République Algérienne Démocratique et Populaire

Ministère de l'enseignement supériure et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2017/2018

Estimation non paramétrique de la fonction de régression : cas des données ergodiques

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité : Analyse Stochastiques, Statistique des Processus et Applications

par

Mekkaoui Assma [1]

Sous la direction de

Dr R. Rouane

Soutnue le 21/06/2018 devant le jury composé de

N. Ait Ouali	Université Dr Tahar Moulay - Saïda	Présidente
R. Rouane	Université Dr Tahar Moulay - Saïda	Encadreur
F. Madani	Université Dr Tahar Moulay - Saïda	Examinateur
F. Benziadi	Université Dr Tahar Moulay - Saïda	Examinateur

1. e-mail : mekkaouiassma94@gmail.com

Remerciements

- ⋆ Avant tous, nous remercions **Allah** le tout puissant qui nous a guidé tout au long de nos vie, qui nous a permis de nous instruire et d'arriver aussi loin dans les études, qui nous a donné courage et patience pour traverser tous les moments difficiles, et qui nous a permis d'achever ce travail.
- ⋆ J'exprime ma reconnaissance et toute ma sympathie à *D.r* **Rouane Rachida** pour son encadrement, sa confiance et sa gentillesse. Je remercie également tous les autres membres de jury qui ont accepté de s'intéresser à mon travail et j'ai apporté leur jugement d'experts.
- ⋆ Un grand merci à tous ceux qui m'ont aidé à élaborer cette recherche.
- ⋆ A tous mes collègues et amies pour leur encouragement et leur aide.

Dédicace

⋆ Je dédie ce mémoire, à mes parents :

∗ Ma très chère mère, qui a oeuvré pour ma réussite, par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, qu'elle reçoit à travers ce travail aussi modeste l'expression de mes sentiments et de mon éternelle gratitude.

∗ Mon chér père qui ma donné la volonté de réaliser et de finir ce travail.

⋆ Merci pour les valeurs nobles, l'éducation et le soutient permanent venu de mes très chères soeurs Zohra, Fatima, Souad, Nacira, Mabrouka, Hanaa, Hadile, et mes frères Mohammed, Rabeh, Ben Ameur, mabrouk, mokhtar, Adbelmalek, Abdallah, Yaakoub, Kaddour, Slimane.

⋆ A tous mes amis : Aimane, Asmaa, Naima, ...

Table des matières

Introduction générale

L'estimation statistique est un domaine très important de la statistique mathématique qui développe des techniques pour décrire certaines caractéristiques d'ensembles d'observations. On pourra distinguer deux composantes principales, à savoir, l'estimation paramétrique et l'estimation non paramétrique. Dans le cadre de l'estimation non-paramétrique, on est intéressé, essentiellement, par l'estimation d'une fonction inconnue appartenant à une certaine classe de fonctions. Notons que l'approche non-paramétrique est basée sur un procédé indépendant de la loi du processus observé. Dans ce contexte général, la procédure d'estimation ne se restreint pas, comme dans l'estimation paramétrique, à l'estimation d'un nombre fini de paramètres liés à la loi de l'échantillon. De ce fait, l'estimation non-paramétrique offre une très grande flexibilité de modélisation pour les applications réelles. Parmi les problèmes importants de la statistique mathématique, l'estimation de caractéristiques fonctionnelles associées à la loi des observations occupe une place prépondérante. À cet effet, de nombreux travaux ont été consacrés à l'estimation de fonctionnelles comme la fonction de régression.

Les contributions à l'étude des paramètres fonctionnels ont beaucoup enrichi la littérature ces dernières décennies. Nous citons à titre d'exemples les travaux de Rosenblatt (1956), Parzen (1962), Nadaraya (1965), Watson (1964), Banon (1978) ainsi que celui de Castellana-Leadbetter (1986).

En modélisation, il est souvent peu réaliste de supposer l'indépendance. C'est pourquoi, depuis plusieurs années déjà, s'est développée une vaste littérature sur les processus dépendants. Les différentes notions de mélange font sans aucun doute partie des conditions de dépendance faible les plus étudiées. Les trois principales classes de mélange sont le mélange fort (ou α-mélange), le β-mélange et le mélange uniforme (ou ϕ-mélange) introduits respectivement par Rosenblatt (1956), Rozanov & Volkonskii (1959) et Ibragimov(1962). Malheureusement, les conditions de mélange sont souvent difficiles à vérifier. On sait même que certains processus bien classiques ne sont pas mélangeants. Citons le célèbre exemple d'Andrews [1] : Si $(\epsilon_i)_{i\in\mathbb{Z}}$ est une suite de variables aléatoires i.i.d., de marginale $\mathcal{B}(1/2)$, alors la solution stationnaire $(X_i)_{i\in\mathbb{Z}}$ de l'équation

$$X_n = \frac{1}{2}(X_{n-1} + \epsilon_n), X_0 \,\text{indépendante de}(\epsilon_i)_{i\geq 1}$$

n'est pas fortement mélangeante. De là est venue l'idée d'introduire des notions de dépendance moins restrictives en considérant des processus satisfaisant des conditions d'ergodicité qui incluent à la fois des processus mélangeants et non-mélangeants, couvrant plus de modèles, pour lesquelles nous avons su développer une théorie asymptotique intéressante. Ces différents coefficients de dépendance vont nous permettre en particulier d'étudier des séries temporelles classiques en économétrie.

Dans ce travail, on se propose d'étudier l'estimateur à noyau de la fonction de régression associés à des processus stationnaires et ergodiques. La motivation essentielle est d'établir des propriétés asymptotiques tout en considérant un cadre de dépendance des données assez général qui puisse être facilement utilisé en pratique.

Ce mémoire est présenté en trois chapitres :

Dans le premier chapitre, on présente notre modèle et on donne une liste des définitions permettant de fixer le vocabulaire utilisé dans la suite de document. Cette liste contient entre autre, les définitions des différents processus de dépandence fort et faible à savoir le α-mélange et l'ergodicité, les définitions des différents modes de convergence,.... On trouvera aussi dans ce chapitre un nombre très important des outils intéressant pour l'élaboration des résultats obtenus.

L'objectif du chapitre 2, est d'étudier le modèle de régression d'une variable réel Y sur une variable fonctionelle X, tous les deux considérés comme stationnaires et ergodiques. En utilisant l'estimateur à noyau de Nadaraya-Watson nous établissons les convergences presque sûres, ponctuelles et uniformes avec des vitesses de convergence de cet estimateur.

Dans le troisième chapitre et dans le même contexte ergodique, nous établissons la normalité asymptotique de l'estimateur considéré dans le chapitre précédent.

Chapitre 1

Présentation

Lorsque l'on souhaite décrire l'influence d'une variable quantitative sur un événement, ou le lien entre une variable explicative X et une variable dite variable réponse Y, on utilise des modèles dits de régression. Ayant observé X, la valeur moyenne de Y est donnée par une fonction de régression : c'est cette fonction qui nous renseigne sur le type de dépendance qu'il y a entre ces deux variables. Cette fonction de régression est définie pour tout $X \in \mathbb{R}$ par :

$$r(x) = \mathbb{E}(Y|X = x) \tag{1.1}$$

qui est la moyenne de la distribution conditionnelle de Y sachant $X = x$.

1.1 Présentation de modele

Soit $(X_i, Y_i)_{i\in\mathbb{N}}$ un processus stationnaire mesurable où Y_i est une variable aléatoire réelle et X_i est une variable aléatoire prend ses valeurs dans un espace semi-métrique $(\mathcal{E}, d(.,.))$. Cela concerne le cas d'espace semi-normé de dimension infinie (par exemple, Hilbert ou espace de Banach) avec la norme $\|.\|$ et la distance $d(x, y) = \|x - y\|$. Suppose que $\mathbb{E}(|Y_1|) < \infty$ et que, pour une partie fixe $x \in \mathcal{E}$, la fonction $r(x) = \mathbb{E}(Y_1|X_1 = x)$ existe.
L'estimateur de type Nadaraya-Watson de r a été introduit par Ferraty et

Vieu (2000). Il est défini par

$$\hat{r}_n(x) = \frac{\displaystyle\sum_{i=1}^{n} Y_i K\left(\frac{d(x,X_i)}{h}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h}\right)} = \frac{\hat{r}_{n,2}(x)}{\hat{r}_{n,1}(x)}, \tag{1.2}$$

oú d est la semi-métrique associée à $\mathcal{E}$, h le paramétre de lissage et K un noyau qui est positif et non décroissant sur son support $[0,1]$ et

$$\hat{r}_{n,j}(x) = \frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n} Y_i^{j-1}\Delta_i(x) \quad \textit{pour} \quad j = 1,2 \tag{1.3}$$

où

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

La méthode du noyau introduite par Rozenblatt (1956)[29] pour estimer la fonction de densité a été reprise par Nadaraya (1965)[23] et Watson (1964)[31] pour estimer la fonction de régression. En raison de la forme de l'estimateur à noyau de la fonction de régression, qui est un rapport, les propriétés de convergence sont obtenues sous des conditions plus restrictives que pour l'estimateur de la densité. Notamment, la convergence uniforme n'est obtenue que sur des ensembles bornés. Cet estimateur a été largement étudié par plusieurs auteurs, nous citons entre autres, Collomb (1985)[4] qui a donné une condition nécessaire et suffisante pour la convergence, uniforme presque sûre, sur un ensemble borné, généralisant ainsi les résultats antérieurs de Nadaraja (1964, 1965)[21], [23], Devroye (1978)[7], Mack et Silverman (1982)[19], Gasser et Müller (1984)[11]. Pour plus de détails sur ce sujet, nous renvoyons le lecteur au livre de Bosq (1998, Chapitre 3)[2].

L'étude de l'estimation de la régression non paramétrique dans le cadre des données ergodiques en temps discret a motivé un certain nombre de travaux. Nous nous référons, entre autres, aux travaux de Delecroix et Rosa (1996), Laib (1999, 2005), Morvai et al. (1996)[5] et Yakowitz et al. (1999)[32] où certaines propriétés asymptotiques sont obtenues. Notons également que la consistance avec vitesse ainsi que la normalité asymptotique ont été obtenues par Laib et Louani (2010, 2011)[24], [16] pour l'estimateur de régression à partir des données fonctionnelles en temps discret.

1.2 Définitions

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ un espace probabilisé filtré.

Définition 1.2.1. *(Processus stationnaire). Un processus $(X_t)_{t\in\mathbb{Z}}$, est dit strictement stationnaire ou stationnaire au sens strict si les lois jointes de $(X_{t_1}, \ldots, X_{t_k})$ et de $(X_{t_{1+h}}, \ldots, X_{t_{k+h}})$ sont identiques pour tout entier positif k et pour tous $t_1, \ldots, t_k, h \in \mathbb{Z}$.*

Définition 1.2.2. *(Ensemble invariant). Soit $(X_t)_{t\in\mathbb{Z}}$, défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, l'ensemble $\mathcal{B} \in \mathcal{F}$ est appelé invariant si il existe un ensemble $\mathcal{A}$, tel que $\mathcal{B} = \{(X_n, X_{n+1}, \ldots) \in \mathcal{A}\}$ est vrai pour tout $n \geq 1$.*

Définition 1.2.3. *(Processus ergodique). Un processus $(X_t)_{t\in\mathbb{Z}}$ est dit ergodique si tout ensemble invariant $\mathcal{B}$ on a $\mathbb{P}(\mathcal{B}) = 1$ ou 0.*

Définition 1.2.4. *(Martingales). Soit M un processus adapté avec $\forall t \geq 0$, $M_t \in L^1$. On dit que M est une $(\mathcal{F}_t)_{t\geq 0}$-martingale si $M_t = \mathbb{E}[M_{t+s}|\mathcal{F}_t]$ pour tout $t, s \geq 0$.*

Définition 1.2.5. *(Différence de martingale). La variable aléatoire $M = (M_t, t \geq 0)$ est une différence de martingale par rapport à la filtration $\mathcal{F} = (\mathcal{F}_t, t \geq 0)$ si*

1- *M_t est $\mathcal{F}_t$-mesurable,*

2- $\mathbb{E}[M_t|\mathcal{F}_{t-s}] = 0, \quad t \geq 0, \qquad s \geq 0.$

1.3 Théorie ergodique

La théorie ergodique est une hypothése originaire de la mécanique. Plus précisément, le mot ergodique est mot grec signifie (travail , énergie). Elle a été proposée par Boltzmann en 1885, afin de modéliser la théorie cinétique des gaz. Depuis cette contribution, les chercheurs en mécanique ont montré que cette hypothése est importante pour modéliser plusieurs phénomenes. Ce qui a motivé les mathématiciens Neumann et Birkhoff en 1931 pour chercher la formulation mathématique de l'ergodicité. Et depuis, la théorie ergodique occupe une place dans des différents branches de la mathématique tel l'analyse fonctionnelle et théorie des groupes, calcule des propabilités et plus précisément processus markoviens, théorie de d'estimation .... Par souci de clarté et de compréhension, nous présentons quelques détails qui définissent la propriété ergodique des processus en temps discret, ainsi qu'une notion intermédiaire entre le mélange et l'ergodicité, appelée le mélange faible. Soit $\{X_n, n \in \mathbb{Z}\}$ une séquence stationnaire, considérons les σ-algèbers $\mathcal{B}_n = \sigma(X_k, k \leq n)$ et $\mathcal{H}_m = \sigma(X_k, k \geq m)$. La suite $\{X_i, i = 1, 2, ...\}$ est dite α-mélange (fort mélange) si

$$\sup_{A \in \mathcal{B}_0^k, B \in \mathcal{H}_n} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| = \alpha(n) \to 0 \text{ quand } n \to \infty.$$

La suite est ergodique si

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mathbb{P}(A \cap \tau^{-k} B) - \mathbb{P}(A)\mathbb{P}(B)| = 0$$

où τ est l'opérateur de transformation. Cette convergence est en fait la convergence au sens de Cesàro de $\mathbb{P}(A \cap \tau^{-k} B) \to \mathbb{P}(A)\mathbb{P}(B)$(see Rosenblatt (1972)). Par conséquent, le mélange fort entrâine l'ergodicité. Cependant, l'inverse n'est pas vrai : il existe des processus ergodiques qui ne sont pas fortement mélangées. La condition d'ergodicité est alors une condition inférieure à tout type de mélange pour lequel les estimateurs non paramétriques usuels (densité, régression, ...) sont convergents.

La question naturelle qui se pose est : quand les moyennes des grandeurs générés de façon stationnaire convergent ?. Dans la situation classique la stationnarité est décrite par une mesure préservant la transformation τ, et l'on considère les moyennes prises sur une séquence $f, f\ o\ \tau, f\ o\ \tau^2, ...$ pour une fonction f intégrable. Cela correspond au concept probabiliste de la stationnarité.

Nous nous sommes intéressés au théorème ergodique, qui s'applique aux processus stationnaires et ergodiques. Le théorème ergodique de Birkhoff, défini pour les processus stationnaires au lieu d'endomorphismes de τ, a maintenant la forme suivante (voir Krengel (1985)[13], p. 26, théorème 4.4.)

Théorème 1.3.1. *[30] (Théorème ergodique). Si $X = (X_t)_{t \in \mathbb{Z}}$ est un proces-*

sus ergodique stationnaire et si X_1 est intégrable, on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mathbb{E}(X_1).$$

1.4 Convergences

Définition 1.4.1. *(Convergence en probabilité). La suite $X_n \to X$ en probabilité ($\mathbb{P}$.),*

$$\forall \epsilon > 0 \qquad \lim_{n\to\infty} \mathbb{P}\{d(X_n, X) \geq \epsilon\} = 0.$$

Définition 1.4.2. *(Convergence presque sûre). La suite (X_n) converge presque sûrement (p.s.) vers X, si*

$$\mathbb{P}\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1.$$

Remarque : La convergence presque sûre entraîne la convergence en probabilité.

Définition 1.4.3. *(Convergence uniforme). Soit I un intervalle de $\mathbb{R}$, (X_n) une suite de fonctions définies sur I, et X définie sur I. On dit que (X_n) converge uniformément vers X sur I si*

$$\forall \epsilon > 0, \quad \exists\, \omega \in I, \forall n \geq n_0, |X_n(\omega) - X(\omega)| < \epsilon.$$

Définition 1.4.4. *Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements on note*

$$\limsup_{n\to\infty} A_n = \cap_{n=0}^{\infty}(\cup_{k=n}^{\infty} A_k),$$
$$\liminf_{n\to\infty} A_n = \cup_{n=0}^{\infty}(\cap_{k=n}^{\infty} A_k).$$

Lemme 1.4.1. *[30] (Lemme de Borel-Cantelli). Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.*

(i) *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n \leq \infty)$, alors*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0,$$

ou de manière équivalente,

$$\textit{presque sûrement,} \qquad \{n \in \mathbb{N}, \omega \in A_n\} \quad \textit{est fini.}$$

(ii) *Si* $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$ *et si les événements* A_n *sont indépendants, alors*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 1,$$

ou de manière équivalente,
presque sûrement, $\{n \in \mathbb{N}, \omega \in A_n\}$ *est infini.*

Lemme 1.4.2. *(L'inégalité - C_r)(Loève, 1963). Soit X une B-variable aléatoire, oú B est espace de banack de type $p(p \geq 1)$, alors on a :*

$$\mathbb{E}\|X - \mathbb{E}X\|^p \leq 2^{p-1}[\mathbb{E}\|X\|^p + \|\mathbb{E}X\|^p].$$

1.5 Inégalités exponentielles

Lemme 1.5.1. *(Peña, V. H. and Giné, E. (1999). [17]). Soit $(Z_n)_{n\geq 1}$ une suite de différences de martingales réelles par rapport à la séquence de σ-algèbre $(\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n))_{n\geq 1}$, où $\sigma(Z_1, \ldots, Z_n)$ est la tribu engendrée par les variables aléatoires $Z_1, \ldots, Z_n$. On pose $S_n = \sum_{i=1}^{n} Z_i$. Pour tout $p \geq 2$ et tout $n \geq 1$, supposons qu'il existe certaines constantes non négatif C et d_n telles que*

$$\mathbb{E}(|Z_n|^p|\mathcal{F}_{n-1}) \leq C^{p-2}p!d_n^2, \qquad \textit{presque sûrement.} \tag{1.4}$$

Alors, pour tout $\epsilon > 0$, on a

$$\mathbb{P}(|S_n| > \epsilon) \leq 2\exp\left\{ -\frac{\epsilon^2}{2(D_n + C\epsilon)}\right\},$$

où $D_n = \sum_{i=1}^{n} d_i^2$.

Le prochain lemme nous donne une inégalité exponentielle similaire pour des sommes partielles de différences de martingales bornées.

Lemme 1.5.2. *(Laïb(1999)[15]). Soit* $\{Z_n : n \in \mathbb{N}\}$ *une suite des différences de martingale par rapport à la filtration* $\{\mathcal{F}_n = \sigma(Z_1, \ldots, Z_n) : i \in \mathbb{N}\}$, *où* $\sigma(Z_1, \ldots, Z_n)$ *est la tribu engendrée par les variables aléatoires* $Z_1, \ldots, Z_n$, *telles que* $\| Z_n \| \leq B$ *p.s. pour* $1 \leq i \leq n$. *Pour tout* $\epsilon > 0$, *on a*

$$\mathbb{P}\bigg\{ \max_{1\leq i\leq n} \bigg\| \sum_{j=1}^{i} Z_j \bigg\| > \epsilon \bigg\} \leq 2 \exp\bigg\{ -\frac{\epsilon^2}{2nB^2} \bigg\}. \tag{1.5}$$

Chapitre 2

Consistance ponctuelle et uniforme presque sûre

Le principal objectif de ce chapitre est d'étudier l'estimateur à noyau de la fonction de régression lorsque le processus est supposé strictement stationnaire et ergodique. Ce chapitre est divisé en deux sections : Dans la première section, on étudie la convergence ponctuelle presque sûre. On traite la convergence uniforme presque sûre de notre estimateur dans la deuxième section.

2.1 Consistance ponctuelle presque sûre

2.1.1 Notations et hypothèses

Soient $\mathcal{F}_i$ est le σ-algèbre engendré par $((X_1, Y_1), ..., (X_i, Y_i))$ et $\mathcal{G}_i$ le σ-algèbre engendré par $((X_1, Y_1), .., (X_i, Y_i), X_{i+1})$. On note $B(x, u)$ la boule centrée en x et de rayon u et $D_i = d(x, X_i)$ de sorte que D_i est une variable aléatoire à valeur réelle non négative. Sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, soit $F_x(u) = \mathbb{P}(D_i \leq u) = \mathbb{P}(X_i \in B(x, u))$ et $F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(D_i \leq u|\mathcal{F}_{i-1}) = \mathbb{P}(X_i \in B(x, u)|\mathcal{F}_{i-1})$ la fonction de distribution et de la fonction de distribution conditionnelle sachant σ-algèbre $\mathcal{F}_{i-1}$ de $(D_i)_{i\geq 1}$, respectivement. A partir de maintenant, pour $j = 1, 2$ et $x \in \mathcal{E}$, on note

$$\bar{r}_{n,j}(x) = \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[Y_i^{j-1}\Delta_i(x)|\mathcal{F}_{i-1}],$$

où $\mathbb{E}(X|\mathcal{F})$ est l'espérance conditionnelle de la variable aléatoire X sachant le σ-algèbre $\mathcal{F}$.

En suite, nous supposons que notre modèle non-paramétrique satisfait les hypothèses suivantes :

(A1) K est un noyau borné non négatif de classe $\mathcal{C}^1$ sur son support $[0, 1]$. La dérivée K' existe sur $[0, 1]$ et satisfait les conditions $K'(t) < 0$, $\forall t \in [0, 1]$ et $|\int_0^1 (K^j)'(u)du| < \infty$ pour $j \geq 1$.

(A2) Pour $x \in \mathcal{E}$, il existe une suite de fonctions aléatoires non négatives $(f_{i,1}(x))_{i\geq 1}$ bornées presque sûrement par une suite de quantités déterministes $(b_i(x))_{i\geq 1}$ en conséquence, une suite de fonctions aléatoires $(\mathrm{g}_{i,x})_{i\geq 1}$, une fonction bornée déterministe non négative f_1 et une fonction réelle non négative ϕ tend vers zéro lorsque son argument est nul, de sorte que

(i) $F_x(u) = \phi(u)f_1(x) + o(\phi(u))$ quand $u \to 0$,

(ii) pour tout $i \in \mathbb{N}, F_x^{\mathcal{F}_{i-1}}(u) = \phi(u)f_{i,1}(x) + \mathrm{g}_{i,x}(u)$ avec $\mathrm{g}_{i,x}(u) = o_{p.s}(\phi(u))$ quand $u \to 0$, $\mathrm{g}_{i,x}(u)/\phi(u)$ bornée presque sûrement et $n^{-1}\sum_{i=1}^{n} \mathrm{g}_{i,x}(u) = o_{p.s}(\phi(u))$ quand $n \to \infty$ et $u \to 0$,

(iii) $n^{-1}\sum_{i=1}^{n} f_{i,1}(x) \to f_1(x)$ presque sûrement quand $n \to \infty$,

(iv) il existe une fonction bornée non décroissante τ_0 telle que, uniformément dans $u \in [0, 1]$
$\frac{\phi(hu)}{\phi(h)} = \tau_0(u) + o(1)$ quand $h \downarrow o$ et $\int_0^1 (k^j(u))'\tau_0(u)du < \infty$,

(v) $n^{-1}\sum_{i=1}^{n} \mathrm{b}_i(x) \to \mathrm{D}(x) < \infty$ quand $n \to \infty$.

(A3) La moyenne conditionnelle de Y_i sachant $\mathcal{G}_{i-1}$ ne dépend que de X_i, c'est-à-dire, pour tout $i \geq 1$, $\mathbb{E}(Y_i|\mathcal{G}_{i-1}) = r(X_i)$.

(A4) (i) Pour $x \in \mathcal{E}$, $|r(x) - r(v)| \leq c_1 d(x,v)^\beta$, pour certains $\beta > 0$ et une constante $c_1 > 0$, v est un voisinage de x.

(ii) Pour tout $m \geq 2$, $\mathbb{E}(|Y_i|^m|\mathcal{G}_i) = \mathbb{E}(|Y_i|^m|X_i)$ et la fonction $\mathrm{g}_m(x) = \mathbb{E}(|Y_1|^m|X_1 = x)$ est continue au point x.

Commentaires sur les hypothèses :
La condition (A1) est très habituelle dans la littérature d'estimation non paramétrique fonctionnelle tandis que l'hypothèse (A2) joue un rôle important dans le contexte fonctionnel et ergodique de ce document. Dans le cas de dimension finie $\mathcal{E} = \mathbb{R}^d$ muni de la norme euclidienne, (voir, [25]), la condition (A2)(ii) apparaît sous la forme suivante

$$\mathbb{P}(X_{i,d} \in B(x,u)|\mathcal{F}_{i-1}) = C_0 f_{i,d} u^d \quad p.s \quad u \to 0, \tag{2.1}$$

où $X_{i,d} = (X_i, \ldots, X_{i-d+1})$ et C_0 est une constante positive. Cela signifie que la probabilité conditionnelle de la boule de dimension d sachant le σ-algèbre $\mathcal{F}_{i-1}$, est gouverné asymptotiquement par une dimension locale lorsque le rayon u tend vers zéro. Cette hypothèse peut être interprétée en termes de dimension fractale qui permet d'obtenir des taux de convergence et la normalité asymptotique pour la fonction de régression du noyau sans supposer l'existence des densités marginales et conditionnelles. C'est d'un intérêt particulier pour les modèles chaotiques où le processus sous-jacent ne possède pas de densité. Notez que l'équation (2.1) est vraie à chaque fois que la distribution conditionnelle $\mathbb{P}_{X_{i,d}}^{\mathcal{F}_{i-1}}$ a une densité conditionnelle continue $f_{X_{i,d}}^{\mathcal{F}_{i-1}}(x) = f_{i,d}(x)$ à tout moment de l'ensemble $\{x : f_{X_{i,d}}^{\mathcal{F}_{i-1}}(x) > 0\}$ et la constante C_0 prend la valeur $\pi^{d/2}/\Gamma((d+2)/2)$, où Γ est la fonction gamma.
Notez également que prendre $\mathrm{g}_{i,x}(u) = C_i\zeta(u)$ avec C_i est une variable aléatoire et ζ une fonction telle que $\zeta(u)/\phi(u) \to 0$ quand $u \to 0$, la condition (A2)(ii) est clairement satisfaite en utilisant le théorème ergodique. Dans les exemples 1 et 2 ci-dessous, nous considérons le processus de d-vecteur et les processus fonctionnels associés aux modèles autorégressifs d'ordre 1 et découvrons des formes explicites de chaque élément apparaissant dans la condition (A2)(ii). Les conditions (A3) et (A4)(ii) sont de type Markovien. Ils sont satisfaits quand on considère, par exemple, le modèle de régression hétéroscédastique habituel $Y_i = r(X_i) + \sigma(X_i)\epsilon_i$ où les variables aléatoires centrées ϵ_i est indépendantes de X_i, tandis que (A4)(i) est une condition de

lissage faible sur la fonction de régression r.

Exemple1. Soit $\mathcal{C}$ un espace séparé muni d'une semi-distance. Considérons le modèle autorégressif d'ordre 1 défini, pour tout $i \geq 1$, par

$$X_i = \rho(X_{i-1}) + \epsilon_i,$$

où $\epsilon_i = \eta_i h$ est une variable aléatoire réelle avec η_i indépendant de X_{i-1} et $h \in \mathcal{C}$ et ρ est un opérateur fonctionnel sur $\mathcal{C}$.

Pour $(x, y) \in \mathcal{C}^2$, considérer la semi-distance entre x et y donné par

$$d(x, y) = \left| \int (x(t) - y(t))dt \right|.$$

Observer pour tout $u > 0$, que nous avons

$$F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d(x, X_i) \leq u|\mathcal{F}_{i-1}) = \mathbb{P}(d(x, X_i) \leq u|X_{i-1}).$$

Par conséquent, pour $0 \neq \int h(t)dt < \infty$, nous avons

$$\begin{aligned}
F_{X_i|X_{i-1}=s}(u) &= \mathbb{P}(d(x, X_i) \leq u|X_{i-1} = s) \\
&= \mathbb{P}\left(\left|\int x(t) - X_i(t)dt\right| \leq u|X_{i-1} = s\right) \\
&= \mathbb{P}\left(\left|\int x(t) - \rho(X_{i-1})(t) - \eta_i h(t)dt\right| \leq u|X_{i-1} = s\right) \\
&= \mathbb{P}\left(\left|\int x(t) - \rho(s)(t) - \eta_i h(t)dt\right| \leq u\right) \\
&= \mathbb{P}\left(\frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt} \leq \eta_i \leq \frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right) \\
&= \Phi\left(\frac{u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right) - \Phi\left(\frac{-u + \int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right),
\end{aligned}$$

où Φ est la fonction de répartition de η_i. En supposant que $0 < \int h(t)dt < \infty$, $|\int x(t)dt| < \infty$ et $|\int \rho(s)(t)dt| < \infty$ pour tout $s \in \mathcal{C}$ et en prenant Φ comme la fonction de répartition de la loi $\mathcal{N}(0, 1)$, nous obtenons

$$F_{X_i|X_{i-1}=s}(u) = \frac{u}{\int h(t)dt}\sqrt{\frac{2}{\Pi}} \exp\left(-\frac{1}{2}\left(\frac{\int x(t)dt - \int \rho(s)(t)dt}{\int h(t)dt}\right)^2\right)(1+o(1)).$$

Ainsi,

$$F_x^{\mathcal{F}_{i-1}} = \frac{u}{\int h(t)dt}\sqrt{\frac{2}{\Pi}} \exp\left(-\frac{1}{2}\left(\frac{\int x(t)dt - \int \rho(X_{i-1})(t)dt}{\int h(t)dt}\right)^2\right)(1 + o(1)),$$

et la condition (A2)(ii) est satisfaite avec

$$\phi(u) = \frac{u}{\int h(t)dt}\sqrt{\frac{2}{\Pi}}.$$

Exemple2. Considérons le modèle AR(1) défini par

$$X_i = AX_{i-1} + \epsilon_i, \quad i \geq 1, \tag{2.2}$$

où $X_i \in \mathbb{R}^d$ et $\epsilon_i \in \mathbb{R}^d$. Supposons que les composants de ϵ_i sont indépendantes et que ϵ_i est indépendante de la σ-algèbre $\mathcal{F}_{i-1}$.
En outre, supposons que A est une matrice diagonale, i.e.,

$$A=\begin{pmatrix} \theta_1 & & & & \\ & . & & & \\ & & . & & \\ & & & . & \\ & & & & \theta_d \end{pmatrix}.$$

Notez que le modèle (2.2) est markovien d'ordre 1 et que le processus $(X_i)_{i\geq 1}$ est stationnaire si $\max\limits_{1\leq i\leq d} |\theta_i| < 1$.
Observer, pour tout $x \in \mathcal{E} = \mathbb{R}^d$ et tout $i \geq 1$, que nous avons,

$$F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(X_i \in B(x,u)|\mathcal{F}_{i-1}) = \mathbb{P}(d(X_i,x) \leq u|\mathcal{F}_{i-1}) \quad u > 0,$$

où $B(x,u)$ est une boule de centre x et de rayon u.
En utilisant la propriété du processus de Markov $(X_i)_{i\geq 1}$, on peut voir que

$$F_x^{\mathcal{F}_{i-1}}(u) = \mathbb{P}(d(X_i,x) \leq u|X_{i-1}).$$

Considérons maintenant la fonction suivante

$$F_{X_i|X_{i-1}=t}(u) = \mathbb{P}(d(X_i,x) \leq u|X_{i-1} = t).$$

De toute évidence, en prenant d comme la norme sup, nous avons

$$\begin{aligned} F_{X_i|X_{i-1}=t}(u) &= \mathbb{P}(x_j - u \leq \theta_j X_{i-1,j} + \epsilon_{i,j} \leq x_j + u, \quad 1 \leq j \leq d|X_{i-1} = t) \\ &= \mathbb{P}(x_j - u - \theta_j t_j \leq \epsilon_{i,j} \leq x_j + u - \theta_j t_j, \quad 1 \leq j \leq d) \\ &= \prod_{j=1}^{d} \mathbb{P}(x_j - u - \theta_j t_j \leq \epsilon_{i,j} \leq x_j + u - \theta_j t_j) \\ &= \prod_{j=1}^{d} \left[G(x_j + u - \theta_j t_j) - G(x_j - u - \theta_j t_j)\right], \end{aligned}$$

où G est la fonction de distribution cumulative de $\epsilon_{i,j}$.
Comme $F_x^{\mathcal{F}_{i-1}}(u) = F_{X_i|X_{i-1}}(u)$, on a

$$F_x^{\mathcal{F}_{i-1}}(u) = \prod_{j=1}^{d} \left[G(x_j + u - \theta_j X_{i-1,j}) - G(x_j - u - \theta_j X_{i-1,j})\right].$$

Supposons que la fonction de distribution est deux fois dérivable, et que la deuxième dérivée est bornée. De toute évidence, cette hypothèse est remplie dans le cas gaussien. Observez également, dans le voisinage de $u = 0$, que

$$\begin{aligned} G(x_j + u - \theta_j X_{i-1,j}) - G(x_j - u - \theta_j X_{i-1,j}) &= 2uG'(x_j - \theta_j X_{i-1,j}) + O(u^2) \\ &= 2u(G'(x_j - \theta_j X_{i-1,j}) + O(u)). \end{aligned}$$

Par conséquent, nous avons

$$\begin{aligned} F_x^{\mathcal{F}_{i-1}}(u) &= 2^d u^d \prod_{j=1}^{d} (G'(x_j - \theta_j X_{i-1,j}) + O(u)) \\ &= 2^d u^d \prod_{j=1}^{d} \mathrm{g}_j(x_j - \theta_j X_{i-1,j}) + O(u^{d+1}) \\ &= (2u)^d \mathrm{g}(x - AX_{i-1}) + O(u^{d+1}), \end{aligned}$$

où g_j est la densité de $\epsilon_{i,j}$ et g est la densité joint de ϵ.
Supposons maintenant que $\mathcal{E}$ est un espace de Hilbert séparable équipé du produit scalaire $<,>$ et soit $(e_j)_{j\geq 1}$ une base orthonormale de $\mathcal{E}$. Soit $x = \sum_{j=1}^{\infty} x_j e_j$ un élément de $\mathcal{E}$. Pour $k \in \mathbb{N}^*$, considérer la semi-métrique d_k défini, pour tout $(\mathbf{x}, \mathbf{y}) \in \mathcal{E}^2$, par

$$d_k(\mathbf{x}, \mathbf{y}) = \left(\sum_{j=1}^{k} < \mathbf{x} - \mathbf{y}, e_j >^2 \right)^{1/2}.$$

Nous nous référons à [9], le Lemme 13.6 page 213, pour la preuve que d_k est en fait une semi-métrique.
Supposons aussi que $X_i = \sum_{i=1}^{\infty} X_i^j e_j$ est un élément aléatoire au carré de $\mathcal{E}$ et mettre $\mathbf{X}_i = (X_i^1, \dots, X_i^k)$ et $\mathbf{x} = (x_1, \dots, x_k)$. De plus, supposons que la fonction de densité conditionnelle $f_{\mathbf{X}_i}^{\mathcal{F}_{i-1}}(\mathbf{x})$ est une fonction aléatoire avec des trajectoires continues presque sûres dans le voisinage de $\mathbf{x}$ et $f_{\mathbf{X}_i}^{\mathcal{F}_{i-1}}(\mathbf{x}) > 0$.

Notez que la décomposition $X_i = \sum_{i=1}^{\infty} X_i^j e_j$ peut résulter de l'expansion de Karhunen-Loève comme un processus gaussien centré et que $\mathbf{X}_i$ est un vecteur aléatoire $\in \mathbb{R}^k$ indépendant de la loi $N(0,1)$. Il s'ensuit alors, pour $u > 0$, que

$$\begin{aligned}
F_x^{\mathcal{F}_{i-1}}(u) &= \mathbb{P}(d_k(x, X_i) \leq u|\mathcal{F}_{i-1}) \\
&= \mathbb{P}\left(\sum_{j=1}^{k} < x_j - X_i^j, e_j >^2 \leq u^2|\mathcal{F}_{i-1}\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{k} (x_j - X_i^j)^2 \leq u^2|\mathcal{F}_{i-1}\right) \\
&= \mathbb{P}(||\mathbf{X}_i - \mathbf{x}||_{Eucl} \leq u|\mathcal{F}_{i-1}),
\end{aligned}$$

où $\|.\|_{Eucl}$ est la norme euclidienne sur $\mathbb{R}^k$. De toute évidence, à partir de la continuité presque sur des trajectoires de densité conditionnelle, nous avons

$$\begin{aligned}
F_x^{\mathcal{F}_{i-1}}(u) &= \int_{B(\mathbf{x},u)} f_{\mathbf{X}_i}^{\mathcal{F}_{i-1}}(t)dt \\
&= \frac{2\Pi^{\frac{k}{2}}}{k\Gamma(\frac{k}{2})} u^k f_{\mathbf{X}_i}^{\mathcal{F}_{i-1}}(\mathbf{x}) + o_{p.s}(u^k).
\end{aligned}$$

Alors que $\mathbf{x} = (x_1, ..., x_k)$ est la projection de $x \in \mathcal{E}$ sur l'espace $\mathbb{R}^k$, il suffit de prendre $f_{i,1}(x) = f_{\mathbf{X}_i}^{\mathcal{F}_{i-1}}(\mathbf{x})$ pour satisfaire l'hypothèse (A2)(ii).

2.1.2 Résultats ponctuels

Théorème 2.1.1. *[16] On suppose que les hypothèses (A1)-(A4) sont vérifiées et que*

$$n\phi(h) \to \infty \quad et \quad \frac{\log n}{n\phi(h)} \to 0 \quad quand \quad n \to \infty. \tag{2.3}$$

Alors, nous avons

$$\hat{r}_n(x) - r(x) = O_{p.s}(h^\beta) + O_{p.s}\left(\sqrt{\frac{\log n}{n\phi(h)}}\right),$$

où β est donnée dans l'hypothèse (A4)(i).

Remarque 2.1.1. *Le taux de convergence est lié, dans cette deuxième composante, aux probabilité de petite boule données par la fonction ϕ qui est étroitement liée à la semi-métrique considérée d.*

Preuve :
Afin d'établir nos résultats, nous utilisons la décomposition ci-après. Pour $x \in \mathcal{E}$,

$$Q_n(x) = (\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)) - r(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)) \tag{2.4}$$

et

$$R_n(x) = -B_n(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)), \tag{2.5}$$

où

$$B_n(x) = \frac{\bar{r}_{n,2}(x)}{\bar{r}_{n,1}(x)} - r(x)$$

est le biais conditionnel de l'estimation de régression $\hat{r}_n(x)$. De toute évidence, nous avons

$$\hat{r}_n(x) - r(x) = B_n(x) + \frac{R_n(x)}{\hat{r}_{n,1}(x)} + \frac{Q_n(x)}{\hat{r}_{n,1}(x)}. \tag{2.6}$$

L'intérêt majeur de la décomposition (2.6) vient du fait que les résumés du terme $Q_n(x)$ forment une différence de martingale qui permet d'établir la convergence presque sûrement avec un taux. La preuve de Théorème 2.1.1 est divisée en plusieurs lemmes établissant la convergence de $\hat{r}_{n,1}(x)$ à 1 et la convergence des termes $B_n(x)$, $R_n(x)$ et $Q_n(x)$.

Lemme 2.1.1. *Supposons que les hypothèses (A1) et (A2)(i), (A2)(ii) et (A2)(iv) vraies. Pour tout nombre rèel $1 \leq j \leq 2+\delta$ et $1 \leq k \leq 2+\delta$ avec $\delta > 0$, quand $n \to \infty$, nous avons*

(i) $\dfrac{1}{\phi(h)}\mathbb{E}[\Delta_i^j(x)|\mathcal{F}_{i-1}] = M_j f_{i,1}(x) + O_{p.s}\left(\dfrac{\mathrm{g}_{i,x}(h)}{\phi(h)}\right)$,

(ii) $\dfrac{1}{\phi(h)}\mathbb{E}[\Delta_1^j(x)] = M_j f_1(x) + o(1)$

et

(iii) $\dfrac{1}{\phi^k(h)}(\mathbb{E}(\Delta_1(x)))^k = M_1^k f_1^k(x) + o(1)$.

où M_j est défini par

$$M_j = K^j(1) - \int_0^1 (K^j)'(u)\tau_0(u)du. \tag{2.7}$$

Remarque 2.1.2. *Dans les conditions (A1), (A2)(iv) et (B1), les résultats du Lemme 2.1.1 se vérifient uniformément en $x \in \mathcal{C}$.*

Remarque 2.1.3. *Les déclarations du Lemme 2.1.1 restent valables pour tous nombre rèel $j \geq 1$ et $k \geq 1$, a condition que l'hypothèse (A2)(iv) est satisfaite avec $j \geq 1$ au lieu de $1 \leq j \leq 2+\delta$.*

Preuve du Lemme 2.1.1. Observez, pour $j \geq 1$, que nous pouvons écrire

$$\begin{aligned}\mathbb{E}[\Delta_i^j(x)|\mathcal{F}_{i-1}] &= \int_0^h K^j\left(\frac{u}{h}\right) d\mathbb{P}^{\mathcal{F}_{i-1}}(d(x,X_1)\leq u)\\ &= \int_0^1 K^j(t) d\mathbb{P}^{\mathcal{F}_{i-1}}\left(\frac{d(x,X_1)}{h}\leq t\right).\end{aligned} \tag{2.8}$$

Comme la fonction K^j est de classe C^1, nous avons évidemment $K^j(t) = K^j(0) + \displaystyle\int_0^t (K^j(u))'du$. De plus, on $\mathbb{P}(d(x,X)\leq 0) = 0$ pour tout $x \in \mathcal{E}$, en combinant ce résultat avec (2.8) et en appliquant le théorème de Fubini, on obtient

$$\begin{aligned}\mathbb{E}[\Delta_i^j(x)|\mathcal{F}_{i-1}] &= K^j(0)F_x^{\mathcal{F}_{i-1}}(h) + \int_0^1\left[\int_0^t (K^j(u))'du\right] d\mathbb{P}^{\mathcal{F}_{i-1}}\left(\frac{d(x,X_1)}{h}\leq t\right)\\ &= K^j(0)F_x^{\mathcal{F}_{i-1}}(h) + \int_0^1\int_0^1 (K^j(u))'1_{[u,1]}(t) d\mathbb{P}^{\mathcal{F}_{i-1}}\left(\frac{d(x,X_1)}{h}\leq t\right)du\\ &= K^j(0)F_x^{\mathcal{F}_{i-1}}(h) + F_x^{\mathcal{F}_{i-1}}(h)[K^j(1)-K^j(0)] - \int_0^1 (K^j(u))'F_x^{\mathcal{F}_{i-1}}(uh)du\\ &= K^j(1)F_x^{\mathcal{F}_{i-1}}(h) - \int_0^1 (K^j(u))'F_x^{\mathcal{F}_{i-1}}(uh)du.\end{aligned}$$

Ainsi, d'après les hypothèses (A2)(ii) et (A2)(iv), on peut écrire

$$\begin{aligned}\mathbb{E}[\Delta_i^j(x)|\mathcal{F}_{i-1}] &= K^j(1)(f_{i,1}(x)\phi(h)+\mathrm{g}_{i,x}(h)) - \int_0^1 (K^j)'(u)[\phi(uh)f_{i,1}(x)+\mathrm{g}_{i,x}(uh)]du,\\ &= \phi(h)\Bigg[K^j(1)f_{i,1}(x) + \frac{\mathrm{g}_{i,x}(h)}{\phi(h)} - \int_0^1 (K^j)'(u)(\tau_0(u)+o(1))\\ &\quad \left[f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(uh)}{\phi(uh)}\right]du\Bigg].\end{aligned}$$

Comme $K^j(1)\frac{\mathrm{g}_{i,x}}{\phi}$ est une fonction bornée presque sûrement, en utilisant le théorème de convergence dominée, on obtient

$$\mathbb{E}[\Delta_i^j(x)|\mathcal{F}_{i-1}] = \phi(h)f_{i,1}(x)\left[K^j(1) - \int_0^1 (K^j)'(u)\tau_0(u)du\right] + O_{p.s}(\mathrm{g}_{i,x}(h)).$$

La preuve de la première partie du Lemme 2.1.1 est alors réalisée. La partie (ii) découle de la partie (i) avec $\mathcal{F}_i$ considérée comme le σ-algèbre trivial. La preuve de la partie (iii) suit en prenant $j = 1$ dans la partie (ii) et en considérant l'exposant K dans les deux côtés de l'égalité. ■

Les deux lemmes suivants décrivent le comportement asymptotique du terme $\hat{r}_{n,1}(x)$, $B_n(x)$ et $R_n(x)$.

Lemme 2.1.2. *Supposons que les hypothèses (A1)-(A2) et la condition (2.3) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$, on a :*

$$\lim_{n\to\infty} \hat{r}_{n,1}(x) = 1 \qquad \textit{presque sûrement.}$$

Preuve du Lemme 2.1.2. On a

$$\hat{r}_{1,n}(x) - 1 = R_{n,1}(x) + R_{n,2}(x), \tag{2.9}$$

où

$$\begin{aligned}
R_{n,1}(x) &= \frac{1}{n\mathbb{E}[\Delta_1(x)]}\sum_{i=1}^n(\Delta_i(x) - \mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}]) = \hat{r}_{1,n}(x) - \bar{r}_{1,n}(x),\\
R_{n,2}(x) &= \frac{1}{n\mathbb{E}[\Delta_1(x)]}\sum_{i=1}^n(\mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}] - \mathbb{E}\Delta_1(x)).
\end{aligned} \tag{2.10}$$

Combinaison le Lemme 2.1.1 avec l'hypothèse (A2)(iii), on voit facilement que $R_{n,2}(x) = o_{p.s}(1)$ quand $n \to \infty$.
Pour gérer le premier terme, puisque le noyau K est borné, observez que $R_{n,1}(x) = 1/n\mathbb{E}[\Delta_1(x)]\sum_{i=1}^n L_{n,i}(x)$ où $\{L_{n,i}(x)\}$ est un tableau triangulaire des différences de martingale bornées par rapport à la suite σ-algèbre $(\mathcal{F}_{i-1})_{i\geq 1}$. En utilisant l'inégalité de Jensen, il s'ensuit que

$$\mathbb{E}[L_{n,i}^2(x)|\mathcal{F}_{i-1}] \leq 2\mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}].$$

Maintenant, de vue dans l'hypothèse (A2), qui suppose que $f_{x,i}$ est bornée presque sûrement par la quantité déterministe $b_i(x)$ et que $\mathrm{g}_{i,x}(h) \leq \phi(h)$ pour n assez grand, et le Lemme 2.1.1(i), nous avons

$$\mathbb{E}[L_{n,i}^2(x)|\mathcal{F}_{i-1}] \leq 2\phi(h)[M_2 b_i(x)+1] = d_i^2 \quad p.s. \tag{2.11}$$

Par conséquent, en utilisant le Lemme 1.5.1, on obtient :

$$\begin{aligned}
\mathbb{P}(|R_{n,1}(x)| > \epsilon) &= \mathbb{P}\left(\left|\sum_{i=1}^{n} L_{n,i}(x)\right| > \epsilon n \mathbb{E}[\Delta_1(x)]\right) \\
&\leq 2\exp\left\{-\frac{(n\epsilon\mathbb{E}(\Delta_1(x)))^2}{2(D_n + C\epsilon n\mathbb{E}[\Delta_1(x)])}\right\} \\
&\leq 2\exp\left\{-n\epsilon^2 \frac{\mathbb{E}(\Delta_1(x))^2}{2D_n/n} \frac{1}{\left[1+\frac{C\epsilon\mathbb{E}[\Delta_1(x)]}{D_n/n}\right]}\right\}.
\end{aligned}$$

Du Lemme 2.1.1 et l'hypothèse (A2)(v), on a :

$$\begin{aligned}
\frac{\mathbb{E}[\Delta_1(x)]}{D_n/n} &= \frac{M_1 f_1(x)\phi(h) + o(\phi(h))}{2\phi(h)[M_2 D(x) + 1 + o(1)]} \\
&= \frac{M_1 f_1(x)}{2(M_2 D(x)+1)} + o(1)
\end{aligned}$$

et

$$\begin{aligned}
\frac{(\mathbb{E}[\Delta_1(x)])^2}{2D_n/n} &= \frac{M_1^2 f_1^2(x)\phi^2(h) + o(\phi^2(h))}{4\phi(h)[M_2 D(x)+1+o(1)]} \\
&= \phi(h)\left[\frac{M_1^2 f_1^2(x)}{4[M_2 D(x)+1]} + o(1)\right].
\end{aligned}$$

Par conséquent, nous avons

$$\mathbb{P}(|R_{n,1}(x)| > \epsilon) = 2\exp\left\{-\epsilon^2 n\phi(h)\left(\frac{M_1^2 f_1^2(x)}{4[M_2 D(x)+1]} + o(1)\right) \times \frac{1}{1+\left(\frac{\epsilon C M_1 f_1(x)}{2(M_2 D(x)+1)} + o(1)\right)}\right\}.$$

En choisissant

$$\epsilon = \epsilon_n = \left(\frac{4[M_2 D(x)+1]\log(n)}{M_1^2 f_1^2(x) n\phi(h)}\right)^{1/2} \epsilon_0 \to 0 \quad \text{quand } n \to \infty,$$

avec $\epsilon_0 > 0$, pour n suffisamment grand, nous obtenons :

$$\begin{aligned}
\mathbb{P}(|R_{n,1}(x)| > \epsilon) &= 2\exp\left\{ -\epsilon_0^2 \log n \times \frac{1}{1+\epsilon_0 \left(\frac{\log n}{n\phi(h)}\right)^{1/2} \frac{C}{[M_2 D(x)+1]^{1/2}} + o(1)} \right\} \\
&\leq 2\exp\{-\log n^{\epsilon_0^2}\} \\
&\leq \frac{1}{n^{\epsilon_0^2}}.
\end{aligned}$$

Par conséquent, en prenant ϵ_0 suffisamment grande et en considérant la condition (2.3), par le lemme de Borel-Cantelli on obtient

$$R_{n,1}(x) = O_{p.s}\left(\left(\frac{\log n}{n\phi(h)}\right)^{1/2}\right) = o_{p.s}(1). \qquad \blacksquare \tag{2.12}$$

Lemme 2.1.3. *Supposons que les hypothèses (A1)-(A2), (A3), (A4)(i) et la condition (2.3) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$, on a :*

$$B_n(x) = O_{p.s}(h^\beta) \tag{2.13}$$

et

$$R_n(x) = O_{p.s}\left(h^\beta \left(\frac{\log n}{n\phi(h)}\right)^{1/2}\right). \tag{2.14}$$

Preuve du Lemme 2.1.3. D'abord, nous évaluons le terme de biais conditionnel. On a

$$B_n(x) = \frac{\bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x)}{\bar{r}_{n,1}(x)}.$$

De même que dans le Lemme 2.1.2, on voit facilement que $\bar{r}_{n,1}(x) = O_{p.s}(1)$. Par conséquent, nous devons établir que

$$\tilde{B}_n(x) = \bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x) = O_{p.s}(h^\beta) \tag{2.15}$$

En utilisant les conditions (A3) et (A4)(i), on peut facilement voir que

$$\begin{aligned}
|\tilde{B}_n(x)| &= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}\mathbb{E}[(Y_i - r(x))\Delta_i(x)|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}\mathbb{E}[\mathbb{E}[(Y_i - r(x))\Delta_i(x)|\mathcal{G}_{i-1}]|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}\mathbb{E}[\mathbb{E}[(Y_i - r(x))\Delta_i(x)|X_i]|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}\mathbb{E}[(r(X_i) - r(x))\Delta_i(x)|\mathcal{F}_{i-1}]\right| \\
&\leq \sup_{u\in B(x,h)}|r(u) - r(x)|\left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}\mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}]\right| = O_{p.s}(h^{\beta}),
\end{aligned}$$

La deuxième partie du Lemme 2.1.3 suit facilement le fait que

$$R_n(x) = -B_n(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)),$$

l'équation (2.15) et l'observation, tirée des équations (2.10) et (2.12), permettent de conclure :

$$\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x) = O_{p.s}\left((\frac{\log n}{n\phi(h)})^{1/2}\right).$$

■

Lemme 2.1.4. *Supposons que les hypothèses (A1)-(A2), (A3), (A4) et la condition (2.3) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$, on a :*

$$\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x) = O_{p.s}\left(\left(\frac{\log n}{n\phi(h)}\right)^{1/2}\right). \tag{2.16}$$

Remarque 2.1.4. *En utilisant l'équation (2.4) et le Lemme 2.1.4, nous avons*

$$Q_n(x) = O_{p.s}\left(\left(\frac{\log n}{n\phi(h)}\right)^{1/2}\right).$$

Preuve du Lemme 2.1.4. Observez que

$$\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x) = \frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^{n}L_{i,n}(x),$$

où $L_{i,n}(x) = Y_i\Delta_i(x) - \mathbb{E}(Y_i\Delta_i(x)|\mathcal{F}_{i-1})$ est une différence de martingale. Par conséquent, nous pouvons utiliser le Lemme 1.5.1 pour obtenir une borne supérieure exponentielle de la quantité $\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)$. Pour tout $p \in \mathbb{N} - \{0\}$, on a

$$L_{i,n}^p(x) = \sum_{k=0}^{p} C_p^k (Y_i\Delta_i(x))^k (-1)^{p-k} [\mathbb{E}(Y_i\Delta_i(x)|\mathcal{F}_{i-1}]^{p-k}.$$

En vue de l'hypothèse (A3), $[\mathbb{E}(Y_i\Delta_i(x)|\mathcal{F}_{i-1}]^{p-k}$ est $\mathcal{F}_{i-1}$ mesurable, il s'ensuit alors que

$$\mathbb{E}(L_{i,n}^p(x)|\mathcal{F}_{i-1}) = \sum_{k=0}^{p} C_p^k \mathbb{E}[(Y_i\Delta_i(x))^k|\mathcal{F}_{i-1}](-1)^{p-k}[\mathbb{E}(Y_i\Delta_i(x)|\mathcal{F}_{i-1})]^{p-k}.$$

Ainsi,

$$\mathbb{E}(L_{i,n}^p(x)|\mathcal{F}_{i-1}) \leq \sum_{k=0}^{p} C_p^k \mathbb{E}[|Y_i\Delta_i(x)|^k|\mathcal{F}_{i-1}][\mathbb{E}(|Y_i\Delta_i(x)||\mathcal{F}_{i-1}]^{p-k}.$$

En utilisant l'inégalité de Jensen, on peut écrire

$$\mathbb{E}[|Y_i\Delta_i(x)|^k|\mathcal{F}_{i-1}][\mathbb{E}(|Y_i\Delta_i(x)||\mathcal{F}_{i-1}]^{p-k} \leq \mathbb{E}(|Y_i\Delta_i(x)|^k|\mathcal{F}_{i-1})\mathbb{E}(|Y_i\Delta_i(x)|^{p-k}|\mathcal{F}_{i-1}).$$

Observez maintenant que pour tout $m \geq 1$

$$\begin{aligned}
\mathbb{E}(|Y_i\Delta_i(x)|^m|\mathcal{F}_{i-1}) &= \mathbb{E}[\mathbb{E}(|Y_i|^m\Delta_i^m(x)|\mathcal{G}_i)|\mathcal{F}_{i-1}] \\
&= \mathbb{E}[\Delta_i^m(x)\mathrm{g}_m(X_i)|\mathcal{F}_{i-1}].
\end{aligned}$$

En vue de l'hypothèse (A4)(ii), nous avons

$$\begin{aligned}
\mathbb{E}(|Y_i\Delta_i(x)|^m|\mathcal{F}_{i-1}) &\leq \mathbb{E}[\Delta_i^m|\mathrm{g}_m(X_i) - \mathrm{g}_m(x)||\mathcal{F}_{i-1}] + \mathrm{g}_m(x)\mathbb{E}[\Delta_i^m(x)|\mathcal{F}_{i-1}] \\
&\leq \mathbb{E}[\Delta_i^m(x)|\mathcal{F}_{i-1}]\left[\sup_{u\in B(x,h)}|\mathrm{g}_m(u) - \mathrm{g}_m(x)| + \mathrm{g}_m(x)\right] \\
&\leq C_0\mathbb{E}[\Delta_i^m(x)|\mathcal{F}_{i-1}],
\end{aligned}$$

où C_0 est une constante positive. Remarque par le Lemme 2.1.1, les hypothèses (A2)(ii) et (A2)(iii) (avec m = k), chaque fois que le noyau K et la fonction τ_0 sont bornés par des constantes positives a_1 et c_0 respectivement, que nous avons :

$$\begin{aligned}
\mathbb{E}[\Delta_i^m(x)|\mathcal{F}_{i-1}] &= \phi(h)M_k f_{i,1}(x) + O_{p.s}(\mathrm{g}_{i,x}(h)) \\
&= c_1\phi(h)a_1^k f_{i,1}(x) + O_{p.s}(\mathrm{g}_{i,x}(h)).
\end{aligned}$$

De même, avec $m = p - k$, on obtient

$$\mathbb{E}(|Y_i\Delta_i(x)|^{p-k}|\mathcal{F}_{i-1}) = c_1\phi(h)a_1^{p-k}f_{i,1}(x) + O_{p.s}(\mathrm{g}_{i,x}(h)).$$

Donc,

$$\begin{aligned}\mathbb{E}[|Y_i\Delta_i(x)|^k|\mathcal{F}_{i-1}][\mathbb{E}(|Y_i\Delta_i(x)|^{p-k}|\mathcal{F}_{i-1})] &= c_1^2a_1^p\phi(h)^2f_{i,1}^2(x) + O_{p.s}(\mathrm{g}_{i,x}(h))\phi(h)f_{i,1}(x)\\ &\quad (a_1^k + a_1^{p-k}) + O_{p.s}(\mathrm{g}_{i,x}^2(h)).\end{aligned}$$

Comme dans l'équation (2.11), étant donnée que $f_{i,1}(x)$ est bornée presque sûrement par une quantité déterministe $b_i(x)$, $\mathrm{g}_{i,x}(h) \leq \phi(h)$ presque sûrement et $\phi(h)^2 < \phi(h)$, pour n suffisamment grand, alors presque sûrement on peut écrire

$$\begin{aligned}|\mathbb{E}(L_{i,n}^p(x)|\mathcal{F}_{i-1})| &= c_1^2a_1^p2^p\phi(h)^2f_{i,1}^2(x) + 2^pO_{p.s}(\mathrm{g}_{i,x}^2(h)) + 2(1+a_1)^pO_{p.s}(\mathrm{g}_{i,x}(h))\phi(h)f_{i,1}(x)\\ &= c_2^2a_1^p2^p\phi(h)f_{i,1}(x) + 2^pO_{p.s}(\mathrm{g}_{i,x}(h)) + 2^p\max(1,a_1^{2p})O_{p.s}(\mathrm{g}_{i,x}(h))\phi(h)(x)\\ &= (2\max(1,a_1^2))^p[c_2^2\phi(h)f_{i,1}(x) + O_{p.s}(\mathrm{g}_{i,x}(h)) + O_{p.s}(\mathrm{g}_{i,x}(h))\phi(h)(x)]\\ &= p!C^{p-2}[M\phi(h)f_{i,1}(x) + O_{p.c}(\mathrm{g}_{i,x}(h))]\\ &\leq p!C^{p-2}\phi(h)[Mb_i(x)+1],\end{aligned}$$

où $C = 2\max(1, a_1^2)$ et $M = (c_2C)^2$. Prenant encore $d_i^2 = \phi(h)[Mb_i(x) + 1]$, $D_n = \sum_{i=1}^n d_i^2$ et en observant par les hypothèses (A2)(ii) et (A2)(v) que $(1/n)D_n = \phi(h)[M+D(x)+1+o(1)]$ quand $n \to \infty$, nous sommes maintenant prêts à utiliser le Lemme 1.5.1 avec $D_n = O(n\phi(h))$ et $S_n = \sum_{i=1}^n L_{n,i}(x)$.

Ainsi, pour tout $\epsilon_0 > 0$, nous avons

$$\begin{aligned}\mathbb{P}\left(|\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)| > \epsilon_0\sqrt{\frac{\log n}{n\phi(h)}}\right) &= \mathbb{P}\left(\left|\sum_{k=1}^n L_{n,i}(x)\right| > n\mathbb{E}(\Delta_1(x))\epsilon_0\sqrt{\frac{\log n}{n\phi(h)}}\right)\\ &\leq 2\exp\left\{-\frac{(n\mathbb{E}(\Delta_1(x)\epsilon_0)^2\frac{\log n}{n\phi(h)}}{2D_n + 2Cn\mathbb{E}(\Delta_1(x))\epsilon_0\sqrt{\frac{\log n}{n\phi(h)}}}\right\}\\ &\leq 2\exp\left\{-\frac{\epsilon_0^2(O(n\phi(h)))^2\frac{\log n}{n\phi(h)}}{O(n\phi(h))\left(1+\sqrt{\frac{\log n}{n\phi(h)}}\right)}\right\}\\ &\leq 2\exp\{-C_1\epsilon_0^2\log n\}\\ &= \frac{2}{n^{C_1\epsilon_0^2}},\end{aligned}$$

où C_1 est une constante positive. Par conséquent, choisir ϵ_0 assez grand, on obtient
$$\sum_{i\geq 1} \mathbb{P}\left(|\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)| > \epsilon_0 \sqrt{\frac{\log n}{n\phi(h)}}\right) < \infty.$$
Finalement, par le lemme de Borel-Cantelli la preuve est achevée. ■

Preuve du Théorème 2.1.1. Utilisant la décomposition (2.4)-(2.6), le résultat suit une conséquence directe des Lemmes 2.1.2 et 2.1.3 et la Remarque 2.1.4. ■

2.2 Convergence uniforme presque sûre

2.2.1 Notations et hypothèses

Afin d'établir la convergence uniforme avec un taux de l'estimateur de la fonction de régression, nous avons besoin des hypothèses supplémentaires. Soit $\mathcal{C}$ une classe d'éléments de $\mathcal{E}$ et considérons, pour tout $\epsilon > 0$, le nombre suivant :

$$\begin{aligned}\mathcal{N}(\epsilon, \mathcal{C}, d) = \min\{n : &\text{ il existe } c_1, \ldots, c_n\, dans\, \mathcal{C} \text{ tel que } \forall x \in \mathcal{C} \\ &\text{il existe } 1 \leq k \leq n \text{ tel que } d(x, c_k) < \epsilon\},\end{aligned}$$

qui mesure la taille de la classe $\mathcal{C}$. Définissez maintenant les hypothèses suivantes.

(**B1**) Pour $x \in \mathcal{E}$, il existe une suite de fonctions aléatoires non-négatives $(f_{i,1}(x))_{i>1}$ bornées presque sûrement par une suite des quantités déterministes $(b_i(x))_{i>1}$, une fonction déterministe f_1 et une fonction réelle ϕ telle que

(i) $0 < b_0 \leq \inf_{x\in\mathcal{C}} f_1(x) \leq \sup_{x\in\mathcal{C}} f_1(x) < \infty$,

(ii) $F_x(u) = \phi(u)f_1(x) + o(\phi(u))$ quand $u \to 0$, où $o(\phi(u))$ est uniforme en x,

(iii) Pour tout $i \in \mathbb{N}, F_x^{\mathcal{F}_{i-1}}(u) = \phi(u)f_{i,1}(x) + \mathrm{g}_{i,x}(u)$ avec $\mathrm{g}_{i,x}(u) = o_{p.s}(\phi(u))$ quand $u \to 0, \mathrm{g}_{i,x}(u)/\phi(u)$ est bornée presque sûrement pour tout $x \in \mathcal{C}$ et $n^{-1}\sum_{i=1}^{n} \mathrm{g}_{i,x}(u) = o_{p.s}(\phi(u))$ quand $n \to \infty$ et $u \to 0$, où $o_{p.s}(\phi(u))$ est uniforme en x,

(iv) $\lim_{n\to\infty} \sup_{x\in\mathcal{C}} |n^{-1}\sum_{i=1}^{n} f_{i,1}(x) - f_1(x)| = 0$, presque sûrement,

(v) $n^{-1}\sum_{i=1}^{n} b_i(x) \to D(x)$, quand $n \to \infty$, et $\sup_{x\in\mathcal{C}} D(x) < \infty$.

(**B2**) Le noyau K satisfait les hypothèses

(i) K est une fonction de Holder d'ordre α avec une constante a_2,

(ii) il existe deux constantes a_0 et a_1 telles que
$0 < a_0 \leq K(x) \leq a_1 < \infty$ pour tout $x \in \mathcal{C}$.

(**B3**) Pour $1 \leq l \leq 2$, la suite des variables aléatoires $(Y_i^l)_{i\geq 1}$ est ergodique et $\mathbb{E}(|Y_1|^l) < \infty$.

(**B4**) Pour tout $(u,v) \in \mathcal{E}^2$, $|r(u)-r(v)| \leq c_1' d(u,v)^\beta$, pour certains $\beta > 0$ et une constante $c_1' > 0$.

Commentaires sur les hypothèses uniformes : Les conditions (B1)(i) et (B2) sont standard dans l'estimation de la régression non paramétrique alors que le fait que le noyau K borné par rapport à 0 est relativement habituel dans le cas des données fonctionnelles. Les hypothèses (B3) et (B4) sont des conditions de régularité assez légère sur la fonction de régression r et les moments de la variable aléatoire (Y_i). Les conditions (B1)(ii) et (B1)(iii) sont des versions uniformes des conditions (A2)(ii) et (A2)(iii). Elles sont également satisfaites en considérant les exemples suivants.

Exemple 1.

Pour $p \geq 1$, considérons une variable aléatoire $X \in \mathbb{R}^p$ et notons f sa densité par rapport à la mesure de Lebesgue λ. De toute évidence, nous avons

$$F_x(u) = \mathbb{P}(D_i \leq u) = \mathbb{P}(d(x,X) \leq u) = \int_{A(u)} f(y)dy,$$

où $A(u) = \{y \in \mathbb{R}^p : d(x,y) \leq u\}$. Notez, pour tout $x \in \mathbb{R}^p$, que $\lambda(A(u)) = (\pi^{p/2}/\Gamma(p/2+1))u^p$. Observez maintenant que

$$F_x(u) = f(x)\int_{A(u)} dy + \int_{A(u)} (f(y) - f(x))dy.$$

En supposant que f est une fonction Lipschitz uniforme d'ordre 1, nous obtenons

$$\int_{A(u)} |f(y) - f(x)|dy \leq C\int_{A(u)} d(x,y)dy \leq Cu\lambda(A(u)),$$

où la constante C est indépendante de x. Par conséquent, nous avons

$$F_x(u) = \phi(u)f(x) + o(\phi(u)),$$

où $o(\phi(u)) = Cu\lambda(A(u))$ est uniforme en x.

Exemple 2.

Considérons un espace de dimension infini et rappelons, dans ce cas, que la mesure de Lebesgue n'existe pas. Soit X le processus aléatoire défini par

$$X_t = \theta t + W_t, \quad -1 \leq t \leq 1,$$

où θ est une variable aléatoire de la loi $N(0,1)$ indépendante du processus de Wiener $W = \{W_t : -1 \leq t \leq 1\}$. Il est bien connu de Lipster et Shiryayev (1972) que la distribution μ_X du processus X est absolument continue par rapport la mesure de Wiener μ_W et que la densité de Radon-Nikodym est donnée par

$$\frac{d\mu_X}{d\mu_W}(x) = \frac{1}{\sqrt{2}} \exp\left\{\frac{x^2(1)}{4}\right\} = f(x).$$

Notez que

$$F_x(u) = \mathbb{P}(d(x,X) \leq u) = \mathbb{P}(X \in B(x,u)),$$

où $B(x,u)$ est la boule de centre x et de rayon u. Par conséquent, nous avons clairement

$$\begin{aligned} F_x(u) = \int_{B(x,u)} d\mu_X(z) &= \int_{B(x,u)} f(z)d\mu_W(z) \\ &= f(x)\mu_W(B(x,u)) + \int_{B(x,u)} (f(z) - f(x))d\mu_W(z). \end{aligned}$$

En prenant $d(x,z) = \sup\limits_{-1\leq t\leq 1} |x(t) - z(t)| < u$, il s'ensuit que $|x(1) - z(1)| < u$. Par conséquent, en augmentant la fonction exponentielle, on obtient

$$\begin{aligned} \exp\left\{\frac{z^2(1)}{4}\right\} - \exp\left\{\frac{x^2(1)}{4}\right\} &= \frac{z^2(1) - x^2(1)}{4} + \sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{z^2(1) - x^2(1)}{4}\right)^j \\ &= \frac{(z(1) - x(1))(z(1) + x(1))}{4} \\ &\quad \sum_{j=2}^{\infty} \frac{1}{j!}\left(\frac{(z(1) - x(1))(z(1) + x(1))}{4}\right)^j. \end{aligned}$$

Puisque $z \in B(x,u)$ et u est petit, nous avons

$$\begin{aligned}\left|\exp\left\{\frac{z^2(1)}{4}\right\} - \exp\left\{\frac{x^2(1)}{4}\right\}\right| &\leq \frac{u(2|x(1)|+u)}{4} + \sum_{j=2}^{\infty}\left(\frac{u}{4}\right)^j \frac{(2|x(1)|+u)^j}{j!} \\ &= \frac{u(2|x(1)|+u)}{4} + \left(\frac{u}{4}\right)^2 \exp\{2|x(1)|+u\}.\end{aligned}$$

Ainsi, pour $|x(1)| < \infty$, nous obtenons

$$F_x(u) = f(x)\mu_W(B(x,u)) + o(\mu_W(B(x,u))).$$

D'un autre côté, selon Théorème 3.1 dans Li et Shao (2001), pour tout $x \in H_{\mu_W}$, l'espace de Hilbert reproduit par μ_W, il en résulte, pour tout $u > 0$, que

$$\mu_W(B(x,u)) \leq \mu_W(B(0,u)) = \frac{4}{\pi}\exp\left\{-\frac{\pi^2}{8u^2}\right\}.$$

Donc, en prenant $\mathcal{C}$ comme la classe de la fonction x telle que $\sup_{x\in C}|x(1)| < \infty$, on obtient

$$F_x(u) = \phi(u)f(x) + o(\phi(u)),$$

où $o(\phi(u))$ est uniforme en x.

2.2.2 Résultats uniformes

Le théorème suivant donné la consistance uniforme presque sûre avec un taux de l'estimateur de régression $\hat{r}_n$.

Théorème 2.2.1. *[16] Supposons que les hypothèses (A1), (B1), (A2)(iv), (A3), (B2), (A4)(ii) et (B3)-(B4) soient vraies et que la condition (2.3) est satisfaite. En outre, pour une suite de nombres réels positifs λ_n tend vers zéro, quand $n \to \infty$, $\epsilon = \epsilon_n = o(h)$, supposons que nous avons*

$$(i)\ \lim_{n\to\infty}\frac{\log\mathcal{N}(\epsilon,\mathcal{C},d)}{\lambda_n^2 n\phi(h)} = 0 \quad et \quad (ii)\sum_{n\geq 1}\exp\{-\lambda_n^2 O(n\phi(h))\} < \infty. \quad (2.17)$$

Alors,

$$\sup_{x\in\mathcal{C}}|\hat{r}_n(x) - r(x)| = O_{p.s}(h^\beta) + O_{p.s}(\lambda_n).$$

Remarque 2.2.1. *Remplacer la condition (2.17)(i) par* $\lim_{n\to\infty} \log \mathcal{N}(\epsilon, \mathcal{C}, d)/\lambda_n^2 n\phi(h) = \delta$, *pour certains* $\delta > 0$, *avec* $\lambda_n = O((\log n/n\phi(h))^{1/2})$, *la condition (2.17)(ii) est clairement satisfaite, induisant la consistance uniforme de* $\hat{r}_n$ *avec le taux* $O_{p.s}(h^\beta) + O_{p.s}(\sqrt{\log n/n\phi(h)})$ *qui a été obtenu par Ferraty et Vieu (2006) dans le cas des données indépendantes.*

Remarque 2.2.2. *Notez, pour* $\epsilon = o(h)$ *et* $\lambda_n^{-1} = o((nh\phi(h))^{1/2})$, *que la condition (2.17)(i) prend la forme*

$$\lim_{\epsilon\to 0} \epsilon \log \mathcal{N}(\epsilon, \mathcal{C}, d) = 0. \tag{2.18}$$

Preuve : La preuve du théorème 2.2.1 est basée sur les lemmes suivants.

Lemme 2.2.1. *On suppose que les hypothèses (A1), (B1), (B2), (A2)(iv), (A3), (A4)(ii) et (B3)-(B4) sont vraies. De plus, supposons que* $\epsilon = \epsilon_n = o(h)$ *et la condition (2.17) est satisfaite. Alors nous avons*

$$\sup_{x\in\mathcal{C}} |\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)| = O_{p.s}(\lambda_n).$$

Remarque 2.2.3. *Prenant* $Y_i = 1$, *pour tout* $i \geq 1$, *dans les définitions de* $\hat{r}_{n,2}(x)$ *et* $\bar{r}_{n,2}(x)$, *on obtient sous les hypothèses (A1), (B1), (B2)(i), (B2)(ii), (A2)(iv) et (B4) le résultat suivant :*

$$\sup_{x\in\mathcal{C}} |\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)| = O_{p.s}(\lambda_n),$$

chaque fois que la condition (2.17) est satisfaite pour $\epsilon = \epsilon_n = o(h)$.

Remarque 2.2.4. *Sous les hypothèses du Lemme 2.2.1, en utilisant l'equation (2.4) et l'hypothèse (B1)(i), facilement nous avons*

$$\sup_{x\in\mathcal{C}} |Q_n(x)| = O_{p.s}(\lambda_n).$$

Preuve du Lemme 2.2.1. Pour $\epsilon > 0$, on a

$$\sup_{x\in\mathcal{C}} |\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)| \leq \max_{1\leq k\leq \mathcal{N}(\epsilon,C,d)} \sup_{x\in B(c_k,\epsilon)} |\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)|,$$

où

$$B(c_k, \epsilon) = \{x \in \mathcal{C} : d(x, c_k) < \epsilon\}$$

et

$$\begin{aligned}
\sup_{x\in\mathcal{C}} |\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)| \leq & \max_{1\leq k\leq \mathcal{N}(\epsilon,\mathcal{C},d)} \sup_{x\in B(c_k,\epsilon)} |\hat{r}_{n,2}(x) - \hat{r}_{n,2}(c_k)| \\
& + \max_{1\leq k\leq \mathcal{N}(\epsilon,\mathcal{C},d)} |\hat{r}_{n,2}(c_k) - \bar{r}_{n,2}(c_k)| \qquad (2.19) \\
& + \max_{1\leq k\leq \mathcal{N}(\epsilon,\mathcal{C},d)} \sup_{x\in B(c_k,\epsilon)} |\bar{r}_{n,2}(x) - \bar{r}_{n,2}(c_k)|.
\end{aligned}$$

Noter

$$H_{n,c_k}(\epsilon) = \sup_{x\in B(c_k,\epsilon)} |\hat{r}_{n,2}(x) - \hat{r}_{n,2}(c_k)|.$$

Observez que pour tout $x \in \mathcal{E}$

$$\begin{aligned}
\hat{r}_{n,2}(x) - \hat{r}_{n,2}(c_k) &= \frac{1}{n\mathbb{E}\Delta_1(x)\mathbb{E}\Delta_1(c_k)} \sum_{i=1}^{n} Y_i[\Delta_i(x)\mathbb{E}\Delta_1(c_k) - \Delta_i(c_k)\mathbb{E}\Delta_1(x)] \\
&= \frac{1}{n\mathbb{E}\Delta_1(x)} \sum_{i=1}^{n} Y_i[\Delta_i(x) - \Delta_i(c_k)] \\
&+ \frac{1}{n\mathbb{E}\Delta_1(x)\mathbb{E}\Delta_1(c_k)} \sum_{i=1}^{n} Y_i\Delta_i(c_k)[\mathbb{E}\Delta_1(c_k) - \mathbb{E}\Delta_1(x)] \\
&= I_1(c_k) + I_2(c_k).
\end{aligned} \qquad (2.20)$$

En utilisant l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned}
|I_1(c_k)| &\leq \frac{1}{n|\mathbb{E}\Delta_1(x)|} \sum_{i=1}^{n} |Y_i||\Delta_i(x) - \Delta_i(c_k)| \\
&\leq \frac{\sqrt{n}}{n|\mathbb{E}\Delta_1(x)|} \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^2\right)^{1/2} \left(\sum_{i=1}^{n}(\Delta_i(x) - \Delta_i(c_k))^2\right)^{1/2}.
\end{aligned}$$

Par l'hypothèse (B3), nous avons

$$|I_1(c_k)| \leq \frac{\sqrt{n}}{n|\mathbb{E}\Delta_1(x)|} O_{p.s}(E^{1/2}(Y_1^2)) \left(\sum_{i=1}^{n}(\Delta_i(x) - \Delta_i(c_k))^2\right)^{1/2}.$$

Par la suite, en utilisant l'hypothèse (B2)(i), nous obtenons

$$\begin{aligned}
|I_1(c_k)| &\leq \frac{1}{n^{1/2}|\mathbb{E}\Delta_1(x)|} O_{p.s}(E^{1/2}(Y_1^2)) \left(a_2 \sum_{i=1}^{n} \left| \frac{d(x,X_i) - d(c_k,X_i)}{h} \right|^{2\alpha} \right)^{1/2} \\
&\leq \frac{1}{n^{1/2}|\mathbb{E}\Delta_1(x)|} O_{p.s}(E^{1/2}(Y_1^2) \left(a_2 n \left(\frac{\epsilon}{h}\right)^{2\alpha} \right)^{1/2} \\
&\leq \frac{\epsilon^\alpha}{h^\alpha |\mathbb{E}\Delta_1(x)|} O_{p.s}(E^{1/2}(Y_1^2)).
\end{aligned}$$

De plus, clairement par l'hypothèse (B2)(ii), nous avons

$$\frac{\Delta_i(x)}{|\mathbb{E}\Delta_1(x)|} \leq \frac{a_1}{a_0} \quad \text{pour tout}\, x \in \mathcal{E}.$$

Donc,

$$|I_2(c_k)| \leq \frac{1}{n|\mathbb{E}\Delta_1(x)|} \sum_{i=1}^{n} |Y_i| \left| \frac{\Delta_i(c_k)}{\mathbb{E}\Delta_i(c_k)} \right| |\mathbb{E}\Delta_i(x) - \mathbb{E}\Delta_i(c_k)|\,.$$

Par l'inégalité de Cauchy-Schwarz et la propriété de K Lipschitz, on obtient

$$\begin{aligned}
|I_2(c_k)| &\leq \frac{a_1/a_0}{n^{1/2}|\mathbb{E}\Delta_1(x)|} \left(\sum_{i=1}^{n} Y_i^2 \right)^{1/2} \left(\sum_{i=1}^{n} (\mathbb{E}\Delta_1(c_k) - \mathbb{E}\Delta_1(x))^2 \right)^{1/2} \\
&\leq \frac{(a_1/a_0)\epsilon^\alpha}{h^\alpha |\mathbb{E}\Delta_1(x)|} O_{p.s}(E^{1/2}(Y_1^2)).
\end{aligned}$$

En utilisant l'hypothèse (B2)(ii), il s'ensuit, pour une $1 \leq k \leq \mathcal{N}(\epsilon, \mathcal{C}, d)$, que

$$I_1(c_k) + I_2(c_k) \leq \frac{(1 + a_1/a_0)\epsilon^\alpha}{h a_0} O_{p.s}(E^{1/2}(Y_1^2)),$$

qui est indépendant de x et c_k. Prendre maintenant $\epsilon = \epsilon_n$, nous obtenons

$$\max_{1 \leq k \leq \mathcal{N}(\epsilon,\mathcal{C},d)} \sup_{x \in B(c_k,\epsilon)} |\hat{r}_{n,2}(x) - \hat{r}_{n,2}(c_k)| = o(1).$$

De même et dans les mêmes conditions, on peut obtenir

$$\max_{1 \leq k \leq \mathcal{N}(\epsilon,\mathcal{C},d)} \sup_{x \in B(c_k,\epsilon)} |\bar{r}_{n,2}(x) - \bar{r}_{n,2}(c_k)| = o(1).$$

D'un autre côté, en utilisant le Lemme 1.5.1 et en procédant comme dans le Lemme 2.1.4 lorsque les hypothèses (B1)(i) et (B1)(v) sont supposées être satisfaites, il s'ensuit que

$$\begin{aligned}
P_n(\mathcal{C}) = \mathbb{P}\left(\max_{1\le k\le \mathcal{N}(\epsilon,\mathcal{C},d)} |\hat{r}_{n,2}(c_k)-\bar{r}_{n,2}(c_k)| > \lambda_n\right) &\le \sum_{i=1}^{\mathcal{N}(\epsilon,\mathcal{C},d)} \mathbb{P}(|\hat{r}_{n,2}(c_k)-\bar{r}_{n,2}(c_k)| > \lambda_n)\\
&\le 2\mathcal{N}(\epsilon,C,d)\exp\left\{-\frac{O(n\phi(h))\lambda_n^2}{1+C\lambda_n}\right\}\\
&\le 2\exp\Bigg\{-O(n\phi(h))\lambda_n^2\\
&\qquad\left[1-\frac{\log\mathcal{N}(\epsilon,\mathcal{C},d)}{\lambda_n^2 O(n\phi(h))}\right]\Bigg\}.
\end{aligned}$$

En utilisant maintenant la condition (2.17), on obtient $\sum_{n\ge 1} P_n(\mathcal{C}) < \infty$ et de réaliser la preuve du lemme de Borel-Cantelli. ■

Lemme 2.2.2. *On suppose que les hypothèses (A1), (A2)(iv) et (B1) sont satisfaites. Alors nous avons*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{C}} |\bar{r}_{n,1}(x)-1| = 0 \qquad \textit{presque sûrement.}$$

Si en plus les hypothèses (A2)(iv), (B4) avec des conditions (2.3) et (2.17), avec $\epsilon=\epsilon_n=o(h)$, rester vraies, nous avons

$$\lim_{n\to\infty}\sup_{x\in\mathcal{C}} |\hat{r}_{n,1}(x)-1| = 0 \quad \textit{presque sûrement.}$$

Preuve du Lemme 2.2.2. En utilisant la Remarque 2.1.2, facilement on obtient que

$$|\bar{r}_{n,1}(x)-1| \le \frac{1}{f_1(x)+o(1)}\left|\frac{1}{n}\sum_{i=1}^{n} f_{i,1}(x)-f_1(x)+\frac{1}{\phi(h)}O_{p.s}\left(\frac{1}{n}\sum_{i=1}^{n} \mathrm{g}_{i,x}(h)\right)\right|.$$

Par conséquent, par l'hypothèse (B1)(iii), nous avons

$$\sup_{x\in\mathcal{C}}|\bar{r}_{n,1}(x)-1| \le \frac{1}{\inf_{x\in\mathcal{C}} f_1(x)+o(1)}\left[\sup_{x\in\mathcal{C}}\left|\frac{1}{n}\sum_{i=1}^{n} f_{i,1}(x)-f_1(x)\right|+o_{p.s}(1)\right]. \tag{2.21}$$

Où $o(1)$ et $o_{p.s}(1)$ sont uniformes en x. Nous réalisons la preuve de la première partie en considérant les hypothèses (B1)(i) et (B1)(iv).
Maintenant, depuis

$$\sup_{x\in\mathcal{C}} |\hat{r}_{n,1}(x) - 1| \leq \sup_{x\in\mathcal{C}} |\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)| + \sup_{x\in C} |\bar{r}_{n,1}(x) - 1|,$$

nous concluons la preuve en vue la Remarque 2.2.3 et l'équation (2.21). ∎

Lemme 2.2.3. *On suppose que les hypothèses (A1), (A2)(iv), (B1) et (B4) sont satisfaites. Ensuite, chaque fois que la condition (2.17) est satisfaite avec* $\epsilon = \epsilon_n = o(h)$*, nous avons*

$$\sup_{x\in\mathcal{C}} B_n(x) = O_{p.s}(h^\beta) \tag{2.22}$$

et

$$\sup_{x\in\mathcal{C}} R_n(x) = O_{p.s}(h^\beta \lambda_n). \tag{2.23}$$

Preuve du Lemme 2.2.3. Compte tenu de l'équation (2.5) et de la Remarque 2.2.3, nous devons indiquer uniquement la partie (2.22) impliquant le biais conditionnel.

$$B_n(x) = \frac{\tilde{B}_n(x)}{\bar{r}_{n,1}(x)}.$$

pour que

$$\sup_{x\in\mathcal{C}} |B_n(x)| = \frac{\sup_{x\in\mathcal{C}} |\tilde{B}_n(x)|}{\inf_{x\in\mathcal{C}} |\bar{r}_{n,1}(x)|}.$$

Maintenant, en utilisant l'équation (2.4) et l' hypothèse (B4), on a

$$\sup_{x\in\mathcal{C}} |\tilde{B}_n(x)| = h^\beta \sup_{x\in\mathcal{C}} \left| \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}(\Delta_i(x)|\mathcal{F}_{i-1}) \right|.$$

Dans l' hypothèse (B1), facilement nous avons

$$\sup_{x\in\mathcal{C}} |\tilde{B}_n(x)| = O(h^\beta) \quad \text{p.s.}$$

De plus, comme $||a| - |b|| \leq |a - b|$ pour tout nombre réel a et b, il en résulte que

$$\inf_{x\in\mathcal{C}} |\bar{r}_{n,1}(x)| = \inf_{x\in\mathcal{C}} |\bar{r}_{n,1}(x) - 1 + 1| \geq 1 - \sup_{x\in C} |\bar{r}_{n,1}(x) - 1|. \tag{2.24}$$

Donc, á partir la Remarque 2.2.3, nous obtenons

$$\inf_{x\in\mathcal{C}} |\bar{r}_{n,1}(x)| = O(1) \quad \text{p.s.}$$

Ainsi,

$$\sup_{x\in\mathcal{C}} |B_n(x)| = O(h^\beta) \quad \text{p.s.}$$

■

Preuve du Théorème 2.2.1. Observez ceci

$$\begin{aligned} \sup_{x\in\mathcal{C}} |\hat{r}_n(x) - r(x)| \quad &\leq \quad \sup_{x\in\mathcal{C}} |B_n(x)| + \sup_{x\in\mathcal{C}} \left|\frac{R_n(x)}{\hat{r}_{n,1}(x)}\right| + \sup_{x\in\mathcal{C}} \left|\frac{Q_n(x)}{\hat{r}_{n,1}(x)}\right| \\ &\leq \sup_{x\in\mathcal{C}} |B_n(x)| + \frac{\sup_{x\in\mathcal{C}} |R_n(x)| + \sup_{x\in\mathcal{C}} |Q_n(x)|}{\inf_{x\in\mathcal{C}} |\hat{r}_{n,1}(x)|} \end{aligned} \tag{2.25}$$

De même que dans (2.24), en utilisant le Lemme 2.2.2, on obtient

$$\inf_{x\in\mathcal{C}} |\hat{r}_{n,1}(x)| \geq 1 - \sup_{x\in\mathcal{C}} |\hat{r}_{n,1}(x) - 1| = O(1) \quad \text{presque sûrement.}$$

Maintenant, il suffit d'utiliser la Remarque 2.2.4 et le Lemme 2.2.3 pour obtenir la preuve. ■

Chapitre 3

Normalité asymptotique

La propriété asymptotique abordée dans ce chapitre est la normalité asymptotique, il s'agit d'un sujet très important en statistique. En effet, la normalité asymptotique nous permet de construire les intervalles de confiance et de faire des tests. Ce chapitre est divisé en trois sections. Dans la première section nous regroupons l'ensemble des hypothèses utilisées pour établir notre résultat asymptotique. Le biais et la cohérence sera donnée dans la deuxième section. Dans la troisième section, nous énonçons le théorème et la démonstration détaillée.

3.1 Notations et hypothèses

Nos résultats sont énoncés sous certaines hypothèses que nous rassemblons ci-après pour faciliter la consultation

(H1) K est un noyau borné non négatif de classe C^1 sur son support $[0,1]$ et $K(1) > 0$. K' existe sur $[0,1]$ et satisfait la condition $K'(t) < 0$, $\forall t \in [0,1]$ et $\left|\int_0^1 (K^j)'(u)du\right| < \infty$ pour $j = 1, 2$.

(H2) Pour $x \in \mathcal{E}$, il existe une suite de fonctions aléatoires bornées non négatives $(f_{i,1})_{i\geq 1}$, une suite de fonctions aléatoires $(\mathrm{g}_{i,1})_{i\geq 1}$, une fonction déterministe non bornée déterminative f_1 et une fonction réelle non négative ϕ tendant vers zéro, quand argument tend à 0, de sorte que

(i) $F_x(u) = \phi(u)f_1(x) + o(\phi(u))$ quand $u \to 0$.

(ii) pour tout $i \in \mathbb{N}$, $F_x^{\mathcal{F}_{i-1}}(u) = \phi(u)f_{i,1}(x) + \mathrm{g}_{i,x}(u)$ avec $\mathrm{g}_{i,x}(u) = o_{p.s}(\phi(u))$ quand $u \to 0$ $\mathrm{g}_{i,x}(u)/\phi(u)$ bornée presque sûrement et $n^{-1}\sum_{i=1}^{n} \mathrm{g}_{i,x}^{j}(u) = o_{p.s}(\phi^j(u))$ quand $n \to \infty$, $j = 1, 2$.

(iii) $n^{-1}\sum_{i=1}^{n} f_{i,1}^{j}(x) \to f_1^j(x)$ presque sûrement quand $n \to \infty$ pour $j = 1, 2$.

(iv) Il existe une fonction τ_0 bornée non décroissante telle que, uniformément dans $u \in [0, 1]$,
$\frac{\phi(hu)}{\phi(h)} = \tau_0(u) + o(1)$, quand $h \downarrow 0$ et, pour $1 \leq j \leq 2 + \delta$ avec $\delta > 0$,
$\int_0^1 (K^j(u))'\tau_0(u)du < \infty$.

(H3) (i) La moyenne conditionnelle de Y_i sachant σ-algèbre $\mathcal{G}_{i-1}$ ne dépend que de X_i, i.e., pour tout $i \geq 1$, $\mathbb{E}(Y_i|\mathcal{G}_{i-1}) = r(X_i)$ presque sûrement.

(ii) La variance conditionnelle de Y_i sachant σ-algèbre $\mathcal{G}_{i-1}$ ne dépend que de X_i, i.e., pour tout $i \geq 1$, $\mathbb{E}((Y_i - r(X_i))^2|\mathcal{G}_{i-1}) = W_2(X_i)$ presque sûrement. De plus, la fonction W_2 est continue dans le voisinage de x, c'est-à-dire

$$\sup_{\{u:d(x,u)\leq h\}} |W_2(u) - W_2(x)| = o(1) \quad \text{quand } h \to 0.$$

(H4) (i) $|r(u)-r(v)| \leq c_3 d(u, v)^\beta$ pour tous $(u, v) \in \mathcal{E}^2$ et certaine $\beta > 0$ et une constante $c_3 > 0$.

(ii) Pour certains $\delta > 0$, $\mathbb{E}(|Y_1|^{2+\delta}) < \infty$ et la fonction $\overline{W}_{2+\delta}(u) = \mathbb{E}(|Y_i - r(x)|^{2+\delta}|X_i = u)$, $u \in \mathcal{E}$, est continue dans le voisinage de x.

Nous donnons maintenant quelques autres notations. Pour $j \geq 1$, définir

$$M_j = K^j(1) - \int_0^1 (K^j)'(u)\tau_0(u)du. \tag{3.1}$$

Définir le biais conditionnel de l'estimation de régression $\hat{r}_n(x)$ comme

$$B_n(x) = C_n(x) - r(x),$$

où

$$C_n(x) = \frac{\overline{r}_{n,2}(x)}{\overline{r}_{n,1}(x)}. \tag{3.2}$$

Discussions d'hypothèses :
Notons que, les conditions (H1) et (H2) sont semblables à les hypothèses (A1) et (A2) dans le chapitre précédent. Considérons le modèle de régression $Y_i = r(X_i) + \sigma(X_i)\epsilon_i$ où les variables aléatoires ϵ_i est différences de martingale par rapport au σ-algèbre $\mathcal{G}_i$ engendrée par les éléments aléatoires $((X_1, \epsilon_1), \ldots, (X_i, \epsilon_i), X_{i+1})$. Clairement, nous avons $\mathbb{E}(Y_i|\mathcal{G}_{i-1}) = r(X_i)$ presque sûrement. De plus, si en plus on suppose que $\mathbb{E}(\epsilon_i^2|\mathcal{G}_{i-1}) = 1$ presque sûrement, alors $\mathbb{E}((Y_i - r(X_i))^2|\mathcal{G}_{i-1})$ ne dépend que de X_i. Par conséquent, les énoncés de l'hypothèse (H3) sont satisfaits. L'hypothèse (H4) est une condition de régularité qui caractérise l'espace fonctionnel de notre modèle elle est nécessaire pour évaluer le terme de biais dans les résultats asymptotiques de ce mémoire.

3.2 Biais et consistance

Les résultats suivants donnent l'ordre du biais conditionnel et la consistance avec un taux de convergence de l'estimation de régression $\hat{r}_n(x)$. Avant de les afficher, tenez compte des conditions suivantes

(CB_1) Supposons, pour $i \geq 1$, que $\mathbb{E}(r(X_i) - r(x)|d(x, X_i), \mathcal{F}_{i-1}) = \mathbb{E}(r(X_i) - r(x)|d(x, X_i)) = \psi(d(x, X_i))$, où la fonction ψ est dérivable à 0. De plus, on suppose que $\psi(0) = 0$ et que $\psi'(0) \neq 0$.

(CB_2) Pour tout $i \geq 1$, la fonction aléatoire $f_{i,1}(x)$ apparaissant dans l'hypothèse (H2) est bornée presque sûrement par une fonction déterministe $b_i(x)$ tel que $n^{-1}\sum_{i=1}^{n} b_i(x) \to D(x) < \infty$, quand $n \to \infty$.

Proposition 3.2.1. *Sous les hypothèses (H1), (H2),* (CB_1) *et* (CB_2) *et le fait que* $f_1(x) > 0$ *et* $\left|\int_0^1 (sK(s))'\tau_0(s)ds\right| < \infty$*, nous avons*

$$\begin{aligned} B_n(x) &= C_n(x) - r(x) = \frac{h\psi'(0)}{M_1}\left[K(1) - \int_0^1 (sK(s))'\tau_0(s)ds + o_{p.s}(1)\right] \\ &+ O_{p.s}\left(h\sqrt{\frac{\log n}{n\phi(h)}}\right) + O_{p.s}\left(h\frac{\log n}{n\phi(h)}\right). \end{aligned}$$

Proposition 3.2.2. *Supposons que les hypothèses (H1)-(H4)(i) sont vraies.*

(a) Si

$$\frac{n\phi(h)}{\log\log(n)} \to \infty \quad quand \quad n \to \infty, \tag{3.3}$$

alors, pour tout $x \in \mathcal{E}$ tel que $f_1(x) > 0$, nous avons

$$\lim_{n\to\infty} \left(\frac{n\phi(h)}{\log\log(n)}\right)^{\frac{1}{2}} (\hat{r}_n(x) - C_n(x)) \stackrel{\mathbb{P}}{=} 0.$$

(b) Si en plus

$$\frac{n\phi(h)h^{2\beta}}{\log\log(n)} \to 0 \quad quand \quad n \to \infty, \tag{3.4}$$

où β est specifié dans (H4)(i), alors nous avons

$$\lim_{n\to\infty} \left(\frac{n\phi(h)}{\log\log(n)}\right)^{\frac{1}{2}} (\hat{r}_n(x) - r(x)) \stackrel{\mathbb{P}}{=} 0,$$

où $\stackrel{\mathbb{P}}{=}$ est l'égalité en probabilité.

3.3 Normalité asymptotique

Théorème 3.3.1 ci-dessous traité de la normalité asymptotique de $\hat{r}_n(x)$.

Théorème 3.3.1. *[24] Supposons que les hypothèses (H1)-(H4) sont vraies et que*

$$n\phi(h) \to \infty \quad quand \quad n \to \infty. \tag{3.5}$$

Alors, pour tout $x \in \mathcal{E}$ tel que $f_1(x) > 0$, nous avons

(i) $\sqrt{n\phi(h)}(\hat{r}_n(x) - C_n(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x))$,

où

$$\sigma^2(x) = \frac{M_2}{M_1^2}\frac{W_2(x)}{f_1(x)} \tag{3.6}$$

et $\xrightarrow{\mathcal{D}}$ désigne la convergence en loi.

(ii) Si en plus nous supposons que

$$h^{\beta}(n\phi(h))^{1/2} \to 0 \quad quand \quad n \to \infty, \tag{3.7}$$

où β est spécifié dans l'hypothèse (H4), alors nous avons

$$\sqrt{n\phi(h)}(\hat{r}_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)).$$

Remarque 3.3.1. *(i) Notez que les constantes M_1 et M_2 soient strictement positives. En effet, en utilisant l'hypothèse (H1) et le fait que la fonction τ_0 est non décroissante, il suffit d'effectuer une simple integration par partie. Par consèquent, si $W_2(x) > 0$, nous avons $\sigma^2(x) > 0$.*

(ii) Si $\mathcal{E} = \mathbb{R}^d$, l'expression de la variance asymptotique prend la forme

$$\sigma^2(x) = \frac{1}{d}\frac{W_2(x)}{f_d(x)}\frac{\displaystyle\int_0^1 K^2(u)u^{d-1}du}{\left(\displaystyle\int_0^1 K(u)u^{d-1}du\right)^2},$$

où $f_d(x)$ est la densité marginale du vecteur aléatoire $X_{i,d}$ définie dans 2.1.

Remarquer maintenant dans le Théorème 3.3.1 que la variance asympyotique contient la fonction inconnue f_1 et que la normalisation dépend de la fonction ϕ qui n'est pas explicitement identifiable. De plus, nous devons estimer les quantités W_2 et τ_0.
Donc, le Corollaire 3.3.1, ci-dessous, qui est une légère modification de Théorème 3.3.1, permet d'avoir une forme utilisable de nos résultats dans la pratique.
Comme d'habitude, la variance conditionnelle de $W_2(x)$ est estimée par

$$\begin{aligned} W_{2,n}(x) &= \frac{\displaystyle\sum_{i=1}^n (Y_i - \hat{r}_n(x))^2 K\left(\frac{d(x,X_i)}{h}\right)}{\displaystyle\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h}\right)} \\ &= \frac{\displaystyle\sum_{i=1}^n Y_i^2 K\left(\frac{d(x,X_i)}{h}\right)}{\displaystyle\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h}\right)} - (\hat{r}_n(x))^2 \\ &= \hat{g}_n(x) - (\hat{r}_n(x))^2. \end{aligned} \tag{3.8}$$

En utilisant la dècomposition de $F_X(u)$ dans l'hypothèse (H2)(i), on peut estimer $\tau_0(u)$ par

$$\tau_n(u) = \frac{F_{x,n}(uh)}{F_{x,n}(h)},$$

où

$$F_{x,n}(u) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{d(x,X_i)\leq u\}}.$$

Par la suite, pour un noyau K donnée, les quantités M_1 et M_2 sont estimées par $M_{1,n}$ et $M_{2,n}$ remplaçant respectivement τ_0 par τ_n dans leurs expressions respectives. Introduisons maintenant des conditions supplémentaires nécessaires pour énoncer le Corollaire 3.3.1.

(H5) (i) La moyenne conditionnelle de Y_i^2 sachant σ-algèbre $\mathcal{G}_{i-1}$ ne dépend que X_i, i.e., il existe une fonction g de telle sorte que, pour tout $i \geq 1$, $\mathbb{E}(Y_i^2|\mathcal{G}_{i-1}) = \mathrm{g}(X_i)$ presque sûrement,

(ii) la variance conditionnelle de Y_i^2 sachant σ-algèbre $\mathcal{G}_{i-1}$ ne dépend que X_i, i.e., pour toute $i \geq 1, \mathbb{E}((Y_i^2 - \mathrm{g}(X_i))^2|\mathcal{G}_{i-1}) = U(X_i)$, presque sûrement, pour une certaine fonction U. De plus, la fonction U est continue dans le voisinage de x, c'est-à-dire, $\sup_{\{u:d(x,u)\leq h\}} |U(u) - U(x)| = o(1)$.

Corollaire 3.3.1. *Supposons que les hypothèses (H1)-(H5) sont vraies, K' et $(K^2)'$ sont des fonctions intégrables et*

$$nF_x(h) \to \infty \quad et \quad h^{\beta}(nF_x(h))^{1/2} \to 0 \quad quand\ n \to \infty. \tag{3.9}$$

Alors, pour tout $x \in \mathcal{E}$ tel que $f_1(x) > 0$, nous avons

$$\frac{M_{1,n}}{\sqrt{M_{2,n}}}\sqrt{\frac{nF_{x,n}(h)}{W_{2,n}(x)}}(\hat{r}_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Preuve :

Afin d'etablir nos résultats, introduire quelques notations supplémentaires. Pour $x \in \mathcal{E}$, définir

$$Q_n(x) = (\hat{r}_{n,2}(x) - \bar{r}_{n,2}(x)) - r(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)) \tag{3.10}$$

et

$$R_n(x) = -B_n(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)). \tag{3.11}$$

De toute évidence, nous avons

$$\hat{r}_n(x) - C_n(x) = \frac{Q_n(x) + R_n(x)}{\hat{r}_{n,1}(x)} \tag{3.12}$$

L'intérêt majeur de la décomposition (3.12) vient du fait que les résumés du terme $Q_n(x)$ forment une différence de martingale qui permettent d'établir le théorème central limite pour le terme central $Q_n(x)$. La preuve du Théorème 3.3.1 est divisée en plusieurs lemmes établissant respectivement la convergence de la probabilité de $\hat{r}_{n,1}(x)$ à 1, le fait que $R_n(x)$, avant ageusement normalisée, est en fait égal à $o_{\mathbb{P}}(1)$ et la normalité asymptotique de $Q_n(x)$.

Le lemme suivant décrit le comportement asymptotique du terme $\hat{r}_{n,1}(x)$

Lemme 3.3.1. *Supposons que les hypothèses (H1)-(H2) et (3.5) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$ tel que $f_1(x) > 0$, nous avons*

$$\lim_{n\to\infty} \hat{r}_{n,1}(x) \stackrel{\mathbb{P}}{=} 1.$$

Preuve du Lemme 3.3.1. On a

$$\hat{r}_{n,1}(x) - 1 = R_{1,n}(x) + R_{2,n}(x), \tag{3.13}$$

où

$$\begin{aligned} R_{1,n}(x) &= \frac{1}{n\mathbb{E}[\Delta_1(x)]}\sum_{i=1}^{n}(\Delta_i(x) - \mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}]), \\ R_{2,n}(x) &= \frac{1}{n\mathbb{E}[\Delta_1(x)]}\sum_{i=1}^{n}(\mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}] - \mathbb{E}\Delta_1(x)) \\ &= \frac{1}{n\mathbb{E}[\Delta_1(x)]}\sum_{i=1}^{n}\mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}] - 1. \end{aligned}$$

En combinant le Lemme 2.1.1 avec les hypothèses (H2)(ii) et (H2)(iii), on voit facilement que $R_{2,n}(x) = o_{p.s}(1)$ quand $n \to \infty$.

Pour gérer le premier terme, observez que $R_{n,1}(x) = \sum_{i=1}^{n} L_{ni}(x)$, où $\{L_{ni}(x)\}$ est une matrice triangulaire de différence de martingale par rapport à la σ-algèbre $\mathcal{F}_{i-1}$. La combinaison de Burkholder ([27], page 23) et les inégalités de Jensen, nous obtenons pour tout $\epsilon > 0$ qu'il existe une constante $C_0 > 0$ tel que

$$\begin{aligned} \mathbb{P}(|R_{n,1}(x)| > \epsilon) &\leq C_0 \frac{\mathbb{E}(\Delta_1^2(x))}{\epsilon^2 n(\mathbb{E}(\Delta_1(x)))^2} \\ &= O\left(\frac{1}{\epsilon^2 n\phi(h)} + o(1)\right), \end{aligned}$$

la dernière égalité est déduit du Lemme 2.1.1. Depuis $n\phi(h) \to \infty$ quand $n \to \infty$, nous concluons alors que $R_{n,1}(x) = o_{\mathbb{P}}(1)$ quand $n \to \infty$. ■

Lemme 3.3.2. *Supposons que les hypothèses (H1)-(H2), (H3)(i), (H4)(i) et la condition (3.5) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$ tel que $f_1(x) > 0$, nous avons*

$$B_n(x) = O_{\mathbb{P}}(h^\beta) \tag{3.14}$$

et

$$R_n(x) = O_{\mathbb{P}}\left(\frac{h^\beta}{\sqrt{n\phi(h)}}\right). \tag{3.15}$$

Preuve du Lemme 3.3.2. D'abord, nous évaluons le terme de biais conditionnel. On a

$$B_n(x) = \frac{\bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x)}{\bar{r}_{n,1}(x)}.$$

De même que dans le Lemme 3.3.1, on voit facilement que $\bar{r}_{n,1}(x) - 1 = o_{\mathbb{P}}(1)$. Par conséquent, nous devons établir que

$$\tilde{B}_n(x) = \bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x) = O_{p.s}(h^\beta).$$

En utilisant les hypothèses (H3)(i) et (H4)(i), on peut facilement voir que

$$\begin{aligned}
|\tilde{B}_n(x)| &= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n \mathbb{E}[(Y_i - r(x))\Delta_i(x)|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n \mathbb{E}[\mathbb{E}[(Y_i - r(x))\Delta_i(x)|\mathcal{G}_{i-1}]|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n \mathbb{E}[\mathbb{E}[(Y_i - r(x))\Delta_i(x)|X_i]|\mathcal{F}_{i-1}]\right| \\
&= \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n \mathbb{E}[(r(X_i) - r(x))\Delta_i(x)|\mathcal{F}_{i-1}]\right| \\
&\leq \sup_{u\in B(x,h)}|r(u) - r(x)| \left|\frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n \mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}]\right| = O_{p.s}(h^\beta),
\end{aligned}$$

puisque le support du noyau K est l'intervalle $[0,1]$.
Pour prouver la deuxième partie du Lemme 3.3.2, on a

$$R_n(x) = -B_n(x)(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)),$$

observons que $\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)$ est une différence de martingale. En suivant maintenant les mêmes étapes que dans le Lemme 3.3.3 ci-dessous, nous pouvons établir que

$$\sqrt{n\phi(h)}(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \rho^2(x)),$$

où

$$\rho^2 = \frac{M_2}{M_1^2}\frac{1}{f_1(x)}.$$

Par conséquent, nous avons clairement

$$\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x) = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n\phi(h)}}\right).$$

En combinant cette équation avec la première partie du Lemme 3.3.2 , nous obtenons le résultat désiré. ■

Le lemme suivant établit la normalité asymptotique du processus Q_n.

Lemme 3.3.3. *Supposons que les hypothèses (H1)-(H4) et la condition (3.5) sont satisfaites. Alors, pour tout $x \in \mathcal{E}$ de telle sorte que $f_1(x) > 0$, nous avons*

$$\sqrt{n\phi(h)}Q_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)), \quad \textit{quand}\, n \to \infty.$$

Preuve du Lemme 3.3.3. Tout d'abord, nous présentons quelques notations. On pose

$$\eta_{ni} = \left(\frac{\phi(h)}{n}\right)^{1/2}(Y_i - r(x))\frac{\Delta_i(x)}{\mathbb{E}\Delta_1(x)} \tag{3.16}$$

et $\xi_{ni} = \eta_{ni} - \mathbb{E}[\eta_{ni}|\mathcal{F}_{i-1}]$. On voit facilement que

$$(n\phi(h))^{1/2}Q_n(x) = \sum_{i=1}^{n}\xi_{ni}, \tag{3.17}$$

où, pour tout x dans $\mathcal{E}$, les résumés (3.17) forment un tableau triangulaire de différences de martingale stationnaire par rapport au σ-algèbre $\mathcal{F}_{i-1}$. Cela nous permet d'appliquer le théorème central limite pour les martingales à valeurs réelles en temps discret (voir [6], Page 23) pour établir la normalité asymptotique de $Q_n(x)$. Cela peut se faire que si nous établissons les déclarations suivantes :

(a) $\sum_{i=1}^{n} \mathbb{E}[\xi_{ni}^2|\mathcal{F}_{i-1}] \xrightarrow{\mathbb{P}} \sigma^2(x)$

et

(b) $n\mathbb{E}[\xi_{ni}^2 I_{[|\xi_{ni}|>\epsilon]}] = o(1)$ pour tout $\epsilon > 0$ (condition de Lindeberg).

Preuve de la partie (a). Observez d'abord que

$$\left|\sum_{i=1}^{n} \mathbb{E}[\eta_{ni}^2|\mathcal{F}_{i-1}] - \sum_{i=1}^{n} \mathbb{E}[\xi_{ni}^2|\mathcal{F}_{i-1}]\right| \leq \sum_{i=1}^{n} (\mathbb{E}[\eta_{ni}|\mathcal{F}_{i-1}])^2. \tag{3.18}$$

En utilisant l'hypothèse (H4)(i) et le Lemme 2.1.1, on a

$$\begin{aligned} |\mathbb{E}[\eta_{ni}|\mathcal{F}_{i-1}]| &= \frac{1}{\mathbb{E}\Delta_1(x)}\left(\frac{\phi(h)}{n}\right)^{1/2} |\mathbb{E}[(r(X_i)-r(x))\Delta_i(x)|\mathcal{F}_{i-1}]| \\ &\leq \frac{1}{\mathbb{E}\Delta_1(x)}\left(\frac{\phi(h)}{n}\right)^{1/2} \sup_{u\in B(x,h)} |r(u)-r(x)|\mathbb{E}[\Delta_i(x)|\mathcal{F}_{i-1}] \\ &= O(h^\beta)\left(\frac{\phi(h)}{n}\right)^{1/2}\left(\frac{f_{i,1}(x)}{f_1(x)} + O_{p.s}\left(\frac{\mathrm{g}_{i,x}(h)}{\phi(h)}\right)\right). \end{aligned}$$

Ainsi, par les hypothèses (H2)(ii)-(iii), nous avons

$$\begin{aligned} \sum_{i=1}^{n} (\mathbb{E}[\eta_{ni}|\mathcal{F}_{i-1}])^2 &= O(h^{2\beta})\left(\frac{\phi(h)}{n}\right)\sum_{i=1}^{n}\left(\frac{f_{i,1}(x)}{f_1(x)} + O_{p.s}\left(\frac{\mathrm{g}_{i,x}(h)}{\phi(h)}\right)\right)^2 \\ &= O(h^{2\beta}\phi(h))\left(\frac{1}{f_1^2(x)}\frac{1}{n}\sum_{i=1}^{n} f_{i,1}^2(x) + o_{p.s}(1)\right) \\ &= O_{p.s}(\phi(h)h^{2\beta}). \end{aligned}$$

La déclaration (a) suit alors si nous montrons que

$$\lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}[\eta_{ni}^2|\mathcal{F}_{i-1}] \stackrel{\mathbb{P}}{=} \sigma^2(x). \tag{3.19}$$

On a

$$\begin{aligned} \sum_{i=1}^{n} \mathbb{E}[\eta_{ni}^2|\mathcal{F}_{i-1}] &= \frac{\phi(h)}{n(\mathbb{E}\Delta_1(x))^2}\sum_{i=1}^{n} \mathbb{E}[(Y_i - r(x))^2\Delta_i^2(x)|\mathcal{F}_{i-1}] \\ &= J_{1n} + J_{2n}, \end{aligned} \tag{3.20}$$

où

$$\begin{aligned} J_{1n} &= \frac{\phi(h)}{n(\mathbb{E}\Delta_1(x))^2}\sum_{i=1}^{n}\mathbb{E}\{\Delta_i^2(x)\mathbb{E}[(Y_i - r(X_i))^2|\mathcal{G}_{i-1}]|\mathcal{F}_{i-1}\} \\ &= \frac{\phi(h)}{n(\mathbb{E}\Delta_1(x))^2}\sum_{i=1}^{n}\mathbb{E}[W_2(X_i)\Delta_i^2(x)|\mathcal{F}_{i-1}] \quad par(H3)(ii) \end{aligned} \tag{3.21}$$

et

$$J_{2n} = \frac{\phi(h)}{n(\mathbb{E}\Delta_1(x))^2}\sum_{i=1}^{n}\mathbb{E}[(r(X_i) - r(x))^2\Delta_i^2(x)|\mathcal{F}_{i-1}]. \tag{3.22}$$

Nous donnons maintenant une limite supérieure pour $\mathbb{E}[W_2(X_i)\Delta_i^2(x)|\mathcal{F}_{i-1}]$. À cette fin, nous diviser en $I_{n1} + I_{n2}$ avec

$$I_{n1} = W_2(x)\mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}] \quad \text{et} \quad I_{n2} = \mathbb{E}[(W_2(X_i) - W_2(x))\Delta_i^2(x)|\mathcal{F}_{i-1}]. \tag{3.23}$$

En utilisant l'hypothèse (H3)(ii), on peut écrire

$$\begin{aligned} |I_{n2}| &\leq \sup_{\{u:d(x,u)\leq h\}} |W_2(u) - W_2(x)|\mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}] \\ &= \mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}] \times o(1). \end{aligned}$$

Ainsi, dans la vue du Lemme 2.1.1 partie (i), nous avons

$$\begin{aligned} \mathbb{E}[W_2(X_i)\Delta_i^2(x)|\mathcal{F}_{i-1}] &= (o(1) + W_2(x))\mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}] \\ &= (o(1) + W_2(x))(M_2\phi(h)f_{i,1}(x) + O_{p.s}(\mathrm{g}_{i,x}(h)). \end{aligned} \tag{3.24}$$

La combinaison de nouveau Lemme 2.1.1 et les hypothèses (H2)(ii)-(iii), on voit facilement que

$$\lim_{n\to\infty} J_{1n} = \frac{M_2}{M_1^2}\frac{W_2(x)}{f_1(x)} \quad \text{presque sûrement,} \tag{3.25}$$

chaque fois que $f_1(x) > 0$. Considérez maintenant le terme J_{2n}. En utilisant les hypothèses (H2)(ii)-(iii) et (H4)(i) et le Lemme 2.1.1, on peut écrire

$$\begin{aligned} |J_{n2}| &= O(h^{2\beta})\frac{\phi(h)}{n(\mathbb{E}(\Delta_1(x)))^2}\sum_{i=1}^{n}\mathbb{E}[\Delta_i^2(x)|\mathcal{F}_{i-1}] \\ &= O(h^{2\beta})\left(\frac{M_2}{M_1^2}\frac{1}{f_1(x)} + o_{p.s}(1)\right) \to 0 \quad \text{presque sûrement } quand\ n \to \infty. \end{aligned} \tag{3.26}$$

Donc,
$\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\mathbb{E}[\eta_{ni}^2|\mathcal{F}_{i-1}] = \lim\limits_{n\to\infty}(J_{n1}+J_{n2}) = \dfrac{M_2}{M_1^2}\dfrac{W_2(x)}{f_1(x)} = \sigma^2(x)$ presque sûrement
chaque fois que $f_1(x) > 0$. Ceci termine la preuve de la partie (a).

Preuve de la partie (b). La condition de Lindeberg résulte du Corollaire (9.5.2 [33]) implique que $n\mathbb{E}[\xi_{ni}^2 I(|\xi_{ni}| > \epsilon)] \leq 4n\mathbb{E}[\eta_{ni}^2 I(|\eta_{ni}| > \epsilon/2)]$. Soit $a > 1$ et $b > 1$ de telle sorte que $\frac{1}{a}+\frac{1}{b} = 1$. En utilisant les inégalités de Holder et de Markov on peut écrire, pour tout $\epsilon > 0$,

$$\mathbb{E}[\eta_{ni}^2 I(|\eta_{ni}| > \epsilon/2)] \leq \frac{\mathbb{E}|\eta_{ni}|^{2a}}{(\epsilon/2)^{2a/b}}. \tag{3.27}$$

En prenant C_0 une constante positive et $2a = 2+\delta$ (avec δ comme dans l'hypothèse (H4)(ii)) et en utilisant l'hypothèse (H4)(ii), on obtient

$$\begin{aligned}
4n\mathbb{E}[\eta_{ni}^2 I(|\eta_{ni}| > \epsilon/2)] &\leq C_0\left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2}\frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}}\mathbb{E}([|Y_i - r(x)|\Delta_i(x)]^{2+\delta})\\
&\leq C_0\left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2}\frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}}\mathbb{E}[\mathbb{E}[|Y_i - r(x)|^{2+\delta}(\Delta_i(x))^{2+\delta}|X_i]]\\
&\leq C_0\left(\frac{\phi(h)}{n}\right)^{(2+\delta)/2}\frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}}\mathbb{E}[(\Delta_i(x))^{2+\delta}\overline{W}_{2+\delta}(X_i)]\\
&\leq C_0\left(\frac{\phi(h)}{n}\right)^{2+\delta/2}\frac{n}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}}(\mathbb{E}[(\Delta_i(x))^{2+\delta}|\overline{W}_{2+\delta}(X_i)\\
&- \overline{W}_{2+\delta}(x)|] + |\overline{W}_{2+\delta}(x)|\mathbb{E}[(\Delta_1(x))^{2+\delta}])\\
&\leq C_0\left(\frac{\phi(h)}{n}\right)^{2+\delta/2}\frac{n\mathbb{E}[(\Delta_1(x))^{2+\delta}]}{(\mathbb{E}(\Delta_1(x)))^{2+\delta}}(|\overline{W}_{2+\delta}(x)| + o(1))\\
&\leq C_0(n\phi(h))^{-\delta/2}\frac{(M_{2+\delta}f_1(x)+o(1))}{(M_1^{2+\delta}f_1^{2+\delta}+o(1))}(|\overline{W}_{2+\delta}(x)| + o(1))\\
&= O((n\phi(h))^{-\delta/2}),
\end{aligned} \tag{3.28}$$

la dernière égalité découle du Lemme 2.1.1. Ceci termine la preuve de la partie (b) ($n\phi(h)\to\infty$ quand $n\to\infty$) et donc la preuve du Lemme 3.3.3
. ■

Preuve de la Proposition 3.2.1. Dans un premier temps, observez que le biais conditionnel se peut décomposer comme suit

$$\begin{aligned}
B_n(x) &= \frac{\bar{r}_{n,2}(x) - r(x)\bar{r}_{n,1}(x)}{\hat{r}_{n,1}(x)} + \frac{(\bar{r}_{n,2}(x) - r(x)\hat{r}_{n,1}(x))(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x))}{(\hat{r}_{n,1}(x))^2} \\
&+ \frac{(\hat{r}_{n,1}(x) - \bar{r}_{n,1}(x))^2 C_n(x)}{(\hat{r}_{n,1}(x))^2} \\
&= B_n^* + U_n(x)
\end{aligned}$$

où B_n^* se présente comme le principal en tout terme de $U_n(x)$ est celui résiduel. Considérant $N_n(x)$ comme numérateur de B_n^*, alors par l'hypothèse (H3)(i) nous avons

$$\begin{aligned}
N_n(x) &= \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} [\mathbb{E}(Y_i \Delta_i(x)|\mathcal{F}_{i-1}) - r(x)\mathbb{E}(\Delta_i(x)|\mathcal{F}_{i-1})] \\
&= \frac{1}{n\mathbb{E}(\Delta_1(x))} \sum_{i=1}^{n} \mathbb{E}[\Delta_i(x)(r(X_i) - r(x))|\mathcal{F}_{i-1}].
\end{aligned}$$

Observez maintenant, en utilisant la condition (CB_1), que

$$\begin{aligned}
A_i &= \mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)(r(X_i) - r(x))|\mathcal{F}_{i-1}\right] \\
&= \mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)\mathbb{E}[(r(X_i) - r(x))|d(x,X_i),\mathcal{F}_{i-1}]|\mathcal{F}_{i-1}\right] \\
&= \mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)\psi(d(x,X_i))|\mathcal{F}_{i-1}\right] \\
&= \int_0^1 K(t)\psi(th)dF_x^{\mathcal{F}_{i-1}}(th).
\end{aligned}$$

Comme $\psi(0) = 0$, le développement de Taylor de la fonction ψ jusqu'à l'ordre 1 au voisinage de $t = 0$ donné par

$$\psi(th) = th\psi'(0) + o(h).$$

Par conséquent, nous avons

$$\begin{aligned}
A_i &= h\psi'(0)\int_0^1 tK(t)dF_x^{\mathcal{F}_{i-1}}(th) + o(h)\int_0^1 K(t)dF_x^{\mathcal{F}_{i-1}}(th) \\
&= h\psi'(0)\left[K(1)F_x^{\mathcal{F}_{i-1}}(h) - \int_0^1 (sK(s))' F_x^{\mathcal{F}_{i-1}}(sh)ds\right] \\
&+ o(h)\left[K(1)F_x^{\mathcal{F}_{i-1}}(h) - \int_0^1 K'(s)F_x^{\mathcal{F}_{i-1}}(sh)ds\right].
\end{aligned}$$

En utilisant l'hypothèse (H2)(ii), on obtient

$$\begin{aligned}
A_i &= K(1)(\phi(h)f_{i,1}(x)+\mathrm{g}_{i,x}(h))(h\psi'(0)+o(h))-h\psi'(0)\int_0^1(sK(s))'(\phi(hs)f_{i,1}(x)\\
&+ \mathrm{g}_{i,x}(hs))ds-o(h)\int_0^1 K'(s)(\phi(hs)f_{i,1}(x)+\mathrm{g}_{i,x}(hs))ds\\
&= K(1)\phi(h)\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(h)}{\phi(h)}\right)(h\psi'(0)+o(h))-h\phi(h)\psi'(0)\int_0^1(sK(s))'\frac{\phi(hs)}{\phi(h)}\\
&\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(hs)}{\phi(uh)}\right)ds-o(h)\phi(h)\int_0^1 K'(s)\frac{\phi(hs)}{\phi(h)}\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(hs)}{\phi(uh)}\right)ds.
\end{aligned}$$

En procédant comme dans la preuve du Lemme 2.1.1, nous avons

$$\begin{aligned}
A_i &= K(1)h\phi(h)\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(h)}{\phi(h)}\right)(\psi'(0)+o(1))\\
&- h\phi(h)\psi'(0)\int_0^1(sK(s))'(\tau_0(s)+o(1))\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(hs)}{\phi(hs)}\right)ds\\
&- o(h)\phi(h)\int_0^1 K'(s)(\tau_0(s)+o(1))\left(f_{i,1}(x)+\frac{\mathrm{g}_{i,x}(hs)}{\phi(hs)}\right)ds\\
&= \psi'(0)h\phi(h)f_{i,1}(x)\left[K(1)-\int_0^1(sK(s))'\tau_0(s)ds\right]+O_{p.s}(h\mathrm{g}_{i,x}(h)).
\end{aligned}$$

Ainsi, en utilisant les hypothèses (H2)(ii)-(iii) et le Lemme 2.1.1, on obtient

$$\begin{aligned}
N_n(x) &= \frac{1}{n\mathbb{E}(\Delta_1(x))}\sum_{i=1}^n A_i\\
&= \frac{h\psi'(0)}{M_1}\left[K(1)-\int_0^1(sK(s))'\tau_0(s)ds+o_{p.s}(1)\right].
\end{aligned}$$

Considérant le terme résiduel, on voit facilement que

$$U_n(x)=N_n(x)\frac{\hat{r}_{n,1}(x)-\bar{r}_{n,1}(x)}{\hat{r}^2_{n,1}(x)}+\frac{B_n^*(x)}{\hat{r}_{n,1}(x)\bar{r}_{n,1}(x)}(\hat{r}_{n,1}(x)-\bar{r}_{n,1}(x))^2.$$

Il est indiqué dans [24], dans les hypothèses (H1), (H2) et (CB_2) combinées avec le fait que le noyau K est borné, que

$$\hat{r}_{n,1}(x)-\bar{r}_{n,1}(x)=O_{p.s}\left(\sqrt{\frac{\log n}{n\phi(h)}}\right).$$

Donc, nous concluons la preuve en utilisant du fait que les deux $\hat{r}_{n,1}(x)$ et $\bar{r}_{n,1}(x)$ convergent presque sûrement vers 1. ∎

Preuve de la proposition 3.2.2(a). À la suite de la décomposition (3.12), puisque par le Lemme 3.3.1, $\hat{r}_{n,1}(x) \xrightarrow{\mathbb{P}} 1$ quand $n \to \infty$, il suffit de montrer que $Q_n(x)$ et $R_n(x)$ convergent en probabilité vers zéro avec une vitesse appropriée.

Par le Lemme 3.3.2, nous avons $R_n(x) = O_{\mathbb{P}}\left(\frac{h^\beta}{\sqrt{n\phi(h)}}\right)$. Par conséquent, nous obtenons

$$\begin{aligned}\left(\frac{n\phi(h)}{\log\log(n)}\right)^{\frac{1}{2}} R_n(x) &= O_{\mathbb{P}}\left(\frac{h^\beta}{\log\log n}\right)\\ &= o_{\mathbb{P}}(1).\end{aligned}$$

Puisque $Q_n(x)$ est une somme des différences de martingales centrées, en utilisant successivement l'inégalité de Burkholder, l'inégalité - C_r et l'inégalité de Jensen, il s'ensuit que

$$\begin{aligned}Var(Q_n(x)) &\leq \frac{C_0}{n^2(\mathbb{E}(\Delta_1(x)))^2}\sum_{i=1}^n \mathbb{E}[\Delta_i^2(x)\mathbb{E}[(Y_i - r(x))^2|X_i]]\\ &= \frac{C_0}{n^2(\mathbb{E}(\Delta_1(x)))^2}\sum_{i=1}^n \mathbb{E}[\Delta_i^2(x)W_2(X_i)].\end{aligned}$$

Par la suite, par la stationnarité et l'hypothèse (H3)(ii), on obtient

$$\begin{aligned}Var(Q_n(x)) &\leq \frac{C_0}{n(\mathbb{E}(\Delta_1(x)))^2}[\mathbb{E}(\Delta_1^2(x)(W_2(X_1) - W_2(x))) + W_2(x)\mathbb{E}(\Delta_1^2(x))]\\ &\leq \frac{C_0\mathbb{E}(\Delta_1^2(x))}{n(\mathbb{E}(\Delta_1(x)))^2}[o(1) + W_2(x)].\end{aligned}$$

On utilise le Lemme 2.1.1, nous avons

$$n\phi(h)\, Var(Q_n(x)) = O\left(\frac{M_2}{M_1^2}\frac{W_2(x)}{f_1(x)}\right).$$

L'inégalité de Tchebychev combinée à la condition (3.3) donné alors

$$\mathbb{P}\left((n\phi(h))^{\frac{1}{2}} Q_n(x) > \epsilon(\log\log(n))^{\frac{1}{2}}\right) = O\left(\frac{1}{(\log\log(n))}\right) \to 0 \quad \text{quand}\, n \to \infty.$$

Cela signifie que, quand $n \to \infty$,

$$Q_n(x) = o_{\mathbb{P}}\left(\left(\frac{\log\log(n)}{n\phi(h)}\right)^{\frac{1}{2}}\right).$$

Ceci complète la preuve de la partie (a).
La preuve de la partie (b) suit facilement de la première partie du Lemme 3.3.2 et la condition (3.3). ■

Preuve du Théorème 3.3.1. Par la décomposition (3.12), la première partie du Théorème 3.3.1 suit en utilisant le Lemme 3.3.1, la seconde partie du Lemme 3.3.2 et ensuite le Lemme 3.3.3. Si en plus nous utilisons la première partie du Lemme 3.3.2, alors nous obtenons la deuxième partie du Théorème 3.3.1. ■

Le lemme suivant donne la consistance de la probabilité des estimateurs $\hat{r}_n$, $\hat{\mathrm{g}}_n$ et $W_{2,n}$, qui est nécessaire pour prouver le Corollaire 3.3.1.

Lemme 3.3.4. *Supposons que les hypothèses (H1), (H2), (H3), (H5) et (3.5) sont vraies. Alors nous avons*
(i) $\lim_{n\to\infty} \hat{r}_n(x) \overset{\mathbb{P}}{=} r(x)$, *(ii)* $\lim_{n\to\infty} \hat{\mathrm{g}}_n(x) \overset{\mathbb{P}}{=} \mathrm{g}(x)$ *et (iii)* $\lim_{n\to\infty} W_{2,n}(x) \overset{\mathbb{P}}{=} W_2(x)$.

Preuve du Lemme 3.3.4. La preuve des deux premières déclarations utilise les Lemmes 3.3.1 et 3.3.2 et des arguments similaires à ceux utilisés dans la preuve de la Proposition 3.2.2. L'énoncé (iii) est une conséquence directe des résultats (i)-(ii) et de la décomposition (3.8). ■

Preuve de Corollaire 3.3.1. Observez ceci

$$\frac{M_{n,1}}{\sqrt{M_{2,n}}}\sqrt{\frac{nF_{x,n}(h)}{W_{2,n}(x)}}(\hat{r}_n(x) - r(x)) = \frac{M_{n,1}\sqrt{M_2}}{M_1\sqrt{M_{2,n}}}\sqrt{\frac{nF_{x,n}(h)W_2(x)}{W_{2,n}(x)n\phi(h)f_1(x)}}\frac{M_1}{\sqrt{M_2}}\sqrt{\frac{n\phi(h)f_1(x)}{W_2(x)}} (\hat{r}_n(x) - r(x)).$$

En utilisant le Théorème 3.3.1 (ii), an a

$$\frac{M_1}{\sqrt{M_2}}\sqrt{\frac{n\phi(n)f_1(x)}{W_2(x)}}(\hat{r}_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Par conséquent, nous devons établir que

$$\frac{M_{1,n}\sqrt{M_2}}{M_1\sqrt{M_{2,n}}}\sqrt{\frac{F_{x,n}(h)W_2(x)}{W_{2,n}(x)\phi(h)f_1(x)}} \xrightarrow{\mathbb{P}} 1, \quad \text{quand}\, n \to \infty.$$

De toute évidence, de la consistance de la fonction de répartition empirique et de la décomposition en (H2)(i), pour $0 < u \leq 1$, nous avons

$$\frac{F_{x,n}(h)}{\phi(h)f_1(x)} \xrightarrow{\mathbb{P}} 1, \quad \text{quand}\, n \to \infty.$$

Par conséquent, par l'hypothèse (H2)(iv), s'ensuit que

$$\frac{F_{x,n}(uh)}{F_{x,n}(h)} = \frac{F_{x,n}(uh)}{\phi(uh)f_1(x)} \times \frac{\phi(h)f_1(x)}{F_{x,n}(h)} \times \frac{\phi(uh)}{\phi(h)} \xrightarrow{\mathbb{P}} \tau_0(u), \quad \text{quand}\, n \to \infty.$$

Comme $0 \leq \frac{F_{x,n}(uh)}{F_{x,n}(h)} \leq 1$, d'aprés le théorème de convergence dominée, lorsque $(K)'$ est intégrable, nous avons

$$M_{1,n} \xrightarrow{\mathbb{P}} M_1, \quad \text{quand}\, n \to \infty.$$

De même, lorsque $(K^2)'$ est intégrable, nous avons

$$M_{2,n} \xrightarrow{\mathbb{P}} M_2, \quad \text{quand}\, n \to \infty.$$

De plus, la consistance de $W_{n,2}$ établi dans le Lemme 3.3.4 donne

$$\frac{W_{n,2}(x)}{W_2(x)} \xrightarrow{\mathbb{P}} 1, \quad \text{quand}\, n \to \infty.$$

De toute évidence, la preuve de Corollaire 3.3.1 est terminée . ■

Conclusion

Dans ce mémoire, nous nous sommes intéressés à l'estimation non paramétrique, par la méthode du noyau, de la fonction de régression associés à des processus stationnaires et ergodiques. La motivation essentielle est d'établir des propriétés asymptotiques tout en considérant un cadre de dépendance des données assez général qui puisse être facilement utilisé en pratique.

Plus précisément, nous avons établi des propriétés asymptotiques de consistance presque sûre, ponctuelle et uniforme, avec des vitesses de convergence ainsi que la normalité asymptotique de cet estimateur. Nous avons obtenu nos résultats, en faisant usage d'une approche basée sur l'utilisation des différence de martingales en liaison avec des inégalités exponentielles.

Bibliographie

[1] Andrews D.W.K. (1984). Nonstrong mixing autoregressive processes. J. Appl. Probab. 21, p. 930-934.

[2] Bosq D. (1998). Nonparametric statistics for stochastic processes : Estimation and prediction. Springer New york.

[3] Chaouch, M., Laib, N. and Louani, D. (2016). Rate of uniform consistency for a class of mode regression on functional stationary ergodic data. To appear Statistical Methods & Applications, DOI : 10.1007/s10260-016-0356-9.

[4] Collomb,G. (1985). Nonparametric regression : an up-to-date bibliography. Statistics. A Journal of Theoretical and Applied Statistics. 16, p. 309-324.

[5] Delecroix, M. and Rosa, A. C. (1996). Nonparametric estimation of a regression function and its derivatives under an ergodic hypothesis. J. Nonparametr. Statist., 6, p. 367-382

[6] Deville, J.C.(1974). Méthodes statistiques et numériques de l'analyse harmonique. Ann. Insee, 15, 3-101.

[7] Devroye, L.P. (1978). The uniform convergence of the Nadaraya-Watson regression function estimate. Canad. J. Statist., 6, p. 179-191.

[8] Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis theory and practice. Springer-Verlag.

[9] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis. Theory and Practice, in : Springer Series in Statistics, Springer, New York, 2006.

[10] F. Ferraty, A. Mas, P. Vieu, Nonparametric regression of functional data : inference and practical aspects, Aust. N. Z. J. Stat. 49 (2007) 267-286.

[11] Gasser, T. and Müller, H. (1984). Estimating regression functions and their derivatives by the kernel method. Scand. J. Statist. 11, p. 171-185.

[12] Holmström. (1963). On a method for parametric representation of the state of the atmosphere. Tellus, 15, 127-149.

[13] Krengel, U. (1985). Ergodic theorems, Walter de Gruyter & Co. Berlin.

[14] Krikpatrick, M. and Heckman, N. (1989). A quatitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. J. Math. Biol., 27, 429-450.

[15] Laib, N. (1999). Exponential-type inequalities for martingale difference sequences. Application to nonparametric regression estimation. Comm. Statist. Theory Methods, 28(7), 1565-1576.

[16] Laib, N. and Louani, D. (2011). Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. J. Statist. Plann. Inference, 141(1), 359-372.

[17] La Pena, V.H.,Giné, E.,1999. Decoupling from Dependence to Independence, Probability and its Applications.Springer-Verlag,NewYork.

[18] L. Györfi, G. Morvai, S. Yakowitz, Limits to consistent on line forecasting for ergodic time series, IEEE Trans. Inform. Theory 44 (1998) 868-892.

[19] Mack, Y. P. and Silverman, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates. Z. Wahrsch. Verw. Gebiete. 61,p. 405-415.

[20] M. Ezzahrioui, E. Ould-Saïd, Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data, J. Nonparametr. Stat. 20 (2008) 3-18.

[21] Nadaraya, E.N. (1964). On a regression estimate, Teor. Verojatnost. i Primenen. 9, p 157-159.

[22] Nadaraya, E.N. (1965). On nonparametric estimates of density functions and regression curves. Teor. Verojatnost. i Primenen., 10, p. 199-203.

[23] Newey, W. K. (1994). Kernel estimation of partial means and a general variance estimator. Econometric Theor, 10(2), p. 233-253.

[24] N. Laïb, D. Louani, Strong consistency of the regression function estimator for functional stationary ergodic data, J. Statist. Plann. Inference (2010), in press.

[25] N. Laib, Kernel estimates of the mean and the volatility functions in a nonlinear autoregressive model with ARCH errors, J. Statist. Plann. Inference 134 (2005) 116-139.

[26] Obhukov, A. (1960). The Statistically Orthogonal Expansion of Empirical Functions. Bull. Acad. Sci. USSR, Geophys. Ser., 3, 288-291.

[27] P. Hall, C. Heyde, Martingale Limit Theory and its Application, Academic Press, New York, 1980.

[28] Ramsay, J.O. and Silverman, B.W. (2005). Functional data analysis, Second Ed. Springer, New-York.

[29] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, Ann. Math. Statist. 27, p. 832-837,

[30] Sultana Didi. Quelques propriétés asymptotiques en estimation non paramétrique de fonction- nelles de processus stationnaires en temps continu. Statistiques [math.ST]. Université Pierre et Marie Curie - Paris VI, 2014. Francais. < NNT : 2014PA066191

[31] Watson, G. S., (1964) Smooth regression analysis 26 of Sankhy¯a Ser. A, Sankhy ¯a (Statistics). The Indian Journal of Statistics. Series A, p. 359-372.

[32] Yakowitz, S., Györfi, L., Kieffer, J. and Morvai, G. (1999). Strongly consistent nonparametric forecasting and regression for stationary ergodic sequences. J. Multivariate Anal., 71, p. 24-41.

[33] Y.S. Chow, H. Teicher, Probabilty Theory, 2nd ed., Springer, New York, 1998.