République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année: 2015

M-estimation non paramétrique de la fonction de régression

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université Dr Tahar Moulay - Saïda

Discipline : MATHEMATIQUES

Spécialité : Probabilité et Applications

par

Ben Sahraoui Bachir [1]

Soutenu le 15 Juin 2015 devant le jury composé de

F. Mokhtari	Université Dr Tahar Moulay - Saïda	Président
R. Rouane	Université Dr Tahar Moulay - Saïda	Rapporteur
S. Rahmani	Université Dr Tahar Moulay - Saïda	Examinateur
F. Benziadi	Université Dr Tahar Moulay - Saïda	Examinateur

1. e-mail : bachir111333@gmail.com

Table des matières

REMERCIEMENTS

A l'echeance de ce projet de fin d'étude. Je doit en premier lieu l'enorme remerciement à notre dieu mesicorde, aussi que mes très chers parents pour leurs patience et le grand sacrifice pour l'aboutissement de nos études.

Aussi mes remerciement les plus sincères à notre encadreur Rouane Rachida. qui a été toujours fidèle , qui nous a aidé du debut à la fin pour finir ce travail.

Aussi au jury à leur tète le president et les examinateurs pour avoir honorer de leurs presences notre soutenance.

Sans oublier tous les professeurs, aussi que toutes les personnes qui ont contribués de près ou de loin pour mener à bien notre travail.

Introduction

L'estimation statistique est une ancienne branche de la statistique qui se modernise souvent par le renouvellement et l'importance pratique des applications qui en relèvent. Cette théorie est divisée en deux composantes principales, à savoir, l'estimation paramétrique et l'estimation non-paramétrique. L'approche non paramétrique est fondée sur l'idée de ne pas faire des hypothèses sur la distribution de l'échantillon d'observations. Ceci, contrairement à l'approche paramétrique qui suppose que la distribution suit un certain modèle décrit par un nombre fini de paramètres. Ainsi, le modèle non paramétrique offre une applicabilité plus large que le cas paramétrique. A ce stade, le modèle le plus fréquemment rencontré en statistique non paramétrique est le modèle de régression. L'utilité de ce modèle est notamment dans la prévision pour décrire la relation entre deux ou plusieurs variables aléatoires. Notons que, l'avantage principal de la régression non paramétrique est qu'elle ne suppose aucune forme spécifique pour l'estimateur, ce qui lui donne plus de flexibilité en pratique. En effet, ce modèle est utilisée pour étudier la relation entre deux variables lorsque le modèle linéaire ne s'applique pas, ou pour suggérer la forme que devrait prendre un modèle de régression paramétrique. Cependant, un des principaux inconvénients de régression classique est que l'estimation de la fonction de régression est sensible aux valeurs aberrantes, et peut-être insuffisant dans certains cas, comme lorsque la distribution est asymétrique ou multimodal. A ce sujet, nous oblige de chercher des approche alternative qui soient suffisamment insensibles aux effets de donnes aberrantes. La régression robuste a été introduite pour résoudre ce genre de problèmes.

Du point de vie historique, la statistique robuste ont été initialement développé par Huber (1964), dont il a obtenu la consistance et la normalité asymptotique d'estimateurs de classe M- estimateur pour la fonction da régression. Härdle (1984) établie la consistance faible et forte, ainsi que la normalité asymptotique d'une M-estimateur de la fonction de régression lorsque les observations sont indépendantes et identiquement distribuées. De nombreux auteurs ont ajouté des résultats sur ce sujet en considérant des observations dépendantes.

Objectifs

Le travail pésenté dans ce mémoire a pour objectifs d'etudier le modèle de la régression non paramétrique dans le cas où la variable explicative est fonctionnelle, en utilisant l'approche robuste.

La robustesse d'une procédure statistique usuelle (estimation, test) est une question très importante en statistique. Elle permet de contrôler la stabilité de cette procédure relativement à la déviation du modèle et/ou des observations. Notons que ce problème a fait l'objet d'un long débat à la fin du XIX siècle, plusieurs scientifiques avaient déjà une idée relativement claire de cette notion de robustesse.

Organisation de mémoire

Nous présentons notre travail dans trois chapitres.

Dans le premier chapitre nous commençons, tout d'abord par présenter le modèle de la régression non paramétrique robuste avec une variable d'intérêt réelle et une variable explicative fonctionnelle, par la suite, nous avons rassemblé le bagage nécessaire pour effectuer une étude de ce modèle fonctionneles, tout en donnont quelques notions et définitions et outils utilisés le long de cette mémoire.

Dans le deuxième chapitre Nous rappelons quelques résultats concernant l'estimation non paramétrique de la fonction de régression, en utilisant l'approche robuste. Nous établissons la convergence presque complète et la normalité asymptotique d'un estimateur à noyau.

Dans le chapitre trois, on établit également à partir du résultat de normalité asymptotique et d'un résultat d'uniforme intégrabilité l'expression explicite des termes asymptotiquement des erreurs $\mathbb{L}_q$ de notre estimateur.

Chapitre 1

Présentation

1.1 Présentation de l'estimateur

Soit un couple (X, Y) de variables aléatoires à valeur dans $\mathcal{F} \times \mathbb{R}$, où $\mathcal{F}$ est un espace semi-métrique. On note d la semi-métrique sur $\mathcal{F}$. Pour $x \in \mathcal{F}$, on consédére un fonction mesurable ψ_x. Le paramètre fonctionnel étudie dans ce travail, noté θ_x, est la solution de léqution en t définie par

$$\Psi(x, t) := \mathbb{E}\left[\psi_x(Y_i, t) \mid X_i = x\right] = 0.$$

On suppose que cette équation admet θ_x comme solution unique (voir, par exemple, Boente et Fraiman (1989)) pour des conditions suffisantes sur l'existence et l'unicité de θ. Le paramètre θ_x, appelé ψ-régression dans Laïb et Ould-Saïd (2000), est une généralisation de la fonction de régression classique. En effet, il suffit de prendre $\psi_x(t) = t$ pour retrouver la fonction de régression.

Etant donné un échantilion $(X_i, Y_i)_{i=1,\ldots,n}$ de méme loi que (X, Y), un estimateur à noyau de $\Psi(x, t)$ est donné par

$$\widehat{\Psi}(x, t) := \frac{\sum_{i=1}^{n} K(h^{-1}d(X_i, x))\psi_x(Y_i, t)}{\sum_{i=1}^{n} K(h^{-1}d(X_i, x))}, \quad \forall t \in \mathbb{R}$$

Où K est un noyau et $h = h_n$ est une suite de nombres réels positifs. L'estimateur naturel du paramètre θ_x, noté $\widehat{\theta_x}$, est tel que

$$\widehat{\Psi}(x,t) = 0. \tag{1.1}$$

On remarque que, sous la condition $\sum_{i=1}^{n} K(h^{-1}d(X_i,x)) \neq 0$, la définition de l'estimateurpar 1.1 est équivalente à

$$\hat{\rho}_n(x,t) = \sum_{i=1}^{n} K(h^{-1}d(X_i,x))\psi_x(Y_i,t) = 0 \tag{1.2}$$

1.2 Notations et définitions

Dans cette section, nous rappelons quelques définitions et théorèmes utilisés le long de cette mémoire.

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

Définition 1.2.1. *On dit que la famille $\{\Delta_i, i \in \mathcal{Z}\}$ est φ-mélangent (resp. ρ-mélangent) si la suite*

$$\varphi_n = \sup_{k\in\mathbb{Z}} \sup_{A\epsilon\sigma_{-\infty}^k, B\epsilon\sigma_{n+k}^{\infty}} |P(B/A) - P(B)|$$

$$\rho_n = \sup_{k\in\mathbb{Z}} \sup_{\{X\in L_2(\sigma_{-\infty}^k), Y\in L_2(\sigma_{n+k}^{\infty})\}} \frac{E[(X-EX)(Y-EY)]}{[Var(X)Var(Y)]^2}$$

tend vers 0 quand n tend vers l'infinie.

Définition 1.2.2. *On dit que la famille* $\{\Delta_i, i \in \mathcal{Z}\}$ *est* α*- mélangent si la suite*

$$\alpha(n) = \sup_{\left\{k\in\mathcal{Z}, A\in\sigma^k_{-\infty}, B\in\sigma^{+\infty}_{n+k}\right\}} |P(A\cap B) - P(A)P(B)|$$

tend vers 0 *quand* n *tend vers l'infinie. La suite* α_n *est appelée coefficient de mèlange forte.*

La relation entre ces trois types des processus est donnée par le lemme suivant :

Lemme 1.2.0.1. *Soit* $\{\Delta_i, i \in \mathcal{Z}\}$ *une famille des variables aléatoires définie sur* $(\Omega, \mathcal{A}, P)$ *à valeurs dans un espace probabilisable* $(\mathbb{E}, \xi)$*. Pour tout* $n \in \mathbb{N}$*, on a,*

$$\alpha_n \leq \rho_n \leq 2\varphi_n^{1/2}.$$

En vertu de ce lemme on peut dire que :

$$\begin{aligned} \text{le processus } \varphi - \text{mélange } &\Rightarrow \text{le processus } \rho - \text{mélange} \\ &\Rightarrow \text{le processus } \alpha - \text{mélange} \end{aligned}$$

Autrement dit, le processus α -mélangeant est le plut fort. Ainsi, nous allons modélisé la notion de la dépendance par ce processus α-mélangeant, car elle englobe les autres processus définies ci-dessus.

1.2.1 Convergence presque complète

Définition 1.2.3. *On dit que la suite* $(X_n)_{n\in\mathbb{N}}$ *converge presque complètement vers* X *si*

$$\forall \epsilon > 0, \qquad \sum_{n=0}^{\infty} P(|X_n - X| > \epsilon) < \infty$$

et on note $X_n \longrightarrow X$ *en p.co.*

Définition 1.2.4. *On dit que la suite* $X_n = O(Y_n)$ *en p.co. s'il existe un* $\epsilon > 0$ *vérifiant :*

$$\sum_{n=0}^{\infty} P(X_n > \epsilon Y_n) < \infty.$$

Définition 1.2.5. *Soient* X_n, Y_n *deux suites des variables aléatoires. La suite*

$$\left(\frac{X_n}{Y_n}\right)_{n\in\mathbb{N}} \longrightarrow 0 \text{ en p.co si}$$

$$X_n \longrightarrow 0 \quad \text{en p.co.}$$

et

$$\exists \delta > 0, \qquad \sum_{n=0}^{\infty} P(|Y_n| < \delta) < \infty.$$

1.3 Outils

1.3.1 Inégalité de type Bernstein

Il existe plusieurs versions d'inégalités de ce type. Nous nous contentons de rappeler dans le lemme ci dessous une version simplifie, qui nous suffit dans ce travail et dont la preuve est donnée dans l'article de Hoeffding (1963) :

Lemme 1.3.1.1. *Soit* $(\Delta_n)_{n\in\mathbb{N}}$ *une suite des variables aléatoires réelles centrées, indépendantes et identiquement distribuées, telles qu'il existe deux réelles positifs* d *et* δ *vérifiant :*

$$|\Delta_1| \le d \text{ et } \quad E\Delta_1^2 \le \delta^2,$$

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

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

Le lemme suivant donne l'inégalité de Fuk Nagave

Lemme 1.3.1.2. *Soit* $\{\Delta_i, i \epsilon \mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeante, de coefficient de mélange* α_n *vérifiant :*

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

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

$$P\left[|\sum_{k=1}^{n} \Delta_k| > 4\varepsilon\right] \leq \left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{2r}{\varepsilon}\right)^{a+1} \tag{1.3}$$

où

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

L'inégalité du lemme suivant s'appelle inégalité de covariance et elle est trés utile pour le calcule de S_n^2, définie dans le lemme précèdent

Lemme 1.3.1.3. *Soit* $\{\Delta_i, i \epsilon \mathbb{N}\}$ *une suite des variables aléatoires réelle* α*-mélangeante, de coefficient des mélange* α_n*, telle que* $\|\Delta_i\|_\infty < \infty, \forall i$*. On a, pour tout* $i \neq j$ *:*

$$|Cov(\Delta_i, \Delta_j)| \leq 4\|\Delta_i\|_\infty \|\Delta_j\|_\infty \alpha|i - j|.$$

L'inégalité de Marcinkiewicz-Zygmund est donnée par le lemme suivant

Lemme 1.3.1.4. *Soit* $\{X_n, n \geq 1\}$ *une suite des variables aléatoires indépendants telle que* $\mathbb{E}X_n = 0$*, alors, pour tout* $p \geq 1$*, il existe deux constants positive* A_p *,* B_p *telle que*

$$A_p \left\| \left(\sum_{j=1}^{n} X_j^2\right)^{1/2} \right\|_p \leq \left\| \sum_{j=1}^{n} X_j \right\|_p \leq B_p \left\| \left(\sum_{j=1}^{n} X_j^2\right)^{1/2} \right\|_p .$$

Théorème 1.1. *(Inégalité de Minkowski) Soit X et Y des variables aléatoires telles que X et $Y \in L^p(\Omega, \mathcal{A}, P)$, avec $p \geq 1$. Alors, $X + Y \in L^p(\Omega, \mathcal{A}, P)$ et*

$$(\mathbb{E}|X+Y|^p)^{1/p} \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}(|Y|^p))^{1/p}.$$

L'inégalité de Chauchy-Shwartz est donnée par le lemme suivant

Lemme 1.3.1.5. *Soit $\{X_n, n \geq 1\}$ et $\{Y_n, n \geq 1\}$ deux suites des variables aléatoires indépendants telle que*

$$\sum_{i=1}^{n} X_i Y_i \leq \left(\sum_{i=1}^{n} X_i^2\right)^{1/2} \left(\sum_{i=1}^{n} Y_i^2\right)^{1/2}.$$

Chapitre 2

M-estimation non paramétrique de la fonction de régression : Convergence presque complète et Normalité Asymptotique

Nous rappelons dans ce chapitre quelques résultats concernant l'estimation non paramétrique de la fonction de régression, en utilisant l'approche robuste. Nous établissons la convergence presque complète et la normalité asymptotique d'un estimateur à noyau. Ces propriétés asymptotiques sont obtenues sous certaines conditions d'indépendance et de dépendance faible (mélange fort).

2.1 Cas i.i.d

Dans cette partie les observations seront considérées indépendantes.

2.1.1 Convergence presque complète

Hypothèses

Les hypothèses necessaires pour établir la convergence presque complète sont :

On fixe un point x dans $\mathcal{F}$, N_x désigne un quartier fixe de x, et nous présenter le les hypothèses suivantes

- (H1) $\mathbb{P}(X \in B(x,h)) = \phi_x(h) > 0 \quad \forall h > 0$ est $\lim_{h \longrightarrow 0} \phi_x(h) = 0$
- (H2) Il existe $C_1 > 0$ est $b > 0$ de sorte que $\forall x_1, x_2 \in N_x$, $\forall t \in \mathbb{R}$
$$| \Psi(t,x_1) - \Psi(t,x_2) | \leq C_1 d^b(x_1,x_2).$$
- (H3) La fonction ψ_x est strictement monotone, bornée, continûment différentiable, et sa dérivée vérifie,
$$|\psi_x'(t)| > C_2 > 0, \forall t \in \mathbb{R}$$
- (H4) K est une fonction continue à intervale $[0,1]$ telle que
$$0 < C_3 < K(t) < C_4 >< \infty.$$
- (H5) $\lim_{n\to\infty} h_K = 0$ et $\lim_{n\to\infty} \dfrac{logn}{n\phi_x(h_K)} = 0$

Notre principal résultat est donnée dans le théorème suivant

Théorème 2.1.1. *Sous les hypothèses $(H1) - (H5)$, l'estimateur $\widehat{\theta}_x$ existe et est unique pour n assez grand, et*

$$\widehat{\theta}_x - \theta_x = O(h_K^b) + O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \quad p.co. \tag{2.1}$$

Preuve

Dans ce qui suit, on note par C une constante strictement positive et $K_i = K\left(\frac{d(x,X_i)}{h_K}\right)$. En vertu de $(H3)$, nous avons

$$\hat{\Psi}(\hat{\theta}_x, x) = \hat{\Psi}(\theta_x, x) + (\hat{\theta}_x - \theta_x)\hat{\Psi}'(\xi_{x,n}, x)$$

pour certains $\xi_{n,x}$ entre $\hat{\theta}_x$ et θ_x La condition sur la dérivée de ψ_x dans (H3), nous conduit à écrire

$$\exists C > 0, \epsilon_0 > 0 \mathbb{P}(|\hat{\theta}_x - \theta_x| \geq \epsilon_0 \left(h^b + \sqrt{\frac{logn}{n\phi_x(h)}}) \right)$$

$$\leq \mathbb{P}(|\hat{\Psi}(\hat{\theta}_x, x) - \Psi(\theta_x, x)| \geq C^{-1}\epsilon_0 \left(h^b + \sqrt{\frac{logn}{n\phi_x(h)}}) \right).$$

Alors, il suffit de montrer que :

$$\hat{\Psi}(\hat{\theta}_x, x) - \Psi(\theta_x, x) = 0 \left(h_K^b + \sqrt{\frac{logn}{n\phi_x(h_K)}} \right) \quad p.co \tag{2.2}$$

La preuve de 2.2 est basé sur la décomposition suivante :

$$\forall t \in \mathbb{R}, \hat{\Psi}(t,x) - \Psi(t,x) = \frac{1}{\hat{\Psi}_D(x)}[(\hat{\Psi}_N(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)] - (\Psi(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)])]$$

$$-\frac{\Psi(t,x)}{\hat{\Psi}(t,x)}[\hat{\Psi}_D(x) - \mathbb{E}[\hat{\Psi}_D(x)]]$$

où

$$\hat{\Psi}_D(x) = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} K_i, \quad \hat{\Psi}_N(t,x) = \frac{1}{n\mathbb{E}(K_1)} \sum K_i \Psi_x(Y_i - t)$$

Il est clair que $\hat{\Psi}(t,x) = \frac{\Psi_N(t,x)}{\hat{\Psi}_D(x)}$ et $\mathbb{E}[\hat{\Psi}_D(x)] = 1$.

Finalement, la preuve du théorème 2.1.1 est réalisé avec les lemmes suivants

Lemme 2.1. *Sous des hypothèses (H1), (H4) et (H5), on a,*

$$\hat{\Psi}_D(x) - \mathbb{E}[\hat{\Psi}_D(x)] = O\left(\sqrt{\frac{logn}{n\phi_x(h_K)}} \right) p.co$$

Preuve La preuve de ce lemme longe les lignes du lemme 3.1 dans Ferraty et al . (2005). Soit $\tilde{\Delta}_i = \dfrac{K_i}{\mathbb{E}[K_1]}$ de (H1) et (H4) nous déduisons $|\tilde{\Delta}_i| < C/\phi_x(h_K)$ et $\mathbb{E}[|\tilde{\Delta}_i|^2] < C'/\phi_x(h_K)$ Donc , nous appliquons l'inégalité exponentielle Bernstein pour obtenir pour tous $\eta > 0$

$$\mathbb{P}\left(|\hat{\Psi}_D(x) - \mathbb{E}[(\hat{\Psi}_D(x)]| > \eta\sqrt{\frac{logn}{n\phi}_x(h_K)}\right) \leq C'n^{-c\eta^2}$$

Ce lemme donne carrément le corollaire suivant

Corollaire 2.1.1. *Sous les hypothèses du lemme 2.1, on a*

$$\sum_{n\geq 1}\mathbb{P}(|\hat{\Psi}_D(x)| \leq 1/2) \leq \sum_{n\geq 1}\mathbb{P}(|\hat{\Psi}_D(x) - 1| > 1/2) < \infty.$$

Preuve corollaire 2.1.1 On a $\mathbb{E}[\hat{\Psi}_D(x)] = 1$. Alors

$$\mathbb{P}(|\hat{\Psi}_D(x)| \leq 1/2) \leq \mathbb{P}(|\hat{\Psi}_D(x) - 1| > 1/2) \leq \mathbb{P}(|\hat{\Psi}_D(x) - \mathbb{E}[\hat{\Psi}_D(x)]| > 1/2)$$

Le résultat est une conséquence du lemme précédent.

Lemme 2.2. *Sous des hypothèses (H1),(H2),(H4) et (H5) , on a pour tout* $t \in \mathbb{R}$

$$\Psi(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)] = O(h_K^b).$$

Preuve

Le équidistribution des couples (X_i, Y_i) et $(H4)$ impliquent

$$\Psi(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)] = \frac{1}{\mathbb{E}(K_1)}\mathbb{E}[(K_1 1\!\!1_{B(x,h_K)}(X_1))(\Psi(t,x) - \mathbb{E}[\Psi(Y_1 - t)\backslash X = X_1])]$$

où $1\!\!1$ est fonction de l'indicateur. L'hypothès (H2) nous permet d'écrire que

$$K_1 1\!\!1_{B(x,h_K)}(X_1)|\Psi(t,X_1) - \Psi(t,x)| \leq C_1 h_K^b$$

puis

$$|\Psi(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)]| \leq C_1 h_K^b$$

Lemme 2.3. *Sous des hypothèses (H1) et (H3) -(H5), on a pour tous $t \in \mathbb{R}$*

$$\hat{\Psi}_N(t,x) - \mathbb{E}[\hat{\Psi}_N(t,x)] = O\left(\sqrt{\frac{logn}{n\phi_x(h_K)}}\right) p.c$$

Preuve La preuve de ce résultat est similaire à la preuve du lemme 2.1. On prend

$$\Lambda_i = \frac{K_i\Psi_x(Y_i - t) - \mathbb{E}[K_1\Psi_x(Y_1 - t)]}{\mathbb{E}[K_1]}$$

comme Ψ_x est bornée, alors $|\Lambda_i| \leq \acute{C}\phi_x(h_K)$ et $\mathbb{E}[\Lambda_i^2] \leq \acute{C}\phi_x(h_K)$, pour tout $i \leq n$. Comme dans le lemme 2.1, il suffit d'appliquer l'inégalité de Bernstein pour obtenir le resultat.

Lemme 2.4. *Sous des hypothèses du théorème 2.1.1, $\hat{\theta}_x$ existe et est unique presque sûrement pour n suffisamment grand.*

Preuve : Nous prouvons ce lemme par des arguments similaires à ceux utilisés dans le théorème 1 de Collomb et Hardle (1986). En effet, pour tout $\epsilon > 0$, la monotonie stricte de ψ_x implique

$$\Psi(\theta_x - \epsilon, x) < \Psi(\theta_x, x) < \Psi(\theta_x + \epsilon, x)$$

Les lemmes 2.1, 2.2, 2.3 et le corollaire 2.1.1 montrent que

$$\hat{\Psi}(\theta_x, x) - \Psi(\theta_x, x) = O\left(h_K^b + \sqrt{\frac{logn}{n\phi_x(h_K)}}\right) p.co$$

pour tout t fixé, alors, pour n suffisament grand,

$$\hat{\Psi}(\theta_x - \epsilon, x) < 0 < \hat{\Psi}(\theta_x + \epsilon, x) \quad p, co$$

Puisque Ψ_x et K sont des fonctions continues alors $\hat{\Psi}(t,x)$ est une fonction continues, alors il existe $t_0 = \hat{\theta}(x) \in [\theta_x - \epsilon, \theta_x + \epsilon]$ tel que $\hat{\Psi}(\hat{\theta}_x, x) = 0$.

Enfin, l'unicité de $\hat{\theta}_x$ est une conséquence directe de la monotonie stricte de ψ_x et la positivité de K.

2.1.2 Normalité asymptotique

Hypothèses et résults

On suppose que $\lambda_\gamma(u,t) = \mathbb{E}[\psi^\gamma(Y,t)/X = u]$ et $\Gamma_\gamma(u,t) = \mathbb{E}[\acute{\psi}_\gamma(Y,t)/X = u]$, pour $\gamma \in [1,2]$, on gardant les même notations que la section précédent, la même hypothèse $(H3)$ et on remplace (H1), (H2), (H4) et (H5) par :

(H1')Il existe une une fonction différentiable non négatif ϕ et une fonction non négative g telle que :

$$\mathbb{P}(X \in B(x,r)) = \phi_r.g(x) + o\phi(r) \quad \text{où} \quad B(x,r) = \acute{x} \in \mathrm{F}/d(x,\acute{x}) < r$$

(H2') i) La fonction $\lambda_\gamma(.,.)$ satisfait la condition de Lipschitz par rapport à la première composante, qui est la suivante : il existe une constante b_γ strictement positive tel que :

$$\forall (u_1,u_2) \in N_x \times N_x, \forall t \in \mathbb{R}, |\lambda_\gamma(u_1,t) - \lambda_\gamma(u_2,t)| \leq C_1 d(u_1,u_2)^{b\gamma}$$

ii) La fonction $\Gamma_\gamma(.,.)$ satisfait la condition de Lipschitz par rapport à la première composante, qui est la suivante : il existe une constante d_γ strictement positive de telle sorte que :

$$\forall (u_1,u_2) \in N_x \times N_x, \forall t \in \mathbb{R}, |\Gamma_\gamma(u_1,t) - \Gamma_\gamma(u_2,t)| \leq C_2 d(u_1,u_2)^{d\gamma}$$

(H4') Le paramètre de lissage h satisfait :

$$h \downarrow 0, \forall t \in [0,1] \lim_{h \longrightarrow 0} \frac{\phi(th)}{\phi(h)} = \beta(t) \quad \text{et} \quad n\phi(h) \longrightarrow \infty \quad \text{quand} \quad n \longrightarrow \infty$$

(H6) Le noyau K est une fonction différentiable sur le support $[0,1]$, sa dérivée K' existe et tels qu'il existe deux constantes C_3 et C_4 avec, $-\infty < C_3 < K'(t) < C_4 < 0$ pour$0 \leq t \leq 1$.

Théorème 2.1. *Sous les hypothèse (H1') - (H2'), (H3) et (H4')-(H5'),* $\hat{\theta}_x$ *existe et est unique avec une probabilité tend vers 1, et pour tout* $x \in A$, *on a*

$$\left(\frac{n\phi(h)}{\sigma^2(x,\theta_x)}\right)^{1/2}(\hat{\theta}_x - \theta_x - B_n(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \textit{quandn} \longrightarrow \infty$$

où

$$B_n(x) = \frac{h}{\phi(h)\alpha_1\Gamma_1(x,\theta_x)} \int_0^1 K(t)\varphi_x(th) + \phi'(th) + o(1)$$

(avec $\varphi_x(s) = \mathbb{E}[\psi(Y,\theta_x)(X,x) = s])$,

$$\sigma^2(x,\theta_x) = \frac{\alpha_2\lambda_2(x,\theta_x)}{\alpha_1^2 g(x)(\Gamma_1(x;\theta_x))^2}$$

(avec $\alpha_j = -\int_0^1 K_j'(s)\beta(s)ds, pour, (j=1,2)$*)*

$$\mathcal{A} = \{x \in \mathcal{F}, g(x)\lambda_2(x,\theta_x)\Gamma_1(x,\theta_x) \neq 0\}$$

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

Corollaire 2.1.2. *Sous les hypothèses du théorème 2.1 et si les paramètres de bande passante h satisfait*

$nh^{2b_1}\phi(h) \longrightarrow 0$ *pour,* $n \longrightarrow \infty$ *nous avons*

$$\left(\frac{n\phi(h)}{\sigma^2(x,\theta_x)}\right)^{1/2}(\hat{\theta}_x - \theta_x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \textit{pour} \quad n \longrightarrow \infty$$

Preuve . Pour $i = 1, ..., n$, nous considérons quantités

$$K_i = K(h^{-1}d(x,X_i)), \psi_i(t) = \psi(Y_i,t)$$

et

$$\hat{\Psi}_N(x,t) = \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} K_i\psi_i(t)$$

On utilise le développement de Taylor d'ordre un autour de θ_x , nous obtenons

$$\hat{\Psi}_N(x,\hat{\theta}_x) = \hat{\Psi}_N(x,\theta_x) + (\hat{\theta}_x - \theta_x)\hat{\Psi}'(x,\xi_n)$$

avec $\xi_n \in (\hat{\theta}_x, \theta_x)$. En raison de la définition $\hat{\theta}_x$, on a

$$\hat{\theta}_x - \theta_x = \frac{-\hat{\Psi}_N(x,\theta_x)}{\hat{\acute{\Psi}}_N(n,\xi_n)}$$

Enfin, nous avons la décomposition suivante :

$$\sqrt{n\phi(h)}(\hat{\theta}_x - \theta_x) = \frac{-\sqrt{n\phi(h)}(\hat{\Psi}(x,\theta_x) - \mathbb{E}[\hat{\Psi}_N(x,\theta_x]))}{\hat{\acute{\Psi}}_N(x,\theta_x)} - \frac{\sqrt{n\phi(h)}\mathbb{E}[\hat{\Psi}_N(x,\theta_x)]}{\hat{\acute{\Psi}}\Psi_N(x,\theta_x)}$$

Ensuite, la preuve de théorème 2.1 est basée sur les lemmes suivants.

Lemme 2.5. *Sous des hypothèses (H1') - (H2'), (H3) et (H4')-(H5'), on a pour tout $x \in \mathcal{A}$*

$$\left(\frac{n\phi(h)\alpha_1^2 g(x)}{\alpha_2\lambda_2(x,\theta_x)}\right)^{1/2}(\hat{\Psi}_N(x,\theta_x) - \mathbb{E}[\hat{\Psi}_N(x,\theta_x])) \longrightarrow \mathcal{N}(0,1) pour n \longrightarrow \infty$$

Preuve :La preuve de ce lemme est basé sur la version du théorème central limite donnée dans Loève (1963, p. 275) où le point principal est de calculer la limite suivante

$$\forall t \in \mathbb{R}, n\phi(h)Var[\hat{\Psi}_N(x,t)] \longrightarrow \frac{\alpha_2\lambda_2(x,t)}{\alpha_1^2 g(x)}$$

Par définition de $\hat{\Psi}_N(x,t)$, on a

$$\begin{aligned}
Var[\hat{\Psi}_N(x,t)] &= \frac{1}{(n\mathbb{E}[K_1])^2}\sum_{i=1}^{n} Var[K_i\psi_i(t)] \\
&= \frac{1}{n\mathbb{E}^2[K_1]} Var[K_1\psi_1(t)] \\
&= \frac{1}{n\mathbb{E}^2[K_1]\mathbb{E}(K_1^2\psi^2(Y,t))} - \frac{1}{n\mathbb{E}^2[K_1]\mathbb{E}(K_1\psi(Y,t))^2} \\
&= \frac{\mathbb{E}[K_1^2]}{n\mathbb{E}^2[K_1]}\mathbb{E}\left[\frac{K_1^2\psi^2(Y,t)}{\mathbb{E}[K_1^2]}\right] - \frac{1}{n}\left(\mathbb{E}\left[\frac{K_1\psi(Y,t)}{\mathbb{E}[K_1]}\right]\right)^2
\end{aligned}$$

Donc,

$$n\phi(h)Var[\hat{\Psi}_N(x,t)] = \frac{\phi(h)\mathbb{E}[K_1^2]}{\mathbb{E}^2[K_1]}\left(\mathbb{E}\left[\frac{K_1^2(\psi(Y,t))^2}{\mathbb{E}[K_1^2]}\right]\right) - \phi(h)\left(\mathbb{E}\left[\frac{K_1(\psi(Y,t))^2}{\mathbb{E}[K_1]}\right]\right)^2 \quad (2.3)$$

D'une part, nous évaluons la limite du second terme de la partie droite de 2.3. Puisque $\phi(h) \longrightarrow 0$, il suffit maintenant de montrer que

$$\forall t \in \mathbb{R}, \quad \mathbb{E}\left[\frac{K_1\psi(Y,t)}{\mathbb{E}[K_1]}\right] \longrightarrow \lambda_1(x,t) \quad (2.4)$$

En effet, un simple calcul nous donne :

$$\left|\mathbb{E}\left[\frac{k_1\psi(Y,t)}{\mathbb{E}(K1)}\right] - \lambda_1(x,t)\right| = \frac{1}{\mathbb{E}(K_1)}|\mathbb{E}(K_1 1\!\!1_{B(x,h)})(X_1)\times(\lambda_1(X_1,t)-\lambda_1(x,t))|.$$

En outre, par (H3) (avec $\gamma = 1$), on obtient

$$1\!\!1_{B(x,h)}(X_1)|(\lambda_1(X_1,t) - \lambda_1(x,t))| \leq Ch^{b_1},$$

puis

$$\left|\mathbb{E}\left[\frac{K_1\psi(Y,t)}{\mathbb{E}(K1)}\right] - \lambda_1(x,t)\right| \leq Ch^{b_1} \longrightarrow 0$$

et ainsi de 2.3 est vérifiée.
D'autre part, pour le premier terme de la partie droite de 2.3, nous montrons à l'aide analogue arguments à ceux qui sont considérés tirer 2.4, nous allons montrer que

$$\mathbb{E}\left[\frac{K_1^2\psi_x^2(Y,t)}{\mathbb{E}(K_1^2)}\right] \longrightarrow \lambda_2(x,t). \tag{2.5}$$

En effet, par (H2), nous pouvons facilement obtenir

$$|\mathbb{E}\frac{K_1^2\psi^2(Y,t)}{\mathbb{E}[K_1^2]} - \lambda_2(x,t)| = \frac{1}{\mathbb{E}[K_1^2]}|\mathbb{E}[K_1^2 1\!\!1_{B(x,h)}(X_1) \times (\lambda_2(X_1,t) - \lambda_2(x,t))]$$

Finalement, à nouveau par l'hpothèse (H3') (avec $\gamma = 2$) on obtient :

$$\frac{1}{\mathbb{E}[K_1^2]}|\mathbb{E}[K_1^2 1\!\!1_{B(x,h)}(X_1) \times (\lambda_2(X_1,t) - \lambda_2(x,t))]| \leq Ch^{b_2}$$

Ensuite, en intégrant par rapport à la distribution de la variable réelle Z :=d(x,X), nous pouvons montrer que, en vertu de (H1') et (H6') (voir Ferraty et Vieu (2006) p . 44),

$$\forall j > 0 \ \ \mathbb{E}[K_1^j] = -g(x)\int_0^1 \acute{K}^j(t)\phi(th)dt + o(\phi(th)).$$

Il en résulte que

$$\frac{\phi(h)\mathbb{E}[K_1^2]}{\mathbb{E}^2[K_1]} \longrightarrow \frac{\alpha_2}{\alpha_1^2 g(x)} \text{ pour } n \longrightarrow \infty \tag{2.6}$$

Puis, en utilisant 2.4, 2.5(2.13) et 2.6, nous permet d'obtenir,

$$n\phi(h)Var(\hat{\Psi}_N(x,t)) \longrightarrow \frac{\alpha_2\lambda_2(x,t)}{\alpha_1^2 g(x)}. \tag{2.7}$$

Voyons maintenant prouver le résultat suivant. fixé

$$L_i(x) = \frac{1}{\mathbb{E}[K_1]}[K_i\psi_i(\theta_x)].$$

Il suffit de montrer que pour certains $\delta > 0$

$$\frac{\sum \mathbb{E}[|L_i(x) - \mathbb{E}[L_i(x)]|^{2+\delta}]}{(Var(\sum_{i=1}^n L_i(x)))^{(2+\delta)/2}} \longrightarrow 0$$

De toute évidence, les calculs ci-dessus montrent que $n\psi(h)Var(\sum_{i=1}^n Li(x))$ converge vers $\frac{\alpha_2\lambda_2(x,t)}{\alpha_1^2 g(x)}$ lorsque n tend à l'infini. Par conséquen, pour conclure la preuve de ce lemme, il suffit de montrer que le numérateur de l'expression ci-dessus converge vers 0. Pour cela, nous utilisons l'inégalité C_r (voir Loève (1963), p. 155) nous montrons que,

$$\begin{aligned}(n\phi(h))^{(1+\delta/2)}\sum_{i=1}^n \mathbb{E}[|L_i(x) - \mathbb{E}[L_i(x)]|^{2+\delta}] &\leq 2^{1+\delta}(n\phi(h))^{(1+\delta/2)}\sum_{i=1}^n \mathbb{E}[|L_i(x)|^{2+\delta}] \\ &+ 2^{1+\delta}(n\phi(h))^{(1+\delta/2)}\sum_{i=1}^n |\mathbb{E}[L_i(x)|^{2+\delta}].\end{aligned}$$

Observez que, selon (2.14), nous avons, pour tout $j > 0, \mathbb{E}[K_1^j] = O(\phi(h))$, puis, parce que ψ est bornée et (H2), nous avons

$$(n\phi(h))^{(2+\delta)/2}\sum_{i=1}^n \mathbb{E}[|L_i(x)|^{|2+\delta|}] = 2^{1+\delta}n^{-\delta/2}(\phi(h))^{-1-\delta/2}\left(\frac{\phi(h)}{\mathbb{E}[K_1]}\right)^{2+\delta}\mathbb{E}[K_1^{2+\delta}|\psi_1(\theta_x)|^{2+\delta}] \tag{2.8}$$

$$\leq C(n\phi(h))^{-\delta/2}\left(\frac{\phi(h)}{\mathbb{E}[K_1]}\right)^{2+\delta}(\mathbb{E}[K_1^{2+\delta}]/\psi(h)) \longrightarrow 0 \tag{2.9}$$

De même, le second terme de (2.16) est également limité, comme un est bornée, par

$$(n\phi(h))^{(2+\delta)/2}\sum_{i=1}^{n}\mathbb{E}[|L_i(x)|^{|2+\delta|}] \leq 2^{1+\delta}n^{-\delta/2}(\phi(h))^{(2+\delta)/2}\mathbb{E}^{-2-\delta}[K_1]\,|\mathbb{E}[K_1\psi_1(\theta)]|^{2+\delta} \tag{2.10}$$

$$\leq Cn^{-\delta/2}(\phi(h))^{(2+\delta)/2} \longrightarrow 0, \tag{2.11}$$

qui achève la démonstration.

Lemme 2.6. *Sous les hypothèses (H1') et (H5'), on a,*

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x] = \frac{h}{\phi(h)\alpha_1}\int_0^1 K(t)\varphi_x(th)\phi'(th)dt + o(1)$$

Preuve :Nous commençons par écrit

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] = \frac{\mathbb{E}[K_1\psi_1(\theta_x)]}{\mathbb{E}[K_1]}$$

En conditionnant par rapport à la variable réelle d(x,X1), on obtient

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] = \frac{\mathbb{E}[K_1\mathbb{E}[\psi_1(\theta_x)/d(x,X_1)]}{\mathbb{E}[K_1]}$$

l'intégration par rapport à la distribution de la variable réelle $d(x, X)$ montre que

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] = \frac{hg(x)\int_0^h K(h^{-1}t)\varphi_x(t)\phi'(t)dt + O(\phi(h))}{-g(x)\int_0^1 K'(t)\phi(th)dt + O(\phi(h))}$$

On prend le changement de variable qui $h^{-1}t = s$

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] = \frac{hg(x)\int_0^h K(s)\varphi_x(hs)\phi'(hs)ds + O(\phi(h))}{-g(x)\int_0^1 K'(t)\phi(th)dt + O(\phi(h))}$$

ll est clair que, selon la définition de α_1, le dénominateur normalisé par $g(x)\phi(H)$ converge vers α_1 ce qui implique que

$$\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] = \frac{hg(x)}{\phi(x)\alpha_1 + O(\phi(h))}\left(\int_0^h K(h^{-1}t)\varphi_x(t)\phi'(t)dt + O(\phi(h))\right)$$

Cela achève la démonstration du lemme.

Lemme 2.7. *Sous les hypothèses (H1')-(H2'), (H4') et (H5'), nous avons pour tout* $t \in \mathbb{R}$

$$\mathbb{E}[\hat{\Psi}_N(x,t] = \lambda_1(x,t) + O(h^{b_1}).$$

Preuve :Par stationnarité, nous avons

$$\left|\mathbb{E}\hat{\Psi}_N(x,t) - \lambda_1(x,t)\right| \le \left|\mathbb{E}\left[\frac{1}{\mathbb{E}[K_1]}K_1 1\!\!1_{B(x,h)}(X_1)(\lambda_1(X_1,t) - \lambda_1(x,t))\right]\right|.$$

Sous (H3), nous obtenons

$$|\mathbb{E}[\frac{1}{\mathbb{E}[K_1]}K_1 1\!\!1_{B(x,h)}(X_1)(\lambda_1(X_1,t) - \lambda_1(x,t))]| \le Ch^{b_1}.$$

Cela donne la preuve du lemme.

Corollaire 2.1.3. *Sous les hypothèses du lemme 2.7, et si le paramètre de lissage h satisfait $nh^{2b_1}\phi(h) \longrightarrow 0$, on a*

$$\sqrt{n\phi(h)}\mathbb{E}[\hat{\Psi}_N(x,\theta_x)] \longrightarrow 0, \quad \textit{quand} \quad n \longrightarrow \infty$$

Preuve corollaire :Par le lemme 2.7, on obtient facilement que

$$\sqrt{n\phi(h)}\mathbb{E}[\Psi_N(x,\theta_x)] = \sqrt{n\phi(h)h^{2b_1}}$$

Proposition 2.1.2.1. *(voir Azzedine et al. (2008)) Supposons que (H1') -(H2') ,(H3) et (H4')-(H5') satisfaites, donc $\hat{\theta}_x$ existe et il est unique avec une probabilité tend vers 1, et on a*

$$\hat{\theta}_x - \theta_x \longrightarrow 0, \quad \textit{en probabilité quand} \quad n \longrightarrow \infty.$$

Lemme 2.8. *Sous les hypothèses du théorème 2.1, on a*

$$\hat{\Psi}'_N(x,\xi_n) \longrightarrow \Gamma_1(x,\theta_x) \quad \text{en probabilité, quand} \quad n \longrightarrow \infty$$

Preuve : Nous considérons la décomposition suivante

$$\left|\hat{\Psi}'_N(x,\xi_n) - \Gamma_1(x,\theta_x)\right| \leq \left|\hat{\Psi}'_N(x,\xi_n) - \hat{\Psi}'_N(x,\theta_x)\right| + \left|\hat{\Psi}'_N(x,\theta_x) - \Gamma_1(\theta_x,x)\right| \quad (2.12)$$

En ce qui concerne le premier terme, observe que

$$\left|\hat{\Psi}_N(x,\xi_n) - \hat{\Psi}_N(x,\theta_x)\right| \leq \sup_{y\in\mathbb{R}}\left|\frac{\partial\psi(y,\xi_n)}{\partial t} - \frac{\partial\psi(y,\theta_x)}{\partial t}\right|\hat{\Psi}_D(x)$$

et parce que $\dfrac{\partial\psi(y,t)}{\partial t}$ est continue en θ_x uniformément en y, l'utilisation de la proposition 2.1.2.1 et la convergence en probabilité de $\hat{\Psi}_D(x)$ vers 1 montrent que le premier terme de 2.12 converge en probabilité vers 0. Cependant , la limite du second terme est obtenue en évaluant séparément le biais et les termes de variance de $[\hat{\Psi}'_N(x,\theta_x)]$. De toute évidence, des arguments similaires que ceux invoqués pour prouver le lemme 2.7 peuvent être utilisés pour obtenir

$$\mathbb{E}[\hat{\Psi}'_N(x,\theta_x)] \longrightarrow \Gamma_1(x,\theta_x).$$

En outre, en utilisant l'argument analogue à celle de 2.7, nous pouvons montrer

$$n\phi(h)Var[\hat{\Psi}_N(x,\theta_x)] \longrightarrow \frac{\alpha_2\Gamma_2(x,\theta_x)}{\alpha_1^2 g(x)}$$

Enfin, en utilisant (H1') et (H4'), nous obtenons la preuve de ce lemme.

2.2 Cas de mélange forte

Afin de généraliser les résultats obtenus dans le cas i.i.d à des observations fortement mélangeantes, nous renforçons les hypothèses précédentes, en ajoutant des hypothèses supplémentaires sur la concentration de loi conjointe (X_i, X_j) et sur le coefficient de mélange. Notre premier résultat concerne la convergence presque complète ponctuelle de la fonction de régression robuste.

2.2.1 Convergence presque complète

on garde les mêmes notations que celles de la section précédente, ainsi que les mêmes hypothèses. Cependant, pour ce cas là, on a besoin d'ajouter les hypothèses suivantes

- (H6) $(X_i, Y_i)_{i\geq 1}$ est une suite α mélangeante dont les coefficients satisfficient

$$\alpha(n) = O\alpha(n^{-a} \quad \text{pour certains} \quad a > 0$$

- (H7) $\sup_{i\neq j} \mathbb{P}((X_i, Y_i) \in B(x,h) \times B(x,h)) = O\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right)$.
- (H8) Il existe $\eta > 0$, telle que, $Cn^{\dfrac{3-a}{a+1}+\eta} \leq \phi_x(h) \leq \acute{C}n^{\frac{1}{1-a}}$, avec $a > \frac{5+\sqrt{17}}{2}$.

Théorème 2.2. *Supposons que (H1)-(H7) sont satisfaits. Si¡*$\Gamma(x, \theta_x) \neq 0$ *, alors* $\hat{\theta}$ *existe et est unique pour n suffisamment grand, et nous avons*

$$\hat{\theta}_x - \theta_x = O(h^{b1}) + O\left(\sqrt{\frac{logn}{n\phi_x(h)}}\right) p.co, pourn \longrightarrow \infty$$

La démonstration est essentiellement basée sur les mêmes techniques utilisées pour le cas i.i.d. Rappelons que l'hypothèse i.i.d. d'observations n'aucune influence sur la partie biais de la vitesse de convergence. En vertu de cette conclusion, nous allons évaluer uniquement la partie de dispersion, dont la vitesse de convergence est donné par les lemmes suivants.

Lemme 2.9. *Sous des hypothèses (H1) et (H4) - (H7), nous avons ,*

$$\hat{\Psi}_D(x) - \mathbb{E}[\hat{\Psi}_D(x)] = O\left(\sqrt{\frac{logn}{n\phi_x(h)}}\right) p.co, quandn \longrightarrow \infty$$

Preuve : Soit $\Delta_i(x) = K_i(x) - \mathbb{E}[K_i(x)]$: Alors,

$$\hat{\Psi}_D(x) - \mathbb{E}[\hat{\Psi}_D(x)] = \frac{1}{n\mathbb{E}[K_1(x)]}\sum_{i=1}^{n}\Delta_i(x)$$

Donc, nous appliquons l'ingalité Fuck-Nagaev exponentielle de(Rio , 2000, p. 87) pour obtenir pour tout $r > 0$ et $\varepsilon > 0$,

$$\mathbb{P}\left[|\mathbb{E}(\hat{\Psi}_D(x)) - \hat{\Psi}_D(x)| > \varepsilon\right] \leq \mathbb{P}\left[|\sum_{i=1}^{n}\Delta_i(x)| > \varepsilon n\mathbb{E}[K_1(x)]\right] \leq C(A_1(x) - A_2(x)) \tag{2.13}$$

où

$$A_1(x) = \left(1 + \frac{\varepsilon^2 n^2(\mathbb{E}[K_1(x)])^2}{r\mathrm{S}_n^2}\right)^{-r/2}, A_2(x) = nr^{-1}\left(\frac{r}{n\varepsilon\mathbb{E}[K_1(x)]}\right)^{a+1}$$

et

$$\mathrm{S}_n^2(x) = \sum_{i=1}^{n} Cov(\Delta_i(x), \Delta_j(x)) = \mathrm{S}_n^{2*}(x) + nVar[\Delta_1(x)] \tag{2.14}$$

avec

$$\mathrm{S}_n^{2*}(x) = \sum_{i\neq j} Cov(\Delta_i(x), \Delta_j(x)) \tag{2.15}$$

Ensuite, nous évaluons le comportement asymptotique de $\mathrm{S}_n^{2*}(x)$. Après Masry (1986), nous définissons les séries

$$E_1 = \{(i, j) \quad \text{tel que} \quad 1 \leq |i - j| \leq m_n\}$$

et

$$E_2 = \{(i,j) \quad \text{tel que} \quad m_n + 1 \leq |i-j| \leq n-1\}$$

où $m_n \longrightarrow \infty$ que $n \longrightarrow \infty$, alors

$$J_{1,n} = \sum_{E_1} Cov|(\Delta_i(x), \Delta_j(x))| \leq \sum_{E_1} |\mathbb{E}[K_i(x)K_j(x) - \mathbb{E}^2[K_1(x)]]|.$$

Sous (H1), (H7) et (H8) nous obtenons

$$J_{1,n} \leq Cnm_n\phi_x(h)\left(\left(\frac{\phi_x(h)}{n}\right)^{1/a} + \phi_x(h)\right)$$

Pour E_2, nous utilisons l'inégalité de Davydov-Rio (Rio , 2000, p . 87) pour les processus de mélange, pour $i \neq j$,

$$|Cov(K_i(x), K_j(x))| \leq C\alpha(|i-j|).$$

Par conséquent,

$$J_{2,n} = \sum_{E_2} |cov(K_i(x), K_j(x))| \leq n^2 m_n^{-a}$$

Choisir $m_n = \left(\frac{\phi_x(h)}{n}\right)^{-1/a}$ permet d'obtenir, sous (H8),

$$\mathrm{S}_n^{2*} = J_{1,n} + J_{2,n} = O(n\phi_x(h)) \tag{2.16}$$

Concernant le terme de variance, on déduit de (H1) que

$$Var(\Delta_1(x)) \leq C(\phi_x(h) + (\phi_x(h))^2) \tag{2.17}$$

Enfin, à partir de 2.14, 2.15, 2.16 et 2.17, nous obtenons

$$\mathrm{S}_n^2 = O(n\phi_x(h)). \tag{2.18}$$

Maintenant, nous appliquons 2.13 avec

$$\varepsilon = \lambda \frac{\sqrt{nlogn\phi_x(h)}}{n\mathbb{E}[K_1(x)]} \quad \text{et} \quad r = C(logn)^2. \tag{2.19}$$

Il en résulte que, en vertu de (H7)

$$A_2(x) \leq Cn^{-1-\eta(a+1)/2}(logn)^{(3a-1)/2}.$$

Ainsi, il existe $\nu > 0$ tel que

$$A_2(x) \leq Cn^{-1-\nu}. \tag{2.20}$$

A l'aide de 2.18 et 2.19, nous obtenons

$$\mathrm{A}_1(x) = C\exp(-\lambda^2\frac{logn}{2}) = Cn^{-\lambda^2/2}$$

Ainsi, pourλ assez grande :

$$\exists\nu' > 0, \quad A_1(x) \leq Cn^{-\lambda^2/2} \leq Cn^{-1-\nu'}. \tag{2.21}$$

Finalement, le résultat est déduire à partir de 2.13, 2.20 et 2.21.

Lemme 2.10. *Sous les hypothèses (H1) et (H3) -(H7), nous avons,*

$$\hat{\Psi}_N(x,t) - \mathbb{E}[\hat{\Psi}_N(x,t)] = O\left(\sqrt{\frac{logn}{n\phi_x(h)}}\right) p.co, \quad pour \quad n \longrightarrow \infty$$

Preuve De manière analogue à lemme 1, nous devons étudier le comportement asymptotique de

$$S_n'^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(\Lambda_i, \Lambda_j) = \sum_{i\neq j} Cov(\Lambda_i, \Lambda_j) = nVar(\Lambda_1)$$

avec

$$\Lambda_i = K_i(x)\psi_x(Y_i - t) - \mathbb{E}[K_1(x)\psi_x(Y_1 - t)]$$

On partage cette somme sur les deux ensembles

$$S'_1 = \{(i,j) \quad \text{tel que} \quad 1 \le i-j \le u_n\}$$

et

$$S'_2 = \{(i,j) \quad \text{tel que} \quad u_n+1 \le i-j \le n-1\},$$

sous (H1), (H7) et (H8), nous avons

$$J'_{1,n} \le CnU_n\phi_x(h)\left(\left(\frac{\phi_x(h)}{n}\right)^{1/a} + \phi_x(h)\right)$$

et

$$J'_{2,n} = \sum_{S'_2} |Cov(\Lambda_i, \Lambda_j)| \le n^2 u_n^{-a}$$

De (H7), nous pouvons prouver que si

$$u_n = \left(\frac{\phi_x(h)}{n}\right)^{-1/a}$$

$$\sum_{i=1}^{n}\sum_{i\neq j} Cov(\Lambda_i, \Lambda_j) = O(n\phi_x(h)).$$

Concernant la variance, l'hypothèse (H1)nous permet d'obtenir

$$Var(K_1(x)) \le C(\phi_x(h) + (\phi_x(h))^2)$$

enfin

$$S_n'^2 = O(n\phi_x(h)). \tag{2.22}$$

On applique Fuk-Nagaev pour inégalité. On obtient pour tout $r > 0$ et pour tous $\varepsilon > 0$,

$$\mathbb{P}\left[\left|\mathbb{E}(\hat{\Psi}_N(x,t)) - \hat{\Psi}_N(x,t)\right| > \varepsilon\right] \le \mathbb{P}\left[\left|\sum_{i=1}^{n}\Lambda_i\right| > \varepsilon n \mathbb{E}[K_1(x)]\right]$$

$$\leq C(A_1'(x) + A_2'(x))$$

où

$$A_1'(x) = \left(1 + \frac{\varepsilon^2 n^2 (\mathbb{E}[K_1(x)])}{S_2'^n r}\right)^{(-r/2)}$$

et

$$A_2'(x) = nr^{-1}\left(\frac{r}{\varepsilon n \mathbb{E}[K_1(x)]}\right)^{a+1}.$$

On prend $\varepsilon = \lambda' \frac{\sqrt{nlogn\phi_x(h)}}{n\mathbb{E}[K_1(x)]}$ et $r = C(logn)^2$ on obtient

$$A_2'(x) \leq Cr^a n^{1-(a+1)/2}\phi_x(h)^{-(a+1)/2}(logn)^{-(a+1)/2} \leq Cn^{-1-\nu_1'} \tag{2.23}$$

pour certains $\nu_1' > 0$. D'autre part

$$A_1'(x) \leq C\left(1 + \frac{\lambda'^2 logn}{r}\right)^{-r/2} \leq Cn^{-1-\nu_2'} \tag{2.24}$$

pour certains $\nu_1' > 0$. Alors 2.23 et 2.24 permettre de conclure le résultat.

2.2.2 Normalité asymptotique

Notre dernier résultat est la généralisation au cas de variables fonctionnelles alpha-dépendantes de la normalité asymptotique obtenu dans le cas i.i.d.

Théorème 2.3. *Sous des hypothèses de régularité de la fonction robuste ainsi que des conditions standards sur le noyau et la fenêtre, on a pour tout $x \in \mathcal{A}$, $\hat{\theta}_x$ existe et est unique avec une grande probabilité et*

$$\left(\frac{n\phi_x(h)}{\sigma_2(x,\theta_x)}\right)\left(\hat{\theta}_x - \theta_x - B_n(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad quand \quad n \longrightarrow \infty \tag{2.25}$$

où

$$B_n(x) = \frac{h}{\phi_x(h)\alpha_1\Gamma_1(x,\theta_x)}\int_0^1 K(t)\varphi_x(th)\phi_x'(th)dt + O(1)$$

$$(\varphi_x(h) = \mathbb{E}[\psi_x(Y, \theta_x / d(X, x)) = s])$$

$$\sigma^2(x, \theta_x) = \frac{\alpha_2 \lambda_2(x, \theta_x)}{\alpha_1^2 (\Gamma_1(x, \theta_x))^2}$$

$$\alpha_j = -\int_0^1 (K^j)' \beta(s) ds, pour, j = 1, 2$$

$$\mathbb{A} = \{x \in \mathbb{F}, \lambda_2(x, \theta_x) \Gamma_1(x, \theta_x) \neq 0\}$$

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

La démonstration de ce théorème est détaillée au (Attouch et al. 2008).

Chapitre 3

M-Régression non-paramétrique fonctionnelle : Expressions asymptotiques des erreurs $\mathbb{L}^q$

Ce chapitre a pour objectif de compléter l'éventail des propriétés asymptotiques de de l'estimateur robuste de la fonction de régressionen en exprimant les termes asymptotiquement dominants des moments centrés et des erreurs $\mathbb{L}_q$. Ces résultats innovants dans le cas fonctionnel le sont également dans le cas vectoriel . En plus de leur utilisation comme simples outils de calcul, ils peuvent se révéler particulièrement intéressants dans l'optique du choix optimal du paramètre de lissage par rapport aux erreurs Lq.

3.1 Un résultat d'uniforme intégrabilité

On donne dans ce paragraphe un résultat utile dans la suite pour obtenir la convergence des moments de $\hat{\theta}_x$. On donne le résultat pour des données $(X_i, Y_i)_{i=1,\ldots,n}$ indépendantes et identiquement distribuées. On peut également montrer, sous des hypothèses plus fortes, le même type de résultat pour des données arithmétiquement α-mélangeantes. On considère les hypothèses suivantes.

3.1.1 Le cas indépendant

Notations et principales hypothèses

On considère les hypothèses suivantes.

- (H.1) Il existe $p > 2$ et $C > 0$, tels que, pour X dans un voisinage ouvert de x, on a presque suûrement

$$\mathbb{E}[|\Psi_x(Y,t)|^p/X] \leq C. \tag{3.1}$$

- (H.2) On suppose que

$$\lim_{n\longrightarrow\infty} nF(h_n) = +\infty. \tag{3.2}$$

- (H.3)

$$K \text{ est à supporté compact } [0,1], \text{ est bornée et } K(1) > 0. \tag{3.3}$$

Théorème 3.1. *Sous les hypothèses (H.1)-(H.3), pour* $0 < q < p$*, la quantité*

$$|\sqrt{nF(h_n)}(\Psi_n(x,t) - \mathbb{E}[\Psi_n(x,t)])|^p$$

est uniformément intégrable, où $\Psi_n(x,t) = 1/nF(h_n)\hat{\rho}_n(x,t)$.

Preuve Nous introduisons les variables

$$W_{i,h} = \psi_x(Y_i,t)K(h^{-1}d(Xi,x)) - \mathbb{E}(\psi_x(Y_i,t)K(h^{-1}d(X_i,x)))$$

Nous allons montrer que la famille $((|\sqrt{nF(h_n)}(\Psi_n(x,t) - \mathbb{E}[\Psi_n(x,t)])|)^p)_n$ a tous ses moments d'ordre r bornés, avec $r = p/q > 1$. Par conséquent, nous devons prouver que

$$\sup \mathbb{E}[(|\sqrt{nF(h_n)}(\Psi_n(x,t) - \mathbb{E}[\Psi_n(x,t)])|)^p] < +\infty. \tag{3.4}$$

ce qui impliquera que cette famille est uniformément intégrable pour $0 < q < p$. Pour ce faire, nous utilisons le lemme suivant.

Lemme 3.1. *Si les couples (X_i, Y_i) sont indépendants, sous les hypothèses (H.1) et (H.2) pour tous $0 \leq q \leq p$, il existe des constantes $K1$ et $K2$ (avec $K1 = 0$ dans le cas $q < 2$) tel que pour tout $l \in \mathbb{N}$ et pour tout $h > 0$ (assez petit et indépendant avec rapport à n), on a*

$$\mathbb{E}\left[\left|\sum_{i=1}^{l} W_{i,h}\right|^q\right] \leq K_1 l F(h) + K_2 (l F(h))^{q/2}. \tag{3.5}$$

Nous appliquons ce lemme avec $l = n$ et $h = h_n$ et nous obtenons

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} W_{i,h_n}\right|^q\right] \leq K_1 n F(h_n) + K_2 (n F(h_n))^{q/2} \leq C (n F(h_n))^{q/2}. \tag{3.6}$$

En effet, pour $q > 2$, nous utilisons hypothèse (3.2), et pour $q \leq 2$, la constante K_1 est égal à zéro. Ensuite, il suffit de remarquer que

$$\sqrt{nF(h_n)} |\Psi_n(x,t) - \mathbb{E}[\Psi_n(x,t)]| = \frac{\left|\sum_{i=1}^{n} W_{i,h_n}\right|}{\sqrt{nF(h_n)}} \tag{3.7}$$

Avec les équations 3.6 et 3.7, nous obtenons finalement l'équation 3.4.

3.1.2 Le cas α-mélangeant

Lorsque l'on s'intéresse à certains problèmes, l'hypothèse d'indépendance des variables ne peut être conservée. C'est par exemple le cas lorsque l'on étudie une série temporelle. Typiquement, on veut prévoir certaines des valeurs futures à partir des valeurs dont on dispose. Plus précisément, nous considérons les variables arithmétiquement α-mélange de commander un.

Notations et hypothèses supplémentaires

Dans le cas où les variables sont dépendantes il nous faut pouvoir contrôler la somme des covariances ; on utilise pour cela l'hypothèse :

$$\exists u \leq p, \exists M, \forall q_j \in \mathbb{N}, \sum_{j=1}^{u} q_j \leq u, \max_{1\leq i_j \leq n} \mathbb{E}\left[\left|\prod_{j=1}^{u} \Psi_x(Y_{i_j}, t)^{q_j}\right| / X_{i_1}, \ldots, X_{i_u}\right] \leq M \tag{3.8}$$

Il nous faut encore introduire quelques notations :

$$k \geq 2, \Theta_k(s) := \max\left(\max_{1\leq i_1 \leq i_k \leq n} p(d(X_{i_j}, x) \leq s, 1 \leq j \leq k, F(s)^k)\right). \tag{3.9}$$

S'ajoutent également des hypothèses portant sur les fonctions Θ_k et l'ordre a des coefficients de mélange. et nous supposons qu'il existe $v < p$ tel que pour tout $k \leq v$, il existe $\nu > 0$ satisfaisant :

$$\Theta_k(s) := O(F(s)^{1+\nu k}) \quad \text{avec} \quad \nu_{k-1}+1 \geq \nu_k \quad \text{et} \quad a > \max_{2\leq k\leq v}(k-1)\frac{(1+\nu_k)p - v}{\nu_k(p-v)} \tag{3.10}$$

$$\exists \gamma > 0, nF(h_n)^{1+\gamma} \longrightarrow \infty. \tag{3.11}$$

Avant de donner le résultat de l'intégrabilité uniforme dans ce cas dépendante, on pose

$$l = 2\left[\frac{\min(v,u)}{2}\right],$$

où $[x]$ représente le plus grand nombre entier inférieur ou égal à x.

Théorème 3.2. *On suppose que les couples (X_i, Y_i) sont arithmétiquement α-mélangeants et que les hypothèses (H.1)-(H.3) et 3.8-3.11 sont vérifiées . On a alors, pour tout $0 \leq q < l$*

$$\left|\sqrt{nF(h_n)}(\Psi_n(x,t) - \mathbb{E}[\Psi_n(x,t)])\right|^q$$

est uniformément intégrable.

Preuve

Le schéma de la preuve est le même que celui la démonstration du Théorème 3.1 , en remplaçant le lemme 3.1 par le lemme suivant.

Lemme 3.2. *Si les couples (X_i, Y_i) sont arithmétiquement α-mélange d'ordre a, sous les hypothèses (3.1) et (9)-(12), puis, pour tout $0 \leq q \leq l$, il existe une constante a de telle sorte que K_1*

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} W_{i,h_n}\right|^q\right] \leq K_1 (nF(h_n))^{q/2}$$

3.2 Expression asymptotique des erreurs $\mathbb{L}_q$

A partir des résultats généraux de la partie précédente, il est possible de donner l'expression asymptotique des moments d'ordre q de l'estimateur à noyau. Les lemmes dont nous avons eu besoin pour montrer les résultats précédents permettent d'en exprimer les termes dominants. Nous supposons maintenant que les hypothèses (3.1) et (3.7) sont satisfaits avec $t = \theta_x$. On note ξ_n la variable aléatoire (entre θ_x et $\hat{\theta}_n$) telle que

$$B_n = -\frac{\mathbb{E}[\Psi_n(n, \theta_n)]}{\mathbb{E}[\partial \Psi_n / \partial t(x, \zeta_n)]}.$$

On suppose que

$$Z_n := \sqrt{\frac{nF(h_n)}{V_n}}(\hat{\theta}_n - \theta_x - B_n) \longrightarrow \mu. \tag{3.12}$$

où μ est une distribution de probabilité connue et V_n est la variance (ou son terme dominant) de la quantité $\sqrt{nF(h_n)}(\hat{\theta}_n - \theta_x - B_n)$. On suppose que les hypothèses (H.1)-(H.3) sont vérifiées (avec $t = \theta_x$)), ainsi que certaines conditions techniques données ci-dessous.

– (H.4)

$$t \longmapsto \sup(\frac{\partial \Psi_x}{\partial t}(y,t) - \frac{\partial \Psi_x}{\partial t}(y,\theta_x)) \quad \text{est continue au voisinage de} \theta_x. \tag{3.13}$$

– (H.5) Il existe une constante N telle que, presque suêrement dans un voisinage de x

$$\max \mathbb{E}\left[\left(\frac{\partial \Psi_x}{\partial t}(Y_i,\zeta_n) - \frac{\partial \Psi_x}{\partial t}(Y_i,\theta_x)\right)^{\beta_i}\left(\frac{\partial \Psi_x}{\partial t}(Y_i,\zeta_n) - \frac{\partial \Psi_x}{\partial t}(Y_i,\theta_x)\right)^{\beta_j}\right] \leq N. \tag{3.14}$$

– (H.6) Il existe des constantes γ et δ telles que

$$\mathbb{E}\left[\frac{\partial \Psi_x}{\partial t}(Y,\theta_x)/X\right] 1\!\!1_{d(X,x)\leq\delta} \geq \gamma 1\!\!1_{d(X,x)\leq\delta} \tag{3.15}$$

– (H.7) En ce qui concerne la quantité B_n, nous supposons que

$$\sqrt{nF(h_n)} B_n = O(1) \tag{3.16}$$

– (H.8) Il existe $p > 2$ et une constante $0 < C < +\infty$ telle que, pour X dans un voisinage ouvert de x, on a presque suŕement

$$\mathbb{E}\left[\left|\frac{\partial \Psi_x}{\partial t}(Y,\zeta_n)\right|^p /X\right] \leq C \leq \tag{3.17}$$

– (H.9) Pour le cas indépendant et il existe $u \leq p$ et une constante M telle que pour tout $q_j \in \mathbb{N}$ avec $\sum_{j=1}^{u} q_j \leq u$, nous avons presque sûrement

$$\max_{1\leq i_j\leq n} \mathbb{E}\left[\left|\prod_{j=1}^{u} \frac{\partial \Psi_x}{\partial t}(Y,\zeta_n)/X_{i_1}, ..., X_{i_u}\right|^{q_j}\right] \leq M \tag{3.18}$$

– (H.10) Il existe r et une constante $0 < M_0 < +\infty$ telle que

$$\mathbb{E}\left[|\hat{\theta}_n - \theta_x|^r\right] \leq M_0 \tag{3.19}$$

Théorème 3.3. *Sous les hypothèses du théorème 3.1 ou les hypothèses du théorème 3.2, ainsi que les hypothèses (H.4)-(H.10), nous avons pour tout* $q < (lr/r + l)$

$$\mathbb{E}[|\hat{\theta}_n - \theta_x|^q] = \mathbb{E}\left[\left|B_n + \sqrt{\frac{V_n}{nF(h_n)}}W\right|^q\right] + O\left(\frac{1}{\sqrt{nF(h_n)}^q}\right)$$

Nous pouvons remarquer que, lorsque W est une variable gaussienne standard, les erreurs peuvent être explicitées $\mathbb{L}^q$ de la même manière que dans l'article de Delsol [21]. La preuve du résultat basée sur les résultats de l'uniforme intégrabilité et de convergence de $\sqrt{nF(h_n)}(\hat{\theta}_n - \theta_x - B_n)$.

Preuve

La preuve du théorème 3.19 repose sur les lemmes suivants. Nous nous concentrons ici sur la preuve pour les variables de α-mélange (le cas des variables indépendantes est similaire).

Lemme 3.3. *Si les couples* (X_i, Y_i) *sont arithmétiquement* α*-mélange et identiquement distribuées, si les hypothèses (H.4)-(H.10) sont vérifiées, si* $0 \leq q < lr/(r + l)$, $|Z_n|^q$ *est uniformément intégrable.*

Lemme 3.4. *Soit* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *une fonction continue et* $\mathbb{F}_{C,M} = \{f_{\zeta,M}(x) = \mathbb{1}_{|x|\leq M}\}$

Si U_n *converge en loi à* U*, dont la fonction est continue ,*

$$\sup_{|\zeta|\leq C} |\mathbb{E}[f_{\zeta,M}(U_n)] - \mathbb{E}[f_{\zeta,M}(U)]| \longrightarrow 0$$

Nous commençons par

$$T_n := \left|\sqrt{\frac{nF(h_n)}{V_n}}(\hat{\theta}_n - \theta_x)\right|^q = \left|\sqrt{\frac{nF(h_n)}{V_n}}B_n\right|^q.$$

En utilisant le lemme 3.3, nous avons $|Z_n|^q$ uniformément intégrable. En outre, l'hypothèse (14) donne le la convergence dans la distribution de Z_n a

W. Assumption (18) permet d'obtenir $\zeta_n = \sqrt{nF(h_n/V_n)}B_n$ délimitée par une constante C. En conséquence, T_n est uniformément intégrable car

$$T_n \leq 2^q \left| \sqrt{\frac{nF(h_n)}{V_n}} B_n \right|^q + 2^q |Z_n|^q$$

D'ailleurs, si nous fixons $\epsilon > 0$, nous considérons la décomposition suivante :

$$\mathbb{E}[|Z_n + \zeta_n|^q] - \mathbb{E}[|W + \zeta_n|^q] = \mathbb{E}[|Z_n + \zeta_n|^q \mathbb{1}_{|Z_n| \leq M}] - \mathbb{E}[|W + \zeta_n|^q \mathbb{1}_{|w| \leq M}]$$

$$+ \mathbb{E}[|Z_n + \zeta_n|^q \mathbb{1}_{|Z_n| > M}] - \mathbb{E}[|W + \zeta_n|^q \mathbb{1}_{|w| > M}] \qquad (3.20)$$

Tout d'abord, nous remarquons que $|Z_n| > M$ (respectivement $|W| > M$) implique que $|Z_n + \zeta_n| > M - C$ (respectivement $|W + \zeta_n| > M - C$). Donc, nous pouvons donner une limite supérieure pour la dernière ligne de l'équation (3.20)

avec

$$|\mathbb{E}[|Z_n + \zeta_n|^q \mathbb{1}_{|Z_n| > M}] - \mathbb{E}[|W + \zeta_n|^q \mathbb{1}_{|w| > M}]|$$

$$\leq \mathbb{E}[|Z_n + \zeta_n|^q \mathbb{1}_{|Z_n + \zeta_n| > M - C}] - \mathbb{E}[|W + \zeta_n|^q \mathbb{1}_{|w + \zeta_n| > M - C}]$$

Enfin, l'intégrabilité uniforme de T_n et $|W + \zeta_n|^q$ permet d'obtenir, sous forme suffisamment grande ($M \geq M_0$), les limites supérieurese

$$\limsup_{n \longrightarrow \infty} \mathbb{E}[|Z_n + \zeta_n|^q \mathbb{1}_{\{|Z_n + \zeta_n| > M - C\}}] < \epsilon/3. \limsup_{n \longrightarrow \infty} \mathbb{E}[|W + \zeta_n|^q \mathbb{1}_{\{|W + \zeta_n| > M - C\}}] < \epsilon/3. \qquad (3.21)$$

Nous appliquons le lemme 3.3 au $\mathbb{F}_{C,M}$ de la famille, définie par $f(x) = |x|^q$, et nous concluons que, pour tous les M

$$|\mathbb{E}[f_{\zeta_n,M}(Z_n)] - \mathbb{E}[f_{\zeta_n,M}(W)] \leq \sup_{|\zeta| \leq C} |\mathbb{E}[f_{\zeta,M}(Z_n)] - \mathbb{E}[f_{\zeta_n,M}(W)]| \leq O(1).$$

En conséquence, il existe N_0 elle que pour tout $n \geq N_0$, nous avons

$$|\mathbb{E}[f_{\zeta_n,M}(Z_n)] - \mathbb{E}[f_{\zeta_n,M}(W)] \leq \epsilon/3 \qquad (3.22)$$

En utilisant les équations (3.20), (3.21) et (3.22), nous pouvons conclure que

$$\forall \epsilon > 0, \exists N_0, \forall n_0 \geq N_0, |\mathbb{E}[|Z_n + \zeta_n|^q] - \mathbb{E}[|W + \zeta_n|^q] \leq \epsilon$$

qui réalise la preuve du théorème 3.19.

Conclusions

Grâce à l'expression explicite de la loi asymptotique de M-estimateur à noyau de la fonction de régression, on a pu obtenir les expressions asymptotiques des moments. Ces résultats sont très motivants car ils ouvrent d'intéressantes perspectives, en particulier vis-à-vis du choix de fenêtre au travers de l'erreur $\mathbb{L}_q$. Il serait notamment intéressant d'étudier comment l'estimateur se comporte sur des données réelles lorsque l'on choisit d'utiliser d'autres critères que l'erreur $\mathbb{L}^1$ ou $\mathbb{L}_2$. De plus les résultats établis sont innovants dans le cadre fonctionnel général mais aussi dans le cadre vectoriel où ils s'appliquent sous des hypothèses plus simples.

Bibliographie

Attouch.M , A. Laksaci, E. Ould-Saïd. (2008). Asymptotic distribution of robust estimator for functional nonparametric models. Comm. Statist. Theory and Methods, in press.

Attouch, M., Laksaci, A. et Ould-Saïd, E. (2007) Strong uniform convergence rate of robust estimator of the regression function for functional and dependent processes. Technical Report L.M.P.A.

Attouch, M., Laksaci, A., Ould-Saïd, E. (2007b).Asymptoti cnormality of robust estimator of the regression function for functional time series data.

Attouch, M. ; Gheriballah A. et A. Laksaci. (2007). Convergence Presque complète d'un estimateur robuste de la régression non paramétrique fonctionnelle : Cas spatial, Pub. Inst. Stat. Univ. Paris LI. **56**. 3-16.

Azzeddine, N., Laksaci, A., Ould-Saïd, E. (2008). On the robust nonparametric regression estimation for functional regressor. Statistics & Probability Letters. **78**, 3216-3221.

Besse, P., Cardot, H. et Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations. Scand. J. Stat. **27**, 673-687.

Barrientos-Marin, J., Ferraty, F., Vieu, P. (2010). Locally modelled regression and functional data. J. of Nonparametric Statistics. **22**, 617-632.

Boente, G., Gonzalez-Manteiga, W., Gonzalez, A. (2009).

Robust nonparametric estimation with missing data. J. Statist. Plann. Inference. **139**, 571-592.

Boente, G. and Fraiman, R. (1989). Nonparametric regression estimation. J. Multivariate Anal. **29**, 180-198.

Boente, G. and Fraiman, R. (1990). Asymptotic distribution of robust estimators for nonparametric models from mixing processes. Ann. Statist. **18**, 891-906.

Boente, G., Fraiman, R., Meloche, J. (1997). Robust plug-in bandwidth estimators in nonparametric regression. J. Statist. Planning and Inf. **57**, 109-142.

Collomb, G., Härdle, W. (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction : Kernel regression estimation from dependent observations. Stochastic Process. Appl. **23**, 77-89.

Crambes, C, Delsol, L., Laksaci, A. (2008). Robust nonparametric estimation for functional data. J. Nonparametr. Stat. **20**, 573-598.

Carbon, M., Tran, L.T., Wu, B. (1997). Kernel density estimation for random fields : the L_1 theory, J.Nonparametric Statist. **6**, 157-170.

Carbon,M., Francq, C., Tran, L.T. (2007). Kernel regression estimation for random fields. J. Statist. Planning and Inf. **137**, 778-798.

Dabo-Niang, S. and Laksaci, A. (2011a). Nonparametric quantile regression estimation for functional dependent data. Comm in Statist. Theory & Methods. **41**, 1254-1268.

Devroye, L. (1978). The uniform convergence of the Nadaraya-Watson regression function estimate. Canad. J. Statist. **6**, No.2, 179-191.

Fan, J., Hu, T. C., Truong, Y. K. (1994). Robust nonparametric function estimation. Scand. J. Statist. **21**, 433-446.

Ferraty, F., Vieu, P. (2006). Nonparametric functional data analysis. Theory and Practice . Springer-Verlag.

Guessoum, Z. ; Ould Saïd, E. (2010). Kernel regression uniform rate estimation for censored data under α-mixing condition. Electron. J. Stat. **4**, 117-132.

Guessoum, Z. ; Ould Saïd, E. (2012). Central limit theorem for the kernel estimator of the regression function for censored time series. J. Nonparametr. Stat. **24**, 379–397. Guyon, X. (1995). Random Fields on a Network-Modeling, Statistics and applications. Springer, New York.

Györfi, L., Härdle, W., Sarda, P. et Vieu, P. (1989). Nonparametric curve estimation for time series, Lecture Notes in Statistics. **60**, Springer-Verlag.

Hardle, W. and Vieu, P. (1992). Kernel regression smoothing of times series. *J. of Times Series Analysis.* **13**, 209-232.

Heng, D. and Leung, Y. (2005). Cross-validation in nonparametric regression with outliers. *Ann. Statist.* **33**, 2291-2310.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35**, 73-101.

Laïb, N. and Ould-Saïd, E. (2000). A robust nonparametric estimation of the autoregression function under an ergodic hypothesis. *Canad. J. Statist.* **28**, 817-828.

Li, J. and Tran, L.T. (2009). Nonparametric estimation of conditional expectation. *J. Statist. Plan. and Inference.* **139**, 164-175.

Laïb, N. and Louani, D. (2011). Strong consistency of the regression function estimator for functional stationary ergodic data J. Statist. Plan. and Inference. **141**, 359-372.

Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. Stoch. Proc. and their Appl. **115**, 155-177.

Nadaraya, E. (1964). On estimating regression. Theory Prob. Appl. **10**, 186-196.

Nakhapetyan, B.S. (1987). An approach to the proof of limit theorems for dependent random variables. Theory of Probab. and its Appl. **32**, 535-539.

ROBINSON, R. (1984). *Robust Nonparametric Autoregression.* Lecture Notes in Statistics, **26**, 247-255. Springer-Verlag, New-York.

S. A. OULD ABDI, S. DABO-NIANG, A.DIOP, AND A. OULD ABDI. (2010). Consistency of a Nonparametric Conditional Quantile Estimator for Random Fields. Mathematical Methods of Statistics. **19**, 1-21.

VIEU, P. (1991). Quadratic errors for nonparametric estimates under dependence. J. Multivariate Anal. **39** (2), 324-347.

WATSON, G. S. (1964). Smooth regression analysis. Sankhya Ser. A. **26**, 359-372.