est vérifiée,

alors :

i) Sous les conditions du théorème 19 on a :

$$\lim_{n \rightarrow \infty} \hat{r}(x) = r(x), p.co.$$

ii) Sous les conditions du théorème 29 on a :

$$\hat{r}(x) - r(x) = O\left(h^\beta\right) + O\left(\frac{\sqrt{s_n^2 \log n}}{n}\right) p.co.$$

**Preuve.** En faisant les mêmes notations introduites dans la décomposition (4.37)

et comme on peut écrire :

$$\hat{r}_2(x) = \frac{1}{n} \sum_{i=1}^n Y_i \Delta_i,$$

la condition (4.149) (resp. (4.150)) nous permet d'appliquer le corollaire 12.i (resp. le corollaire 13.i) et d'avoir directement

$$\hat{r}_2(x) - E\hat{r}_2(x) = O\left(\frac{\sqrt{s_n^2 \log n}}{n}\right), p.co. \quad (4.151)$$

Avec les mêmes arguments, puisque

$$\hat{r}_1(x) = \frac{1}{n} \sum_{i=1}^n \Delta_i \quad (4.152)$$

on a

$$E\hat{r}_1(x) = 1$$

et

$$\hat{r}_1(x) - 1 = O\left(\frac{\sqrt{s_n^2 \log n}}{n}\right), p.co. \quad (4.153)$$

Finalemnt la preuve du théorème 41-i) ( resp.41-ii)) vient dirèctement de (4.37) , (4.151) , (4.153) et du lemme 32 (resp.du lemme 3.12). ■

Les deux résultats donnés par le théorème 41 peuvent être donnés sous plusieurs formulations, selon la connaissance qu'on a sur le terme de covariance  $s_n^2$ . Ce terme dépend dirèctement de la distribution conjointe des deux couples différents  $(X_i, Y_i)$  et  $(X_j, Y_j)$  et du comportement des coefficients de mélange. On s'intéresse dans notre étude ci-dessous à une formulation simple des résultats, en faisant deux ensembles de condition (une de coefficient géométrique l'autre de coefficient algébrique)

### 4.3.1 Application du cas de mélange géométrique

On a besoin d'ajouter les hypothèses suivantes sur la distribution des deux couples distincts  $(X_i, Y_i)$  et  $(X_j, Y_j)$  et sur la largeur de la *fenêtre*  $h$ . On suppose que :

$$\forall i \neq j, E(Y_i Y_j / (X_i, X_j)) \leq C < \infty, \quad (4.154)$$

et la fonction de distribution conjointe

$$\psi_{i,j}(h) = P((X_i, X_j) \in B(x, h) \times B(x, h)),$$

vérifie

$$\exists \epsilon_1 \in ]0, 1], 0 < \psi_x(h) = O\left(\varphi_x(h)^{1+\epsilon_1}\right), \quad (4.155)$$

où

$$\psi_x(h) = \max_{i \neq j} \psi_{i,j}(h).$$

**Remarque 44** Dans le cas où l'espace  $\mathcal{F} = \mathbb{R}^k$ , la condition (4.155) est vérifiée pour  $\epsilon_1 = 1$  dès que chaque couple  $(X_i, X_j)$  a pour densité  $f_{i,j}$  tel que :  $\sup_{i,j} f_{i,j}(x, x) \leq C < \infty$ .

On suppose aussi que :

$$\exists \epsilon_2 \in ]0, 1], \varphi_x(h) = O(n^{-\epsilon_2}) \quad (4.156)$$

**Lemme 45** Supposons que les conditions du théorème 29 sont vérifiées. supposons aussi que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est géométriquement mélangeante vérifiant (4.154), (4.155) et (4.156). On a pour  $l = 1$  ou 2 :

$$s_{n,l} = O\left(\frac{n}{\varphi_x(h)}\right).$$

**Preuve.** Les conditions (4.154) et (4.155) nous permet d'écrire pour tout  $i \neq j$  :

$$\begin{aligned} E(Y_i \Delta_i Y_j \Delta_j) &= E(\Delta_i \Delta_j E(Y_i Y_j(X_i, Y_j))) \leq CE \Delta_i \Delta_j \\ &\leq \frac{C}{\varphi_x(h)^2} EK(h^{-1}d(x, X_i)) K(h^{-1}d(x, X_j)) \end{aligned}$$

la dernière inégalité vient de la définition de  $\Delta_i$  et de l'application du résultat (2.40) pour  $m = 1$ . Comme  $K$  est borné et à support compact, on a par définition de  $\psi_x(\cdot)$  :

$$\begin{aligned} E(Y_i \Delta_i Y_j \Delta_j) &\leq \frac{C}{\varphi_x(h)^2} P((X_i, Y_j) \in B(x, h) \times B(x, h)) \\ &\leq \frac{C \psi_x(h)}{\varphi_x(h)^2}. \end{aligned}$$

En utilisant maintenant le résultat (4.39) avec  $m = 1$  et en faisant la notation,

$$\Psi(h) = \max\left(\frac{\psi_x(h)}{\varphi_x(h)^2}, 1\right),$$

on obtient :

$$\begin{aligned} \text{cov}(Y_i \Delta_i, Y_j \Delta_j) &\leq E(Y_i \Delta_i Y_j \Delta_j) + |E Y_i \Delta_i| |E Y_j \Delta_j| \\ &\leq O(\Psi(h)). \end{aligned} \quad (4.157)$$

On va maintenant traiter le terme  $s_{n,2}^2$ . On peut écrire :

$$\begin{aligned} s_{n,2}^2 &= \sum_{i=1}^n \text{var}(Y_i \Delta_i) + \sum_{i=1}^n \sum_{0 < |i-j| \leq v_n} \text{cov}(Y_i \Delta_i Y_j \Delta_j) \\ &\quad + \sum_{i=1}^n \sum_{|i-j| > v_n} \text{cov}(Y_i \Delta_i Y_j \Delta_j), \end{aligned} \quad (4.158)$$

où  $(v_n)$  est une suite de nombres réels positifs. Le premier terme de droite de l'égalité (4.158) peut être traité au moyen de (4.38). Le second terme sera traité au moyen de (4.157). Le troisième peut être traité en utilisant l'inégalité de covariance donnée par la proposition 10. On arrive finalement à :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n \Psi(h)) \\ &\quad + O\left(\left(|E(Y_i^p \Delta_i^p)| |E(Y_j^p \Delta_j^p)|\right)^{\frac{1}{p}} \sum_{i=1}^n \sum_{|i-j| > v_n} \alpha(|i-j|)^{1-\frac{2}{p}}\right) \end{aligned} \quad (4.159)$$

En utilisant (4.39) avec l'hypothèse géométrique sur les coefficients de mélange, on obtient pour  $0 < t < 1$  :

$$s_{n,2}^2 = O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n \Psi(h)) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 t^{v_n} \left(1-\frac{2}{p}\right)\right) \quad (4.160)$$

Ce dernier résultat est disponible pour tout  $p > 2$ , car le noyau  $K$  est borné et la variable  $Y$  vérifie la condition des moments (4.34). Si on choisit  $v_n = \varphi_x(h)^{-\epsilon_1}$  alors (4.160) devient :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 t^{\nu \varphi_x(h)^{\epsilon_1}}\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} \left(\frac{1}{\varphi_x(h)^{\frac{p-2}{p}}}\right) n e^{-\frac{b}{\varphi_x(h)^{\epsilon_1}}}\right), \end{aligned}$$

pour  $b > 0$  et pour tout  $\eta > 0$ , on a :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} (n\varphi_x(h)^\eta)\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right). \end{aligned} \quad (2.161)$$

Ce dernier résultat vient de la condition (4.156) et pour  $\eta$  suffisamment grand. De la même manière comme ci-dessus en faisant  $Y = 1$ , on arrive à :

$$s_{n,1}^2 = O\left(\frac{n}{\varphi_x(h)}\right) \quad (4.162)$$

■

**Corollaire 46** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est géométriquement  $\alpha$ -mélangeante.*

*Supposons que les conditions du théorème 29 avec (4.154), (4.155) et (4.156) sont vérifiées.*

*Alors on a :*

$$\hat{r}(x) - r(x) = O\left(h^\beta\right) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) \quad (4.163)$$

**Preuve.** Puisque la condition (4.150) est vérifiée en utilisant (4.156) et le lemme 43 alors la combinaison du théorème 41 et du lemme 43 suffit de prouver ce corollaire. ■

### 4.3.2 Application du cas de mélange algébrique

On va présenter quelque résultat concernant les variables algébriquement  $\alpha$ -mélangeante.

**Lemme 47** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est algébriquement  $\alpha$ -mélangeante avec*

*le coefficient de mélange*

$$a > \frac{1 + \epsilon_1}{\epsilon_1 \epsilon_2}. \quad (4.164)$$



Supposons que les conditions du théorème 29 sont vérifiées. Si en plus les conditions (4.154), (4.155) et (4.156) sont vérifiées. Alors pour  $l = 1$  ou  $2$  on a :

$$s_{n,1}^2 = O\left(\frac{n}{\varphi_x(h)}\right)$$

**Preuve.** En suivant les mêmes étapes de la preuve du lemme 43, on a :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n\Psi(h)) \\ &\quad + O\left(\left(|E(Y_i^p\Delta_i^p)| |E(Y_j^p\Delta_j^p)|\right)^{\frac{1}{p}} \sum_{i=1}^n \sum_{|i-j|>v_n} \alpha(|i-j|)^{1-\frac{2}{p}}\right) \end{aligned}$$

En utilisant (4.39) avec l'hypothèse algébrique sur les coefficients de mélange, on obtient pour tout  $p > 2$  :

$$s_{n,2}^2 = O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n\Psi(h)) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 v_n^{-a\left(1-\frac{2}{p}\right)}\right)$$

En choisissant  $v_n = \varphi_x(h)^{-\epsilon_1}$ , cette dernière inégalité s'écrit :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 \varphi_x(h)^{a\epsilon_1\left(1-\frac{2}{p}\right)}\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} \left(n\varphi_x(h)^{(a\epsilon_1-1)\left(1-\frac{2}{p}\right)}\right)\right) \end{aligned}$$

Puisque de la condition (4.160), on peut toujours choisir  $p$  tel que :

$$\epsilon_2(a\epsilon_1 - 1) \left(1 - \frac{2}{p}\right) > 1.$$

Donc pour une valeur choisie de  $p$  et en utilisant la condition (4.156) on obtient :

$$\begin{aligned} s_{n,2}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} \left(n^{1-\epsilon_2(a\epsilon_1-1)\left(1-\frac{2}{p}\right)}\right)\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right). \end{aligned}$$

En appliquant ces résultats pour le cas où  $Y = 1$  on obtient :

$$s_{n,1}^2 = O\left(\frac{n}{\varphi_x(h)}\right)$$

et la preuve du lemme est achevée. ■

**Corollaire 48** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est algébriquement  $\alpha$ -mélangeante d'ordre  $a$ . Supposons que les conditions du théorème 29 avec (4.154), (4.155), (4.156) et (4.164) sont vérifiées. Alors on a :*

$$\hat{r}(x) - r(x) = O(h^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right) \quad (4.165)$$

**Preuve.** Puisque la condition (4.149) est vérifiée en utilisant (4.156) et le lemme 45 alors la combinaison du théorème 41 et du lemme 45 suffit de prouver ce corollaire. ■

On va étudier le problème de prédiction du scalaire  $Y$  sachant le prédicteur fonctionnel  $X$  au moyen des quantiles conditionnels. Le but est de montrer comment les résultats asymptotiques de la fonction de répartition conditionnelle et des quantiles conditionnels peuvent s'étendre en cas de mélange fort des variables fonctionnelles. Pour cette raison on a besoin de quelques notations générales et définitions :

La fonction de répartition conditionnelle est définie par :

$$\forall y \in \mathbb{R}, F_Y^X(x, y) = P(Y \leq y / X = x). \quad (4.166)$$

Le quantile conditionnel d'ordre  $\alpha \in (0, 1)$  est :

$$t_\alpha(x) = \inf \{y \in \mathbb{R}, F_Y^X(x, y) \geq \alpha\}. \quad (4.167)$$

La médiane conditionnelle est :

$$m(x) = \inf \left\{ y \in \mathbb{R}, F_Y^X(x, y) \geq \frac{1}{2} \right\}. \quad (4.168)$$

On a déjà vu que les estimateurs à noyau de  $F_Y^X$ ,  $t_\alpha(x)$  et  $m(x)$  sont définis respectivement par :

$$\hat{F}_Y^X = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i)) H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}, \quad (4.169)$$

$$\hat{t}_\alpha(x) = \inf \left\{ y \in \mathbb{R}, \hat{F}_Y^X(x, y) \geq \alpha \right\}, \quad (4.170)$$

et

$$\hat{m}(x) = \hat{t}_{\frac{1}{2}}(x) \quad (4.171)$$

où  $K$  est un noyau symétrique,  $H$  est un noyau intégrable et  $h$  et  $g$  sont des paramètres de lissage qui dépendent de la taille de l'échantillon  $n$ .

## 4.4 La convergence presque complète des quantiles conditionnels

Les résultats généraux donnés par le théorème ci-dessous sont l'extention des deux théorèmes 26 et 36 au cas de mélange fort des variables fonctionnelles. L'influence de la structure de mélange sur les vitesses de convergences apparait à travers les quantités suivantes :

$$s_{n,3,0} = \sum_{i=1}^n \sum_{j=1}^n cov(\Gamma_i(y) \Delta_i, \Gamma_j(y) \Delta_j), \quad (4.172)$$

$$s_{n,3,l} = \sum_{i=1}^n \sum_{j=1}^n cov(\Gamma_i^{(l)}(y) \Delta_i, \Gamma_j^{(l)}(y) \Delta_j), \quad (4.173)$$

et

$$s_{n,1} = \sum_{i=1}^n \sum_{j=1}^n cov(\Delta_i, \Delta_j), \quad (4.174)$$

avec

$$\Delta_i = \frac{K(h^{-1}d(x, X_i))}{EK(h^{-1}d(x, X_i))},$$

et

$$\Gamma_i(y) = H(g^{-1}(y - Y_i)).$$

On note que  $s_n$  est une suite des entiers naturels qui appartient à l'ensemble  $\{s_{n,1}, s_{n,3,l}, l = 0, \dots, j\}$ .

En comparaison avec des résultats du cas indépendant on a besoin d'ajouter les hypothèses suivantes qui servent à contrôler ces covariances :

$$\left\{ \begin{array}{l} (X_i, Y_i)_{i=1, \dots, n} \text{ sont fortement mélangeante, avec} \\ \text{les coefficients algébriques d'ordre } a > 1, \text{ et} \\ \exists \theta > 2, s_n^{-(a+1)} = o(n^{-\theta}) \end{array} \right\} \quad (4.175)$$

où

$$\left\{ \begin{array}{l} (X_i, Y_i)_{i=1, \dots, n} \text{ sont fortement mélangeante} \\ \text{avec le coefficients géométriques, et} \\ \exists \theta > 1, s_n^{-1} = o(n^{-\theta}) \end{array} \right\} \quad (4.176)$$

**Théorème 49** *i) Posons  $s_n = \max\{s_{n,1}, s_{n,3,0}\}$ , et supposons que (4.175) or (4.176) sont vérifiées avec les conditions du théorème 26. Alors on a :*

$$\lim_{n \rightarrow \infty} \hat{t}_\alpha(x) = t_\alpha(x), p.co.$$

*ii) Posons  $s_n = \max\{s_{n,1}, s_{n,3,l}, l = 0, \dots, j\}$ , et supposons que (4.175) or (4.176) sont vérifiées avec les conditions du théorème 36. Alors on a :*

$$\hat{t}_\alpha(x) - t_\alpha(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{s_n^2 \log n}{n^2}\right)^{\frac{1}{2j}}\right).$$

**Preuve.** i) Soit  $\epsilon > 0$  fixé. La continuité de  $\hat{F}_Y^{x-1}$  au point  $\hat{F}_Y^x(t_\alpha(x))$  nous permet d'écrire :

$$\exists \delta(\epsilon) > 0, \forall y, \left| \hat{F}_Y^x(y) - \hat{F}_Y^x(t_\alpha(x)) \right| \leq \delta(\epsilon) \rightarrow |y - t_\alpha(x)| \leq \epsilon.$$

En prenant  $y = \hat{t}_\alpha(x)$ , on arrive finalement au résultat suivant :

$$\begin{aligned} \exists \delta(\epsilon) &> 0, P\left(|\hat{t}_\alpha(x) - t_\alpha(x)| > \epsilon\right) \\ &\leq P\left(\left|\hat{F}_Y^x(\hat{t}_\alpha(x)) - \hat{F}_Y^x(t_\alpha(x))\right| > \delta(\epsilon)\right) \\ &= P\left(\left|F_Y^x(t_\alpha(x)) - \hat{F}_Y^x(t_\alpha(x))\right| > \delta(\epsilon)\right), \end{aligned}$$

En utilisant la convergence *p.co.* de l'estimateur de la fonction de répartition (voir le lemme 48 ci-dessous) on obtient le résultat du théorème énoncé.

ii) Le développement de Taylor de la fonction  $\hat{F}_Y^x$  nous conduit à l'existence de  $t^*$  compris entre  $\hat{t}_\alpha(x)$  et  $t_\alpha(x)$  tel que :

$$\begin{aligned} \hat{F}_Y^x(t_\alpha(x)) - \hat{F}_Y^x(\hat{t}_\alpha(x)) &= \sum_{l=1}^{j-1} \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^l}{l!} \hat{F}_Y^{x(l)}(t_\alpha(x)) \\ &\quad + \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^j}{j!} \hat{F}_Y^{x(j)}(t^*). \end{aligned}$$

Selon (4.96), cette égalité s'écrit :

$$\begin{aligned} \hat{F}_Y^x(t_\alpha(x)) - \hat{F}_Y^x(\hat{t}_\alpha(x)) &= \sum_{l=1}^{j-1} \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^l}{l!} \left( \hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x)) \right) \\ &\quad + \frac{(t_\alpha(x) - \hat{t}_\alpha(x))^j}{j!} \hat{F}_Y^{x(j)}(t^*). \end{aligned}$$

Puisque  $\hat{F}_Y^x(\hat{t}_\alpha(x)) = F_Y^x(t_\alpha(x)) = \alpha$ , on a :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j \hat{F}_Y^{x(j)}(t^*) &= O\left(\hat{F}_Y^x(t_\alpha(x)) - F_Y^x(t_\alpha(x))\right) + \\ &O\left(\sum_{l=1}^{j-1} (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x))\right)\right) \end{aligned}$$

La combinaison du résultat du lemme 33 et du théorème 26 , avec  $t^*$  compris entre  $\hat{t}_\alpha(x)$  et  $t_\alpha(x)$ , nous amène à :

$$\lim_{n \rightarrow \infty} \hat{F}_Y^{x(j)}(t^*) = F_Y^{x(j)}(t_\alpha(x)), p.co.$$

Puisque la deuxième partie de l'hypothèse (4.140) assure que cette limite est différente de 0, alors en utilisant la proposition 6 – ii on arrive à :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j &= O\left(\hat{F}_Y^x(t_\alpha(x)) - F_Y^x(t_\alpha(x))\right), p.co + \\ &O\left(\sum_{l=1}^{j-1} (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\hat{F}_Y^{x(l)}(t_\alpha(x)) - F_Y^{x(l)}(t_\alpha(x))\right)\right), p.co. \end{aligned}$$

Selon (4.110), et pour tout  $l \in \{0, \dots, j\}$  et pour tout  $y$  dans le voisinage de  $t_\alpha(x)$ , il existe  $t^*$  compris entre  $y$  et  $t_\alpha(x)$  tel que :

$$\hat{F}_Y^{x(l)}(y) - F_Y^{x(l)}(t_\alpha(x)) = \frac{(y - t_\alpha(x))^{j-l}}{(j-l)!} \hat{F}_Y^{x(j)}(t^*).$$

Ceci implique que  $F_Y^{x(l)}$  est Lipschitzienne autour de  $t_\alpha(x)$  à l'ordre  $j-l$ . Donc, en utilisant le lemme 48 et le lemme 49 qu'on va prouver ci-dessous avec l'ordre de Lipschitz adapté, on obtient :

$$\begin{aligned} (t_\alpha(x) - \hat{t}_\alpha(x))^j &= O(h^\beta + g^\beta) + O\left(\frac{\sqrt{(\max\{s_{n,1}, s_{n,3,0}\})^2 \log n}}{n}\right) \\ &+ O\left(\sum_{l=1}^{j-1} A_{n,l}^*\right) p.co + O\left(\sum_{l=1}^{j-1} B_{n,l}^*\right), p.co, \end{aligned}$$

où

$$A_{n,l}^* = (t_\alpha(x) - \hat{t}_\alpha(x))^l \left(\frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{ng^{2l-1}}\right)$$

et

$$B_{n,l} = (t_\alpha(x) - \hat{t}_\alpha(x))^l g^{j-l}.$$

En faisant la comparaison de  $A_{n,l}^*$  et  $B_{n,l}^*$  avec  $(t_\alpha(x) - \hat{t}_\alpha(x))^j$ . Si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(A_{n,l})$ , on peut écrire :

$$\begin{aligned} |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C |t_\alpha(x) - \hat{t}_\alpha(x)|^l \left( \frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{ng^{2l-1}} \right), \\ \implies |t_\alpha(x) - \hat{t}_\alpha(x)|^{j-l} &\leq C \left( \frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{ng^{2l-1}} \right), \\ \implies |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C \left( \frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{ng^{2l-1}} \right)^{\frac{j}{j-l}}. \end{aligned}$$

donc selon (4.97), il existe  $l$  pour lequel :  $|t_\alpha(x) - \hat{t}_\alpha(x)|^j = O(A_{n,l})$ , alors on a :

$$|t_\alpha(x) - \hat{t}_\alpha(x)|^j = O\left(\frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{n}\right)$$

De la même manière, si on suppose qu'il existe  $l \in \{1, \dots, j-1\}$  tel que  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(B_{n,l})$ , on peut écrire :

$$\begin{aligned} |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C |t_\alpha(x) - \hat{t}_\alpha(x)|^l g^{j-l} \\ \implies |t_\alpha(x) - \hat{t}_\alpha(x)|^j &\leq C g^j. \end{aligned}$$

Donc il existe  $l$  tel que :  $(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(B_{n,l})$ , alors on a :

$$(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(g^\beta)$$

Finalement la combinaison de (4.132), (4.133) et (4.134) nous donne :

$$(t_\alpha(x) - \hat{t}_\alpha(x))^j = O(h^\beta + g^\beta) p.co + O\left(\left(\frac{s_n^2 \log n}{n^2}\right)^{\frac{1}{2}}\right), p.co.$$

■

**Lemme 50** *Sous les conditions du théorème 47 – i, pour tout réel  $y$  fixé on a :*

$$\lim_{n \rightarrow \infty} \hat{F}_Y^x(y) = F_Y^x(y), p.co.$$

**Preuve.** La preuve de ce lemme est identique à celle du lemme 23, il suffit donc de garder les mêmes notations. Comme les noyaux  $K$  et  $H$  sont supposés bornés, alors le terme  $\hat{r}_3(x)$  est la somme de variables  $\alpha$ -mélangeantes bornées.

$$\hat{r}_3(x) = \frac{1}{n} \sum_{i=1}^n \Gamma_i(y) \Delta_i$$

ensuite la condition (4.175) nous permet d'appliquer le corollaire 12-ii et on obtient :

$$\hat{r}_3(x) - E\hat{r}_3(x) = O\left(\frac{\sqrt{s_{n,3,0} \log n}}{n}\right) \quad (4.177)$$

Enfin la preuve de ce lemme vient directement de (4.48), (4.49), (4.153) et (4.177).

■

**Lemme 51** *Sous les conditions du théorème 47, pour tout nombre réel fixé  $y$  on a :*

$$F_Y^x(y) - \hat{F}_Y^x(y) = O(h^\beta + g^{\beta_0}) + O\left(\frac{\sqrt{(\max\{s_{n,1}, s_{n,3,0}\})^2 \log n}}{n}\right), p.co.$$

**Preuve.** Ce résultat s'obtient en faisant (4.48), (4.105), (4.153) et (4.177). ■

**Lemme 52** *Soit  $l$  un entier de  $\{1, \dots, j\}$  et soit  $y$  un nombre réel fixé. Sous les conditions du théorème 47-ii, on a :*

$$\lim_{n \rightarrow \infty} \hat{F}_Y^{x(l)}(y) = F_Y^{x(l)}(y), p.co.$$

*Si en outre la condition (4.111) est vérifiée alors on a :*

$$F_Y^{x(l)}(y) - \hat{F}_Y^{x(l)}(y) = O(h^\beta + g^{\beta_0}) + O\left(\frac{\sqrt{(\max\{s_{n,1}, s_{n,3,l}\})^2 \log n}}{ng^{2l-1}}\right), p.co.$$



**Preuve.** La preuve de ce lemme est similaire à celle de la preuve du lemme 33. En faisant les mêmes notations que celles utilisées dans la décomposition (4.113) et (4.117). Grace à (4.117), (4.118) et (4.119), et en utilisant l'inégalité de Fuk-Nagaev donnée par le corollaire 13-ii, on obtient :

$$\hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) = O\left(\frac{\sqrt{s_{n,3,l} \log n}}{n}\right), p.co. \quad (4.178)$$

Le premier résultat du lemme 50 vient de (4.113), (4.115), (4.153) et (4.178). Le second vient de (4.113), (4.116), (4.153) et (4.178). ■

Le résultat général donné dans le théorème 47-ii peut être donné sous plusieurs formulations différentes, selon la connaissance que nous avons sur le terme de covariance  $s_n^2$ . On va présenter les deux cas particuliers ci-dessous (pour ces deux cas le terme de covariance a le même comportement que dans le cas i.i.d). Ainsi, les résultats concernant  $s_n^2$  sont donnés par les lemmes 51 et 53 et leurs applications pour la vitesse de convergence de l'estimateur du quantile conditionnel sont données par les corollaires 52 et 54.

Pour les conditions on suppose que :

$$\exists \epsilon_1 \in ]0, 1], 0 < \psi_x(h) = O\left(\varphi_x(h)^{1+\epsilon_1}\right), \quad (4.179)$$

$$\exists \epsilon_2 \in ]0, 1], g\varphi_x(h) = O\left(n^{-\epsilon_2}\right), \quad (4.180)$$

et que pour tout  $i \pm j$  :

$$\text{la fonction de densité conditionnelle } f_{i,j} \text{ de } (Y_i, Y_j) \quad (4.181)$$

sachant  $(X_i, X_i)$  existe et bornée.

#### 4.4.1 Application du cas de mélange géométrique

**Lemme 53** *Supposons que les conditions du théorème 36 sont vérifiées. Supposons aussi que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est géométriquement  $\alpha$ -mélangeante vérifiant (4.179), (4.180) et (4.181), alors on a :*

$$s_{n,3,0} = O\left(\frac{n}{\varphi_x(h)}\right)$$

et pour tout  $l = 1, \dots, j$  :

$$s_{n,3,l} = O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right).$$

**Preuve.** Pour tout  $l \in \{0, 1, \dots, j\}$  en faisant la décomposition suivante :

$$\begin{aligned} s_{n,3,l}^2 &= \sum_{i=1}^n \text{var}\left(\Gamma_i^{(l)}(y) \Delta_i\right) + \sum_{0 < |i-j| \leq v_n} \sum \text{cov}\left(\Gamma_i^{(l)}(y) \Delta_i, \Gamma_j^{(l)}(y) \Delta_j\right) \\ &+ \sum_{|i-j| > v_n} \sum \text{cov}\left(\Gamma_i^{(l)}(y) \Delta_i, \Gamma_j^{(l)}(y) \Delta_j\right), \end{aligned} \quad (4.182)$$

où  $v_n$  est une suite positive des nombres réels.

Montrons le résultat du lemme 51 pour  $l = 0$ . Comme  $\mathbb{H}$  est borné, en appliquant le résultat (4.157) on obtient :

$$\text{cov}\left(\Gamma_i(y) \Delta_i, \Gamma_j(y) \Delta_j\right) = O(\Psi(h)), \quad (4.183)$$

avec

$$\Psi(h) = \max\left\{\frac{\psi(h)}{\varphi_x(h)^2}, 1\right\}$$

Le premier terme de droite de l'égalité (4.182) peut être traité au moyen de (4.115) et (4.119). Le second terme sera traité au moyen de (4.183). Le troisième peut être traité en utilisant l'inégalité de covariance donnée par la proposition 10-ii. On arrive finalement

à :

$$\begin{aligned}
s_{n,3,0}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n\Psi(h)) \\
&\quad + O\left(\left(|E(\Gamma_i(y)^p \Delta_i^p)| |E(\Gamma_j(y)^p \Delta_j^p)|\right)^{\frac{1}{p}} \sum_{i=1}^n \sum_{|i-j| > v_n} \alpha(|i-j|)^{1-\frac{2}{p}}\right) \\
&= O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n\Psi(h)) \\
&\quad + O\left(\frac{1}{(E\Gamma_i(y)^p \Delta_i^p)^{\frac{2}{p}}} n^2 \alpha(v_n)^{1-\frac{2}{p}}\right) \\
&= O\left(\frac{n}{\varphi_x(h)}\right) + O(nv_n\Psi(h)) \\
&\quad + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 \alpha(v_n)^{1-\frac{2}{p}}\right)
\end{aligned} \tag{4.184}$$

En prenant  $v_n = \varphi_x(h)^{-\epsilon_1}$  (pour traiter  $O(nv_n\Psi(h))$ ) et en utilisant la condition géométrique sur les coefficients de mélange on obtient pour  $0 < t < 1$  :

$$\begin{aligned}
s_{n,3,0}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 t^{\frac{p-2}{p\varphi_x(h)^{\epsilon_1}}}\right) \\
&= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 e^{-\frac{b}{\varphi_x(h)^{\epsilon_1}}}\right).
\end{aligned}$$

pour  $b > 0$ . Ceci est suffisant pour montrer :  $s_{n,3,0}^2 = O\left(\frac{n}{\varphi_x(h)}\right)$  car la condition (4.56) est vérifiée.

Montrons maintenant le second résultat du lemme 51 pour  $l \in \{1, \dots, j\}$ . En faisant les mêmes arguments utilisés pour la preuve de (4.119), et en utilisant la condition

(4.181), on obtient pour tout  $i \neq j$  :

$$\begin{aligned}
& E\Gamma_i^{(l)}(y) \Delta_i \Gamma_j^{(l)}(y) \Delta_j \\
&= \frac{1}{g^{2l}} E \left( \Delta_i \Delta_j \int_{\mathbb{R}^2} H^{(l)} \left( \frac{y-u_1}{g} \right) H^{(l)} \left( \frac{y-u_2}{g} \right) f_{i,j}(u_1, u_2) du_1 du_2 \right) \\
&= \frac{1}{g^{2l-2}} E \left( \Delta_i \Delta_j \int_{\mathbb{R}^2} H^{(l)}(v_1) H^{(l)}(v_2) f_{i,j}(y-gv_1, y-gv_2) dv_1 dv_2 \right) \\
&\leq C \frac{1}{g^{2l-2}} E \Delta_i \Delta_j \\
&\leq C \frac{1}{g^{2l-2}} \frac{1}{\varphi_x(h)^2} P((X_i, X_j) \in B(x, h) \times B(x, h)) \\
&= O \left( \frac{\psi_{i,j}(h)}{\varphi_x(h)^2 g^{2l-2}} \right). \tag{4.185}
\end{aligned}$$

On a aussi pour tout  $p > 0$  :

$$\begin{aligned}
E\Gamma_i^{(l)}(y)^p \Delta_i^p &= \frac{1}{g^{pl}} E \left( \Delta_i^p \int_{\mathbb{R}} H^{(l)} \left( \frac{y-u}{g} \right)^p f_Y^X(u) du \right) \\
&= \frac{1}{g^{p(l-1)}} E \left( \Delta_i^p \int_{\mathbb{R}} H^{(l)}(v)^p f_Y^X(y-gv) dv \right) \\
&\leq C \frac{1}{g^{p(l-1)}} E \Delta_i^p \\
&= O \left( \frac{1}{g^{p(l-1}} \frac{1}{\varphi_x(h)^{p-1}} \right). \tag{4.186}
\end{aligned}$$

La borne de la dernière inégalité vient de (4.41). Enfin de (4.185) et (4.186) on obtient :

$$\text{cov} \left( \Gamma_i^{(l)}(y) \Delta_i, \Gamma_j^{(l)}(y) \Delta_j \right) = O \left( \frac{\Psi(h)}{g^{2l-2}} \right). \tag{4.187}$$

D'autre part, le résultat (4.186) avec l'inégalité de covariance donnée par la proposition

10-ii, nous amène pour tout  $p > 0$  à :

$$\begin{aligned}
& \text{cov} \left( \Gamma_i^{(l)}(y) \Delta_i, \Gamma_j^{(l)}(y) \Delta_j \right) \tag{4.188} \\
&= O \left( \frac{1}{E \left( \Gamma_1^{(l)}(y) \Delta_1 \right)^{2p}} \alpha(|i-j|)^{\left(1-\frac{2}{p}\right)} \right) \\
&= O \left( \frac{1}{g^{\frac{2(p(l-1))}{p}} \varphi_x(h)^{p-1}} \alpha(|i-j|)^{\left(1-\frac{2}{p}\right)} \right).
\end{aligned}$$

La variance est identique à celle du cas i.i.d, en utilisant (4.115) et (4.119) on obtient :

$$\text{var} \left( \Gamma_i^{(l)}(y) \Delta_i \right) = O \left( \frac{1}{g^{2l-1} \varphi_x(h)} \right). \quad (4.189)$$

Le premier terme de droite de l'égalité peut être traité par (4.189), le second par (4.187) et le troisième par (4.188). Enfin on a :

$$\begin{aligned} s_{n,3,l}^2 &= O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) + O \left( n v_n \frac{\Psi(h)}{g^{2l-2}} \right) \\ &+ O \left( \frac{1}{g^{\frac{2(p_l-1)}{p}}} \frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} \alpha(|i-j|)^{\left(1-\frac{2}{p}\right)} \right). \end{aligned} \quad (4.190)$$

En choisissant  $v_n = g^{-1} \varphi_x(h)^{-\epsilon_1}$  et en faisant (4.179) on obtient :

$$O \left( n v_n \frac{\Psi(h)}{g^{2l-2}} \right) = O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right)$$

et puisque les coefficients de mélange sont géométriques on obtient pour  $0 < t < 1$  et pour  $b > 0$  :

$$\begin{aligned} s_{n,3,l}^2 &= O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) + O \left( \frac{1}{g^{\frac{2(p_l-1)}{p}}} \frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 t^{\frac{p-2}{p} \frac{1}{g \varphi_x(h)^{\epsilon_1}}} \right) \\ &= O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) + O \left( \frac{1}{g^{\frac{2(p_l-1)}{p}}} \frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 e^{-\frac{b}{g \varphi_x(h)^{\epsilon_1}}} \right) \\ &= O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) + O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \left( \frac{1}{(g \varphi_x(h))^{\frac{p-2}{p}}} n e^{-\frac{b}{g \varphi_x(h)^{\epsilon_1}}} \right) \right). \end{aligned}$$

Qui peut s'écrire pour tout  $\eta > 0$  :

$$\begin{aligned} s_{n,3,l}^2 &= O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) \\ &+ O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \left( \frac{1}{(g \varphi_x(h))^{\frac{p-2}{p}}} n (g \varphi_x(h))^\eta \right) \right), \end{aligned}$$

en utilisant(4.36) on obtient :

$$s_{n,3,l}^2 = O \left( \frac{n}{g^{2l-1} \varphi_x(h)} \right) + O \left( \frac{n}{g^{2l-1} \varphi_x(h)} n^{1-\epsilon_2 \eta \frac{p-2}{p}} \right).$$

Pour  $\eta$  suffisamment grand on arrive à :

$$s_{n,3,l}^2 = O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right).$$

■

**Corollaire 54** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est géométriquement  $\alpha$ -mélangeante.*

*Supposons que les conditions du théorème 36 avec (4.179), (4.180) et (4.181) sont vérifiées alors on a :*

$$\hat{t}_\alpha(x) - t_\alpha(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right).p.co.$$

**Preuve.** La combinaison des résultats des lemmes 49, 50 et 51 nous donne les deux résultats suivants :

$$F_Y^x(y) - \hat{F}_Y^x(y) = O\left(h^\beta + g^\beta\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co.$$

et

$$F_Y^{x(l)}(y) - \hat{F}_Y^{x(l)}(y) = O\left(h^\beta + g^{\beta_0}\right) + O\left(\left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co.$$

La preuve de ce corollaire est exactement celle de la preuve du théorème 26, en utilisant ces deux résultats ci-dessus au lieu des lemmes 32 et 33. ■

#### 4.4.2 Application du cas de mélange algébrique

**Lemme 55** *Supposons que les conditions du théorème 36 sont vérifiées. Supposons aussi que (4.179), (4.180) et (4.181) sont vérifiées et la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est algébriquement  $\alpha$ -mélangeante d'ordre :*

$$a > \frac{1 + \epsilon_1}{\epsilon_1 \epsilon_2} \tag{4.191}$$

on a :

$$s_{n,3,0}^2 = O\left(\frac{n}{\varphi_x(h)}\right),$$

et pour tout  $l = 1, \dots, j$  :

$$s_{n,3,l}^2 = O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right)$$

**Preuve.** La démonstration de ce lemme est similaire à celle du lemme 51 ci-dessus, elle est basée sur la décomposition (4.182).

Montrons que :

$$s_{n,3,0}^2 = O\left(\frac{n}{\varphi_x(h)}\right),$$

le résultat (4.188) reste valide pour ce cas, il suffit maintenant de poser  $v_n = \varphi_x(h)^{-\epsilon_1}$  pour traiter le second terme de droite dans la décomposition (4.184) et qui vient de (4.159). En utilisant la condition algébrique sur les coefficients de mélange et (4.191) on arrive, pour tout  $p$ , à :

$$\begin{aligned} s_{n,3,0}^2 &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}} n^2 \varphi_x(h)^{a\epsilon_1\left(1-\frac{2}{p}\right)}\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} \left(n \varphi_x(h)^{(a\epsilon_1-1)\left(1-\frac{2}{p}\right)}\right)\right) \\ &= O\left(\frac{n}{\varphi_x(h)}\right) + O\left(\frac{n}{\varphi_x(h)} \left(n^{1-\epsilon_2(a\epsilon_1-1)\left(1-\frac{2}{p}\right)}\right)\right). \end{aligned}$$

En choisissant  $p$  de sorte que :  $\epsilon_2(a\epsilon_1-1)\left(1-\frac{2}{p}\right) > 1$ , on obtient :

$$s_{n,3,0}^2 = O\left(\frac{n}{\varphi_x(h)}\right).$$

Montrons maintenant le résultat du lemme 53 pour  $l \in \{1, \dots, j\}$ . Le résultat (4.190) reste inchangé pour ce cas. En prenant  $v_n = \varphi_x(h)^{-\epsilon_1}$  pour traiter le second terme de (4.190),

et puisque les coefficients de mélange sont de décroissance algébrique d'ordre  $a$ , alors pour tout  $p$ , on arrive à :

$$\begin{aligned} s_{n,3,l}^2 &= O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right) + \\ &O\left(\frac{1}{g^{\frac{2(p-1)}{p}}}\frac{1}{\varphi_x(h)^{\frac{2p-2}{p}}}n^2(g\varphi_x(h)^{\epsilon_1})^{a\left(1-\frac{2}{p}\right)}\right) \\ &= O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right) + \\ &O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\left(\frac{1}{(g\varphi_x(h))^{\frac{p-2}{p}}}n(g\varphi_x(h))^{a\epsilon_1\left(1-\frac{2}{p}\right)}\right)\right) \end{aligned}$$

D'autre part selon la condition (4.191) on peut choisir  $p$  tel que :

$$\epsilon_2(a\epsilon_1 - 1)\left(1 - \frac{2}{p}\right) > 1$$

Donc, en choisissant  $p$  et en utilisant (4.180), on a :

$$\begin{aligned} s_{n,3,l}^2 &= O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right) + O\left(\frac{n}{g^{2l-1}\varphi_x(h)}n^{1-\epsilon_2(a\epsilon_1-1)\left(1-\frac{2}{p}\right)}\right) \\ &= O\left(\frac{n}{g^{2l-1}\varphi_x(h)}\right). \end{aligned}$$

■

**Corollaire 56** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est algébriquement  $\alpha$ -mélangeante vérifiant (4.191). Supposons que les conditions du théorème 36 et (4.179), (4.180) et (4.181) sont vérifiées. Alors on a :*

$$\hat{t}_\alpha(x) - t_\alpha(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right), p.co.$$

**Preuve.** La combinaison des résultats des lemmes 49, 50 et 53 nous permet d'écrire :

$$F_Y^x(y) - \hat{F}_Y^x(y) = O\left(h^\beta + g^\beta\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right), p.co.$$



et

$$F_Y^{x^{(l)}}(y) - \hat{F}_Y^{x^{(l)}}(y) = O(h^\beta + g^{\beta_0}) + O\left(\left(\frac{\log n}{ng^{2l-1}\varphi_x(h)}\right)^{\frac{1}{2}}\right), p.co.$$

Donc la preuve du corollaire 54 est exactement celle de la preuve du théorème 26, mais en utilisant ces deux résultats ci-dessus au lieu des lemmes 32 et 33. ■

On va étudier le problème de prédiction du scalaire  $Y$  sachant la variable fonctionnel  $X$  au moyen du mode conditionnel. Le but est de montrer comment les résultats asymptotiques de la fonction de densité conditionnelle et du mode conditionnel peuvent s'étendre en cas du mélange fort des variables fonctionnelles. Pour cette raison on a besoin de quelques notations générales et définitions :

$$\forall y \in \mathbb{R}, f_Y^X(x, y) = \frac{\partial}{\partial y} F_Y^X(x, y) \quad (4.192)$$

Le mode conditionnel, qui est supposé dans le sous ensemble  $S \subset \mathbb{R}$  est défini par :

$$\theta(x) = \arg \sup_{y \in S} f_Y^X(x, y) \quad (4.193)$$

Les estimateurs à noyau de  $f_Y^X$  et  $\theta(x)$  sont respectivement définis par :

$$f_Y^{\hat{x}}(x, y) = \frac{g^{-1} \sum_{i=1}^n K(h^{-1}d(x, X_i)) K_0(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))} \quad (4.194)$$

et

$$\hat{\theta}(x) = \arg \sup_{y \in S} f_Y^{\hat{x}}(x, y) \quad (4.195)$$

avec  $K$  et  $K_0$  sont des noyaux symétriques et  $h$  et  $g$  sont des paramètres de lissages.

## 4.5 La convergence presque complète du mode conditionnelle

Les résultats généraux donnés par le théorème 55 ci dessous est l'extention des deux théorèmes 24 et 34 dans le cas du mélange fort des variables fonctionnelles. L'influence de la structure de mélange sur les vitesses de convergence apparait à travers les quantités suivantes :

$$s_{n,4} = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(\Omega_i(y) \Delta_i, \Omega_j(y) \Delta_j), \quad (4.196)$$

$$s_{n,l} = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(\Delta_i, \Delta_j), \quad (4.197)$$

où

$$\Delta_i = \frac{K(h^{-1}d(x, X_i))}{EK(h^{-1}d(x, X_i))},$$

et

$$\Omega_i(y) = g^{-1}K_0(g^{-1}(y - Y_i))$$

On note que  $s_n$  est une suite des entiers naturels qui appartient à l'ensemble des  $s_{n,j}$  définis ci-dessus. En comparaison avec les résultats dans le cas indépendant on a besoin d'ajouter les hypothèses suivantes qui servent à contrôler les covariances :

$$\left\{ \begin{array}{l} (X_i, Y_i)_{i=1, \dots, n} \text{ sont fortement mélangeante,} \\ \text{avec le coefficient algébrique d'ordre } a > 1, \text{ et} \\ \exists \theta > 2(1 + \zeta), s_n^{-(a+1)} = o(n^{-\theta}), \end{array} \right\} \quad (4.198)$$

Où

$$\left\{ \begin{array}{l} (X_i, Y_i)_{i=1, \dots, n} \text{ sont fortement mélangée} \\ \text{avec le coefficient géométrique, et} \\ \exists \theta > 1, s_n^{-1} = o(n^{-\theta}). \end{array} \right\} \quad (4.199)$$

**Théorème 57** *i) Posons  $s_n = \max\{s_{n,1}, s_{n,4}\}$  et supposons que (4.198) ou (4.199) soit vérifiée avec les conditions du théorème 24. Alors on a :*

$$\lim_{n \rightarrow \infty} \hat{\theta}(x) = \theta(x), p.co.$$

*ii) Posons  $s_n = \max\{s_{n,1}, s_{n,3,l}, l = 0, \dots, j\}$ , et supposons que (4.198) ou (4.199) soit vérifiée avec les conditions du théorème 34. Alors on a :*

$$\hat{\theta}(x) - \theta(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{s_n^2 \log n}{n^2}\right)^{\frac{1}{2j}}\right).$$

**Preuve.** Le résultat du lemme 25 reste vrai dans le cas du mélange fort. Ceci nous amène à énoncer le lemme 56 ci-dessous, et la preuve de i) du théorème 55 vient directement de (4.63) et du lemme 56.

La preuve de ii) a besoin de s'étendre le lemme 35 aux variables dépendantes, c'est le lemme 56. Le résultat ii) du théorème 55 vient directement de (4.128) et de l'application du lemme 56 pour  $l = 1$  et  $S = (\theta(x) - \xi, \theta(x) + \xi)$ . ■

**Lemme 58** *Sous les conditions du théorème 55 – i, on a pour tout compact  $S \subset \mathbb{R}$  :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} \left| f_Y^{\hat{x}}(y) - f_Y^x(y) \right| = 0, p.co.$$

**Preuve.** La preuve du lemme 56 est similaire à celle du lemme 25, en effet on s'intéresse aux étapes qui sont affectées par le mélange fort. Puisque la décomposition (2.65)

reste valide, en gardant les mêmes notations du lemme 25. Donc pour prouver le lemme 56, il suffit de démontrer les deux résultats suivants :

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |E\hat{r}_4(x, y) - f_Y^x(y)| = 0, p.co. \quad (4.200)$$

et

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - E\hat{r}_4(x, y)| = 0, p.co. \quad (4.201)$$

Comme le terme du biais reste inchanger en cas de dépendance, le résultat (4.70) reste vrai et (4.200) s'obtient en faisant (4.70), (4.153) et la proposition 5 - i.

Montrons (4.201), en effet puisque S est un compact alors  $S \subset \cup_{k=1}^{z_n} S_k$  où  $S_k = (t_k - l_n, t_k + l_n)$  et  $l_n = Cz_n^{-1} = n^{-2a}$ . En utilisant la décomposition :

$$\begin{aligned} & \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - E\hat{r}_4(x, y)| \\ \leq & \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - \hat{r}_4(x, t_y)|}_{A_1} \\ & + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, t_y) - E\hat{r}_4(x, t_y)|}_{A_2} \\ & + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |E\hat{r}_4(x, t_y) - E\hat{r}_4(x, y)|}_{A_3} \end{aligned} \quad (4.202)$$

-  $A_1$  reste inchanger même en cas de dépendance des variables donc (4.73) est valide c.à.d  $\lim_{n \rightarrow \infty} A_1 = 0$ .

-  $A_3$  se fait en utilisant (4.74) avec (4.135) et la proposition 5-i, on arrive donc à (4.75) c.à.d  $\lim_{n \rightarrow \infty} A_3 = 0$ .

Pour le terme  $A_2$ , en utilisant la décomposition(4.76) c.à.d

$$P \left( \sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon \right) \leq \\ z_n \max_{j=1, \dots, z_n} P \left( \left| \frac{1}{n} \sum_{i=1}^n (U_i - EU_i) \right| > \epsilon \right)$$

comme les variables aléatoires  $U_i$  vérifient la condition de bornitude (4.77), on peut appliquer la proposition 11 – ii, en faisant  $b = a$  sous la condition algébrique (4.198) ou  $b = +\infty$  sous la condition géométrique (4.199) et en prenant  $r = (\log n)^2$ . On a :

$$P \left( \sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon \sqrt{\frac{s_n^2 \log n}{n^2}} \right) \\ \leq C n^{2\zeta} \left\{ \left( 1 + \frac{\epsilon^2}{\log n} \right)^{-\frac{(\log n)^2}{2}} + n (\log n)^{-2} \left( \frac{\sqrt{\log n}}{\epsilon s_n} \right)^{b+1} \right\} \\ \leq C n^{2\zeta} \left\{ e^{-\frac{\epsilon^2 \log n}{2}} + n (\log n)^{-2 + \frac{1+b}{2}} s_n^{-b-1} \epsilon^{-b-1} \right\} \\ \leq C \left\{ n^{2\zeta - \frac{\epsilon^2}{2}} + n^{2\zeta+1} s_n^{-b-1} (\log n)^{-2 + \frac{1+b}{2}} \right\}$$

Maintenant la condition sur  $s_n$  nous permet d'écrire pour  $\epsilon$  suffisamment grand, il existe  $\eta > 0$  tel que :

$$P \left( \sup_{y \in S} |\hat{r}_4(x, t_y) - E \hat{r}_4(x, t_y)| > \epsilon \sqrt{\frac{s_n^2 \log n}{n^2}} \right) = O(n^{-1-\eta}),$$

ce qui prouve que

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |\hat{r}_4(x, y) - E \hat{r}_4(x, t_y)| = 0, p.co.$$

La combinaison de ce dernier résultat avec (4.153) et la proposition 5 – i) nous donne :

$$\lim_{n \rightarrow \infty} A_2 = 0, p.co. \quad (4.203)$$

Enfin la démonstration de (4.201) vient de (4.202), (4.73), (4.75) et (4.203) ■

**Lemme 59** *Sous les conditions du théorème 55 – ii, on a pour tout  $S \subset \mathbb{R}$  et pour tout*

*$l = 1, \dots, j$  :*

$$\sup_{y \in S} \left| F_Y^{x^{(l)}}(y) - \hat{F}_Y^{x^{(l)}}(y) \right| = O\left(h^\beta + g^{\beta_0}\right) + O\left(\sqrt{\frac{s_{n,3,l} \log n}{n^2}}\right), p.co.$$

**Preuve.** - La démonstration de ce lemme est la même que celle du cas indépendant (voir le lemme 35). Cependant, la plupart des étapes de la preuve du lemme 35 restent valable en cas de dépendance, en particulier les résultats (4.113) et (4.131) sont valable c.à.d on a :

$$\begin{aligned} \hat{F}_Y^{x^{(l)}}(y) - F_Y^{x^{(l)}}(y) &= \frac{1}{\hat{r}_1(x)} \left( \hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) \right) \\ &\quad - \frac{1}{\hat{r}_1(x)} \left( F_Y^{x^{(l)}}(y) - E\hat{r}_3^{(l)}(x, y) \right) \\ &\quad + \frac{F_Y^{x^{(l)}}(y)}{\hat{r}_1(x)} (1 - \hat{r}_1(x)). \end{aligned}$$

et

$$\sup_{y \in S} \left| E\hat{r}_3^{(l)}(x, y) - F_Y^{x^{(l)}}(y) \right| = O\left(h^\beta + g^{\beta_0}\right)$$

D'autre part le dénominateur qui apparait dans (4.113) se traite par (4.153) et la proposition 5 – i. Enfin la démonstration est achevée si le résultat suivant est prouvé :

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - E\hat{r}_3^{(l)}(x, y) \right| = O\left(\sqrt{\frac{s_{n,3,l} \log n}{n^2}}\right), p.co. \quad (4.204)$$

Montrons (2.204). Comme  $S$  est un compact alors  $S \subset \cup_{k=1}^{z_n} S_k$  où  $S_k = (t_k - l_n, t_k + l_n)$  et  $l_n$

et  $z_n$  sont choisis de sorte que  $l_n = Cz_n^{-1} \sim Cn^{-(l+1)\varsigma - \frac{1}{2}}$ . En prenant  $t_y = \arg \min_{t \in \{t_1, \dots, t_{z_n}\}} |y - t|$

on a :

$$\begin{aligned}
& \frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - E \hat{r}_3^{(l)}(x, y) \right| \\
\leq & \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, y) - \hat{r}_3^{(l)}(x, t_y) \right|}_{B_1} \\
& + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right|}_{B_2} \\
& + \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} \left| E \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, y) \right|}_{B_3} \tag{4.205}
\end{aligned}$$

-  $B_1$  reste inchanger même en cas de dépendance des variables donc selon (4.135), la dernière condition de (4.61) et la condition sur  $s_{n,3,l}$  (voir (4.198) ou (4.199)) on a :

$$\begin{aligned}
B_1 &= \frac{1}{\hat{r}_1(x)} \left| \hat{r}_3^{(l)}(x, y) - \hat{r}_3^{(l)}(x, t_y) \right| \leq Cl_n g^{-l-1} \tag{4.206} \\
&\leq C (gn^\zeta)^{-l-1} n^{-\frac{1}{2}} \\
&= O\left(n^{-\frac{1}{2}}\right) \\
&= O\left(\sqrt{\frac{s_{n,3,l}^2 \log n}{n^2}}\right)
\end{aligned}$$

-  $B_3$  se fait en utilisant (4.135) avec (4.153) et la proposition 5 - i, ainsi on obtient :

$$B_3 = O\left(\sqrt{\frac{s_{n,3,l}^2 \log n}{n^2}}\right), p.co. \tag{4.207}$$

-  $B_2$  est le seule terme affecté par la dépendance des variables. Notons que nous avons pour tout  $\epsilon > 0$  :

$$\begin{aligned}
& P\left(\sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right| > \epsilon\right) \\
\leq & z_n \max_{j=1, \dots, z_n} P\left(\left| \hat{r}_3^{(l)}(x, t_j) - E \hat{r}_3^{(l)}(x, t_j) \right| > \epsilon\right)
\end{aligned}$$

comme les résultats concernant les bornes obtenus en (4.118) et (4.119) sont inchangés et restent valables en cas de dépendance, donc en appliquant l'inégalité de Fuk Nagaev pour les variables  $\alpha$ -mélangeantes et bornées (voir la proposition 11 – *ii*), en faisant  $b = a$  sous la condition algébrique (4.198) ou  $b = 0$  sous la condition géométrique (4.199) et en prenant  $r = (\log n)^2$ . On a :

$$\begin{aligned}
& P \left( \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_j) - E \hat{r}_3^{(l)}(x, t_j) \right| > \epsilon \sqrt{\frac{s_{n,3,l}^2 \log n}{n^2}} \right) \\
& \leq C n^{2\varsigma} \left\{ \left( 1 + \frac{\epsilon^2}{\log n} \right)^{-\frac{(\log n)^2}{2}} + n (\log n)^{-2} \left( \frac{\sqrt{\log n}}{\epsilon s_{n,3,l}} \right)^{b+1} \right\} \\
& \leq C n^{2\varsigma} \left\{ e^{-\frac{\epsilon^2 \log n}{2}} + n (\log n)^{-2 + \frac{1+b}{2}} s_{n,3,l}^{-b-1} \epsilon^{-b-1} \right\} \\
& \leq C \left\{ n^{2\varsigma - \frac{\epsilon^2}{2}} + n^{1+2\varsigma} s_{n,3,l}^{-b-1} (\log n)^{-2 + \frac{1+b}{2}} \right\}.
\end{aligned}$$

Ces deux derniers résultats prouvent que pour  $\epsilon$  suffisamment grand, il existe  $\eta > 0$  tel que :

$$P \left( \sup_{y \in S} \left| \hat{r}_3^{(l)}(x, t_y) - E \hat{r}_3^{(l)}(x, t_y) \right| > \epsilon \sqrt{\frac{s_{n,3,l}^2 \log n}{n^2}} \right) = O(n^{-1-\eta}),$$

la dernière inégalité vient de la condition sur  $s_{n,3,l}$ . Enfin on arrive à :

$$B_2 = O \left( \sqrt{\frac{s_{n,3,l} \log n}{n^2}} \right), p.co. \quad (2.208)$$

■

Le résultat général donné dans le théorème 55 – *ii* peut être donné sous plusieurs formulations différentes, selon la connaissance que nous avons sur le terme de covariance  $s_n^2$ . On va présenter les 2 cas particuliers de dépendance ci-dessous (pour ces 2 cas le terme de covariance a le même comportement que dans le cas i.i.d). Ainsi, les hypothèses qu'on suppose sont :

$$\exists \epsilon_1 \in ]0, 1], 0 < \psi_x(h) = O \left( \varphi_x(h)^{1+\epsilon_1} \right), \quad (4.209)$$



$$\exists \epsilon_2 \in ]0, 1], g\varphi_x(h) = O(n^{-\epsilon_2}), \quad (4.210)$$

et que pour tout  $i \pm j$  :

$$\text{la fonction de densité conditionnelle } f_{i,j} \text{ de } (Y_i, Y_j) \quad (4.211)$$

sachant  $(X_i, X_j)$  existe et bornée.

#### 4.5.1 Application du cas de mélange géométrique

Le résultat suivant est une conséquence des propriétés obtenues par les sommes des covariances citées dans le lemme 51. La vitesse de convergence de l'estimateur du mode conditionnel pour les variables géométriquement  $\alpha$ -mélangeantes est identique à celle du cas i.i.d.

**Corollaire 60** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est géométriquement  $\alpha$ -mélangeante. Supposons que les conditions du théorème 34 avec (4.209), (4.210) et (4.211) sont vérifiées alors on a :*

$$\hat{\theta}(x) - \theta(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right).p.co. \quad (4.212)$$

**Preuve.** La combinaison des résultats du lemme 57 et 51 nous permet d'écrire :

$$\sup_{y \in S} \left| \hat{F}_Y^{x^{(l)}}(y) - F_Y^{x^{(l)}}(y) \right| = O\left(\left(h^\beta + g^{\beta_0}\right)\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co.$$

Pour  $l = 1$  et  $S = (\theta(x) - \xi, \theta(x) + \xi)$  et en utilisant (2.128) on obtient :

$$\begin{aligned} \left| \hat{\theta}(x) - \theta(x) \right|^j &= O\left(\sup_{y \in (\theta(x) - \xi, \theta(x) + \xi)} \left| \hat{f}_Y^x(y) - f_Y^x(y) \right|\right), p.co. \\ &= O\left(\left(h^\beta + g^\beta\right)\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co. \end{aligned}$$

■

### 4.5.2 Application du cas de mélange algébrique

On va montrer que la vitesse de convergence sous la condition du mélange arithmétique est identique à celle dans le cas i.i.d. On rappelle que  $\epsilon_1$  et  $\epsilon_2$  sont définis par (4.209) et (4.210), et on suppose que :

$$a > \frac{1 + \epsilon_1}{\epsilon_1 \epsilon_2} \quad (4.213)$$

**Corollaire 61** *Supposons que la suite  $(X_i, Y_i)_{i=1, \dots, n}$  est algébriquement  $\alpha$ -mélangeante vérifiant (4.69). Supposons que les conditions du théorème 34 et (4.65), (4.66) et (4.67) sont vérifiées. Alors on a :*

$$\hat{\theta}(x) - \theta(x) = O\left(\left(h^\beta + g^\beta\right)^{\frac{1}{j}}\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2j}}\right).p.co. \quad (4.214)$$

**Preuve.** La combinaison des résultats du lemme 57 et 53 nous permet d'écrire :

$$\sup_{y \in S} \left| \hat{F}_Y^{x^{(l)}}(y) - F_Y^{x^{(l)}}(y) \right| = O\left(\left(h^\beta + g^{\beta_0}\right)\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co.$$

(Ce dernier résultat est une extension du lemme 35 vu dans le cas indépendant dans le cas de mélange algébrique.)

Pour  $l = 1$  et  $S = (\theta(x) - \xi, \theta(x) + \xi)$  et en utilisant (4.128) on obtient :

$$\begin{aligned} \left| \hat{\theta}(x) - \theta(x) \right|^j &= O\left(\sup_{y \in (\theta(x) - \xi, \theta(x) + \xi)} \left| \hat{f}_Y^x(y) - f_Y^x(y) \right|\right), p.co. \\ &= O\left(h^\beta + g^\beta\right) + O\left(\left(\frac{\log n}{n\varphi_x(h)}\right)^{\frac{1}{2}}\right).p.co. \end{aligned}$$

■

### 4.5.3 Complément sur l'estimateur de la distribution conditionnelle

On a déjà vu que tous les lemmes utilisés dépendent des résultats asymptotiques de l'estimateur de la fonction de répartition conditionnelle ou de l'estimateur de la densité

conditionnelle. On peut donc utiliser ces lemmes pour l'étude asymptotique de ces deux fonctions.

**Proposition 62** *i) Sous les conditions du théorème 47 – i, on a pour tout réel fixé  $y$  :*

$$\lim_{n \rightarrow \infty} F_Y^x(y) = \hat{F}_Y^x(y), p.co.$$

*ii) Si de plus la largeur de la fenêtre  $g$  vérifie pour tout  $\zeta > 0$  la condition  $\lim_{n \rightarrow \infty} gn^\zeta = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |F_Y^x(y) - \hat{F}_Y^x(y)| = 0, p.co.$$

**Preuve.** En faisant les mêmes notations que celles utilisées pour la démonstration de la proposition 27 ou la preuve du lemme 48.

i) Ce résultat est déjà prouvé dans le lemme 48.

ii) En faisant les mêmes étapes utilisées pour la preuve de la proposition 27 – ii, et on note que le résultat (4.85) est vrai même pour ce cas de dépendance, alors la seule chose qui nous reste à montrer est :

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |\hat{r}_3(x, y) - E\hat{r}_3(x, y)| = 0, p.co.$$

Comme  $S$  est un compact alors  $S \subset \cup_{k=1}^{z_n} S_k$  où  $S_k = (t_k - l_n, t_k + l_n)$  et  $l_n$  et  $z_n$  sont choisis de sorte que :

$$l_n = Cz_n^{-1} \sim Cn^{-\varsigma}. \quad (4.215)$$

En prenant  $t_y = \arg \min_{t \in \{t_1, \dots, t_{z_n}\}} |y - t|$  on :

$$\begin{aligned} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - E\hat{r}_3(x, y)| &\leq \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - \hat{r}_3(x, t_y)|}_{D_1} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)|}_{D_2} \\ &+ \underbrace{\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |E\hat{r}_3(x, t_y) - E\hat{r}_3(x, y)|}_{D_3} \end{aligned}$$

-  $D_1$  n'est pas affecté par la structure de dépendance, donc le résultat (2.90) est valable.

- Pour traiter  $D_3$ , on utilise (4.89) avec (4.153), la proposition 5 –  $i$  et (4.91) qui reste valable.

- Pour traiter  $D_2$ , on utilise la décomposition (4.76), comme les variables  $\Gamma_i(y)$  et  $\Delta_i$  sont bornées (car  $H$  et  $K$  sont bornés) alors en appliquant la proposition 11 –  $ii$ , en posant  $b = a$  sous la condition algébrique (2.198) ou  $b = +\infty$  sous la condition géométrique (4.199) et  $r = (\log n)^2$ , on a :

$$\begin{aligned} &P \left( \sup_{y \in S} |\hat{r}_3(x, t_y) - E \hat{r}_3(x, t_y)| > \epsilon \sqrt{\frac{s_{n,3,0}^2 \log n}{n^2}} \right) \\ &\leq Cn^\zeta \left\{ \left( 1 + \frac{\epsilon^2}{\log n} \right)^{-\frac{(\log n)^2}{2}} + n (\log n)^{-2} \left( \frac{\sqrt{\log n}}{\epsilon s_{n,3,0}} \right)^{b+1} \right\} \\ &\leq Cn^\zeta \left\{ e^{-\frac{\epsilon^2 \log n}{2}} + n (\log n)^{-2 + \frac{1+b}{2}} s_{n,3,0}^{-b-1} \epsilon^{-b-1} \right\} \\ &\leq C \left\{ n^{\zeta - \frac{\epsilon^2}{2}} + n^{1+\zeta} s_{n,3,0}^{-b-1} (\log n)^{-2 + \frac{1+b}{2}} \right\}. \end{aligned}$$

La condition sur  $s_{n,3,0}$  (voir (4.175) ou (4.176)), nous permet d'écrire que pour  $\epsilon$  suffisamment grand, il existe  $\eta > 0$  tel que :

$$P \left( \sup_{y \in S} |\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)| > \epsilon \sqrt{\frac{s_{n,3,0}^2 \log n}{n^2}} \right) = O(n^{-1-\eta}) \quad (4.116)$$

ce qui implique que :

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |\hat{r}_3(x, t_y) - E\hat{r}_3(x, t_y)| = 0, p.co.$$

la combinaison de ce dernier résultat avec (4.116) et en appliquant la proposition 5 – i on obtient :

$$\lim_{n \rightarrow \infty} A_2 = 0, p.co.$$

Ainsi la preuve est achevée. ■

**Proposition 63** i) *Sous les conditions du théorème 55 – i, on a pour tout réel fixé  $y$  :*

$$\lim_{n \rightarrow \infty} f_Y^x(y) = \hat{f}_Y^x(y), p.co.$$

ii) *Si de plus la largeur de la fenêtre  $g$  vérifie pour tout  $\zeta > 0$  la condition*

*$\lim_{n \rightarrow \infty} gn^\zeta = \infty$  alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} \left| f_Y^x(y) - \hat{f}_Y^x(y) \right| = 0, p.co.$$

**Preuve.** i) est un cas particulier du lemme 50, en posant  $l = 1$ .

ii) est déjà étudié dans le lemme 56. ■

Les deux propositions suivantes caractérisent la vitesse de convergence. On va étudier la vitesse de convergence presque complète de l'estimateur de la fonction de répartition conditionnelle. Dans la première proposition on va présenter le résultat général où la vitesse de convergence est exprimée en fonction des termes de covariance de l'estimateur.

**Proposition 64** *i) Sous les conditions du théorème 47 – ii, on a pour tout réel fixé  $y$  :*

$$F_Y^x(y) - \hat{F}_Y^x(y) = O(h^\beta + g^\beta) + O\left(\frac{\sqrt{(\max\{s_{n,3,0}, s_{n,1}\})^2 \log n}}{n}\right), p.co.$$

*ii) Si de plus la largeur de la fenêtre  $g$  vérifie pour  $\zeta > 0$  la condition  $\lim_{n \rightarrow \infty} gn^\zeta = \infty$ ,*

*alors pour tout compact  $S \subset \mathbb{R}$  on a :*

$$\sup_{y \in S} |F_Y^x(y) - \hat{F}_Y^x(y)| = O(h^\beta + g^\beta) + O\left(\frac{\sqrt{(\max\{s_{n,3,0}, s_{n,1}\})^2 \log n}}{n}\right), p.co.$$

**Preuve.** i) Ce résultat est déjà étudié lors de la preuve du lemme 49.

ii) La preuve de ce résultat est similaire à celle de proposition 37 – ii. Notons que le résultat (4.143) reste valide sous la structure de dépendance. c.à.d

$$\sup_{y \in S} |E\hat{r}_3(x, y) - F_Y^x(y)| = O(h^\beta + g^\beta).$$

Cependant, de (4.153), la proposition 5 – i et (2.216) on obtient :

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_3(x, y) - E\hat{r}_3(x, y)| = O\left(\frac{\sqrt{s_{n,3,0}^2 \log n}}{n}\right), p.co.$$

Ce dernier résultat avec (2.133), (2.48), (2.49) et la proposition 5 – i sont suffisants pour donner le résultat ii). ■

Maintenant on va présenter la vitesse de convergence de l'estimateur de la densité conditionnelle.

**Proposition 65** *i) Sous les conditions du théorème 55 – ii, on a pour tout réel fixé  $y$  :*

$$f_Y^x(y) - \hat{f}_Y^x(y) = O(h^\beta + g^\beta) + O\left(\frac{\sqrt{(\max\{s_{n,4}, s_{n,1}\})^2 \log n}}{n}\right), p.co.$$

ii) Si de plus la largeur de fenêtre  $g$  vérifie pour  $\zeta > 0$  la condition  $\lim_{n \rightarrow \infty} gn^\zeta = \infty$ , alors pour tout compact  $S \subset \mathbb{R}$ , on a :

$$\sup_{y \in S} |f_Y^x(y) - \hat{f}_Y^x(y)| = O(h^\beta + g^\beta) + O\left(\frac{\sqrt{(\max\{s_{n,4}, s_{n,1}\})^2 \log n}}{n}\right), p.co.$$

**Preuve.** Cette proposition n'est qu'un cas particulier du résultat de la seconde partie du lemme 50 et du lemme 57, en prenant  $l = 1$  et  $\beta = \beta_0$ . ■

## Chapitre 5

# Simulation

Cette section est consacrée à la simulation de la fonction de régression dans les deux cas : l'indépendance et le mélange fort.

Pour effectuer cette simulation, nous considérons un échantillon de  $n = 100$  courbes suivant :

$$x_i(t_j) = a_i \cos(2\pi t_j) + b_i \sin(4\pi t_j) + 2c_i (t_j - 0.25) (t_j - 0.5)$$

où  $0 = t_1 < t_2 < \dots < t_{99} < t_{100} = 1$  sont des points équidistants, et  $(a_i)$ ,  $(b_i)$ ,  $(c_i)$  sont des v.a suivant la loi uniforme sur  $[0,1]$ . La figure 1 nous donne l'idée sur la forme de ces courbes. (on peut consulter [AFRV08] pour triter des exemples).

1/ la simulation de la fonction de régression dans le cas indépendant

- i) Choisir une fonction  $r(\cdot)$ ,
- ii) générer les courbes  $x_1, \dots, x_{100}$ ,
- iii) générer indépendamment  $\varepsilon_1, \dots, \varepsilon_{100}$  à partir de la loi normale centrée de variance



égale à 0.05,

iv) simuler la variable à expliquer :

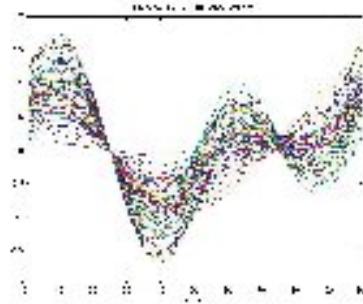
$$Y_i = r(x_i) + \varepsilon_i.$$

En considérant  $r(x) = 100(x - 0.15)^3$  et l'estimateur à noyau

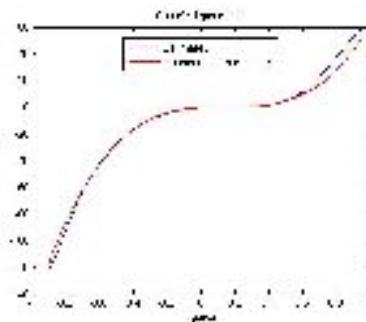
$$\hat{r}(x) = \frac{\sum_{i=1}^n Y_i K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}$$

nous avons obtenu par le logiciel Matlab les graphiques représentés sur les figures 1 et 2.

1



1.jpg



Les deux courbes (Figure 1) sont proches donc les approximations sont bonnes.

Voici programmes utilisé :

```
%%%%%%%%%
```

```
% The Gaussian Kernel function.
```

```

% Returns the standart normal pdf of  $\|x_1-x_2\|/h$ 

function [r]=gau(x1,x2,h)

if nargin==2

h=1;

end;

u=norm((x1-x2)./h);

r=normpdf(u,0,1);

function [yhat]=kern2(y,x,h,xgrid,ker);

[grow gcol]=size(xgrid);

[hrow hcol]=size(h);

if nargin<5 % checking for valid kernel function

ker='epa';

else

if ker~='gau' & ker~='uni' & ker~='tri' & ker~='qua'

ker='epa';

end

end

R=[ker,'(z,0,1)']; %build the string ker(0,z,1) to be evaluated below

if hrow==1 & hcol==1 % global and uniform

hh=ones(grow,gcol)*h;

elseif hrow==1 & hcol==gcol %global but not uniform

hh=kron(ones(grow,1),h);

```

```

elseif hrow==grow & hcol==1 %local and uniform

hh=kron(h,ones(gcol,1));

elseif hrow==grow & hcol==gcol

hh=h;

else

error('bandwidth size don"t match X variable size');

end

w=zeros(grow,size(x,1));

r=zeros(grow,1);

f=zeros(grow,1);

yhat=zeros(grow,1);

for i=1 :grow

for j=1 :size(x,1)

z=norm((x(j, :)-xgrid(i, :))./hh(i, :));

if or(z<=1,ker=='gau')

r(i,1)=r(i,1)+y(j)*eval(R);

f(i,1)=f(i,1)+eval(R);

end;

end;

if f(i,1)==0

f(i,1)=NaN;

end;

```

```

end ;

yhat=r./f;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%a = unifrnd(0,1,1,100);

%b = unifrnd(0,1,1,100);

%c = unifrnd(0,1,1,100);

%for i=1 :1 :100;

%for j=1 :1 :100;

%X(j,i)=a(i)*cos(2*pi*(j/100))+ b(i)*sin(4*pi*(j/100))+ 2*c(i)*((j/100)-.25)*((j/100)-
.5);

%end

%end

%eps=normrnd(0,.05,1,100);

%for i=1 :1 :100

%Y(i)=100*(X(i)-.15).^3+eps(i);

%end

h=.9;

X0=X(1 :100,100);

%for h=.1 :.1 :.9;

yhat = kern(Y, X, h, X0, 'gau');

r=100*(X0-.15).^3;

MSE=[mean((yhat-r).^2)];

```

```

fprintf ('%d.\n',MSE,h);

ytab = sortrows([ MSE h ])

%end

plot(X0,yhat)

hold on;

plot(X0,r,'r')

legend('Estimateur','la fonction de regression' ).

```

La valeur de la largeur de la fenêtre  $h$  a été choisie par balayage selon le critère

$MCE$  (moyenne des carrées des erreurs), comme l'indique le tableau suivant :

$h$	0.1	0.2	0.3	0.4	0.5	0.6	<b>0.7</b>	0.8	0.9
$MCE$	25.0258	21.0330	16.7129	14.1902	11.6192	8.7033	<b>7.1931</b>	8.7739	14.1234

La valeur retenue selon le critère  $MCE = \frac{1}{100} \sum_{i=1}^{100} (Y_i - \hat{Y}_i)^2$  vaut  $h = 0.7$ .

Le choix de l'initialisation de l'algorithme est  $x_0 = x_{100}$ .

2/ La simulation de la fonction de régression dans le cas  $\alpha$ -mélangeant.

i) Choisir une fonction  $r(\cdot)$ ,

ii) générer les courbes  $x_1, \dots, x_{100}$ ,

iii) générer indépendamment  $\eta_1, \dots, \eta_{100}$  à partir de la loi normale centrée de va-

riance égale à 0.05,

iv) simuler les erreurs  $\alpha$ -mélangeantes :

$$e_i = \frac{1}{\sqrt{2}}(e_{i-1} + \eta_i)$$

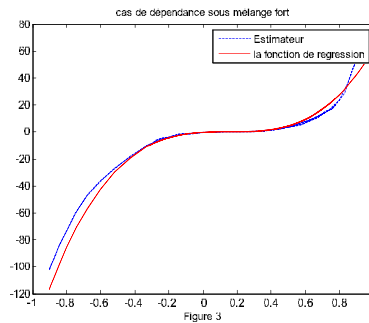
v) simuler la variable à expliquer

$$Y_i = r(x_i) + \varepsilon_i.$$

En considérant  $r(x) = 100(x - 0.15)^3$  et l'estimateur à noyau

$$\hat{r}(x) = \frac{\sum_{i=1}^n Y_i K(h^{-1}d(x, X_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}$$

nous avons obtenu par le logiciel Matlab le graphe suivant :



Voici le programme utilisé :

```
%La regression dans le cas du melange fort
```

```
%a = unifrnd(0,1,1,100);
```

```
%b = unifrnd(0,1,1,100);
```

```
%c = unifrnd(0,1,1,100);
```

```
%for i=1 :1 :100;
```

```
%for j=1 :1 :100;
```

```
%X(j,i)=a(i)*cos(2*pi*(j/100))+ b(i)*sin(4*pi*(j/100))+ 2*c(i)*((j/100)-.25)*((j/100)-.5);
```

```
%end
```

```

%end

%eta=normrnd(0,.05,1,100);

%%%%% Générer alpha-mixing gaussian errors %%%%%%

eps(1)=0;

for i=1 :1 :99;

eps(i+1)=(1/sqrt(2))*(eps(i)+eta(i+1));

end

for i=1 :1 :100

Y(i)=100*(X(i)-.15).^3+eps(i);

end

h=.4;

X0=X(1 :100,100);

%for h=.1 :.1 :.9;

yhat = kern(Y, X, h, X0, 'gau');

r=100*(X0-.15).^3;

MSE=[mean((yhat-r).^2)];

fprintf ('%d.\n',MSE,h);

ytab = sortrows([ MSE h ])

%end

plot(X0,yhat,':')

hold on;

plot(X0,r,'r')

```

legend('Estimateur', 'la fonction de regression' ).

La valeur de la largeur de la fenêtre  $h$  a été choisie par balayage selon le critère  $MCE$ , comme l'indique le tableau suivant :

$h$	0.1	0.2	0.3	<b>0.4</b>	0.5	0.6	<b>0.7</b>	0.8	0.9
$MCE$	6.3910	5.4666	4.4705	<b>4.8462</b>	5.9370	8.2111	14.3888	30.1714	63.7808

La valeur retenu selon le critère  $MCE = \frac{1}{100} \sum_{i=1}^{100} (Y_i - \hat{Y}_i)^2$  vaut  $h = 0.4$

Le choix de l'initialisation de l'algorithme est  $x_0 = x_{100}$ .

On remarque que les courbes obtenues ( Figure 2) sont proches donc les approximations sont bonnes.

## 5.1 Conclusion

Le problème abordé dans ce mémoire se situe autour de la dynamique qui existe actuellement dans la communauté statistique internationale autour de l'estimation non-paramétrique et plus particulièrement autour de la modélisation et du traitement des données et variables fonctionnelles.

On s'est intéressé à l'estimation de la régression, de la fonction de répartition conditionnelle, de la médiane conditionnelle, du mode conditionnelle et des quantiles conditionnelles pour des variables fonctionnelles à partir des régresseurs réels. Plus précisément, on a abordé des aspects de la statistique fonctionnelle en considérant le modèle de régression non-paramétrique pour variables fonctionnelles. Les travaux présentés dans ce mémoire traitent de la modélisation pour variables aléatoires à valeurs dans un espace de dimension infinie (on parle des variables fonctionnelles). Ainsi, la principale source de difficulté, tant d'un point de vue théorique que pratique, provient du fait que les observations de ce type de va-



riables appartiennent à un espace de dimension infinie. Dans ce travail nous avons présenté d'une part des outils mathématiques existants dans la littérature et qui sont nécessaires pour la résolution du problème posé par cette dimension infinie d'autre part la notion du mélange fort qui est utilisée afin d'étendre les résultats de convergence au cadre dépendant dans les deux cas de mélange géométrique et arithmétique. Les vitesses de convergence sont aussi déterminées. Une application pratique est réalisée. Pour cela, nous avons effectué une simulation pour l'estimateur de la régression dans le cadre indépendant et mélangeant. Les résultats de la simulation ont validé les résultats obtenus théoriquement.

# Bibliographie

- [1] [AFKV] Ait Saidi, A., Ferraty, F., Kassa, R. et Vieu, P., Cross-validated estimations in the single-functional index model. *Statistics, Vol. 42, No. 6, December 2008, 475-494.*
- [AP03] Akritas, M. Politis, D. (ed) Recent Advances and trends in nonparametric statistics. Elsevier (2003).
- [B46] Bernstein, S. Probability Theory, 4th ed. (in russian), ed. M.L. Gostechizdat (1946).
- [B75] Bosq, D. Inégalité de Bernstein pour un processus mélangeant (in french). *Compte Rendus Acad. Sci. Paris, Ser. A, 275, 1095-1098 (1975).*
- [B93] Bosq, D. Bernstein's type large deviation inequality for partial sums of strong mixing process. *Statistics, 24, 59-70 (1993).*
- [B00] Bosq, D. Linear Processes in Function Spaces, Theory and Applications. Lecture Notes in statistics, 149, Springer-Verlag New-York (2000).
- [BL87] Bosq, D., Lecoutre, J.P. Théorie de l'estimation fonctionnelle (in french). *Economica (1987).*
- [C83] Carbon, M. Inégalité de Bernstein pour les processus fortement mélangés

non nécessairement stationnaires (in french). *Compte Rendus Acad. Sci.Paris, Ser. A*, 297, 303-306 (1983).

[C84] Collomb, G. Propriétés de convergence presque complète du prédicteur à noyau (in french). *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 441-460 (1984).

[C85] Collomb, G. Nonparametric regression : an up-to-date bibliography. *Statistics*, 2, 309-324 (1985).

[CK95] Cox, D., Kim, T.Y. Moment bounds for mixing random variables useful in nonparametric function estimation. *Stochastic Process Appl.*, 56, (1), 151-158 (1995).

[CFS03] Cardot, H., Ferraty, F., Sarda, P. Spline Estimators for the Functional Linear Model. *Statistica.*, 13, 571-591 (2003).

[CGS04] Cardot, H., Goia, A., Sarda, P. Testing for no effect in functional linear regression models, some computational approaches. *Comm. Statist. Simulation Comput.* 33 179-199 (2004).

[CFF02] Cuevas, A., Febrero, M., Fraiman, R. Linear functional regression : the case of fixed design and functional response. *Canad. J. of Statist.*, 30(2) 285-300 (2002).

[D68] Davydov, Y. Convergence of distributions generated by stationary stochastic processes. *Theory of Probability and its Applications*, 13, 691-696 (1968).

[EQV02] Estèvez-Pérez, G., Quintela-del-Rio, A., Vieu, P. Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples. *J. Statist. Plann. Inference*, 104 (1), 1-30 (2002).

[FN71] Fuk, D.K, Nagaev, S.V. Probability inequalities for sums of independent

- random variables. *Theory Probab. Appl.*, 16, 643-660 (1971).
- [FV05] Ferraty, F., Vieu, P., *Nonparametric Modelling for functional data. Methods, Theory, Applications and Implementations*. Springer, Heidelberg, New York, Hong Kong, London, Milan, Paris, Tokyo (2005).
- [GM84] Gasser, T., Muller, H.G. Kernel estimation of regression function. In : T. Gasser and M. Rosenblatt (Eds) *Smoothing techniques for curve estimation*. Springer Heidelberg, 23-68 (1979).
- [HR47] Hsu, P., Robbins, H. Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA*, 33, 25-31 (1947).
- [HSV98] Hu, T.C., Szynal, D., Volodin, A. I. A note on complete convergence for arrays. *Statistics and Probability Letters* , 38, 27-31 (1998).
- [HV02] Heu, T.C., Volodin, A. Addendum to “A note on complete convergence for arrays”. *Statistics and Probability Letters* , 47, 209-211 (2002).
- [HVZ02] Horova, I., Vieu, P., Zelinka, J. Optimal choice of nonparametric estimates of a density and of its derivatives. *Statistics Decisions* 20 355-378 (2000).
- [I62] Ibragimov., I. Some limit theorems for stationary processes. *Theory of Probability and its Applications* , 7, 349-362 (1962).
- [K94] Kim, T.Y. Moment bounds for non-stationary dependent sequences. *J. Appl. Probab.*, 31 (3), 731-742 (1994).
- [MN89] Maron, J.S. Nolan, D. Canonical Kernels for density estimation. *Statistics and Probability Letters*. 7 195-199 (1989).
- [N97] Nagaev, S. Some refinements of probabilistic and moment inequalities. *Teor :*

- Veroyatnost. i Primenen* (in russian), 42 (4), 832-838 (1997).
- [N98] Nagaev, S. Some refinements of probabilistic and moment inequalities (English translation). *Theory Probab. Appl.*, 42 (4), 707-713 (1998).
- [N64] Nadaraya, E. On estimating regression. *Theory Proba. Appl.* 10 186-196 (1964).
- [RI87] Roussas, G., Ioannid' es, D. Moment inequalities for mixing sequences of random variables. *Stoch. proc. and Appl.*, 5 (1), 61-120 (1987).
- [Ro96] Roussas, G. Exponential probability inequalities with some applications. In : *Statistics, probability and game theory. IMS Lecture Notes Monogr. Ser.*, Inst. Math. Statist., Hayward, CA. 30, 303-319 (1996).
- [R00] Rio, E. Théorie asymptotique des processus aléatoires faiblement dépendants (in french). Springer, *Mathématiques & Applications*, 31, (2000).
- [RS97] Ramsay, J., Silverman, B.W. *Functional Data Analysis*. Springer-Verlag, New-York (1997).
- [SV00] Sarda, P., Vieu, P. Kernel regression, In : M, Schimek (Ed) *Smoothing and regression ; Approches, Computation, and application*. Wiley Series in Probability and statistics, Wiley, New-York, 43-70 (2000).
- [S00] Schimek, M. (Ed.) : *Smoothing and regression ; Approches, Computation, and application*. Wiley Series in Probability and statistics, Wiley, New-York, 43-70 (2000).
- [U37] Uspensky, J. *Introduction to mathematical probability*. Mc Graw-Hill (1937).
- [W64] Watson, G. Smooth regression analysis. *Sankhya Ser. A* 26 359-372 (1964).
- [Y76] Yurinskii, V. Exponential inequalities for sums of random vectors. *J. of Multiv. Analysis*, , 475-499 (1976).

[Y92] Yoshihara, K. Weakly dependent stochastic sequences and their applications.

I : Summation theory for weakly dependent sequences. Sanseido (1992).

[Y94] Yoshihara, K. Weakly dependent stochastic sequences and their applications.

IV : Curve estimation based on weakly dependent data. Sanseido (1994).