REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL-ABBÈS

THESE

Présentée par:

CHAHAD ABDELKADER

Pour obtenir le Diplôme de Doctorat

Spécialité : Mathématiques
Option : Probabilités-Statistiques

Intitulée

Sur la régression relative pour des données incomplètes

Thèse soutenue le / JUIL/2019
Devant le jury composé de :

Président :

Mr MECHAB Boubaker	MCA	à L'Université S.B.A.

Directeur de thèse :

Mr LAKSACI ALI	Professeur	à L'Université S.B.A

Examinateur :

Mr BENCHIKH Tawfik	Professeur	à L'Université S.B.A
Mr GUENDOUZI Toufik	Professeur	à L'Université de Saida
Mr CHIKR EL MEZOUAR Zouaoui	MC A	à L'Université de Béchar
Mr RAHMANI Saâdia	MC A	à L'Université de Saida

Dédicace

À MES CHERS PARENTS

Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma considération pour les sacrifices que vous avez consenti pour mon instruction et mon bien être.
Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours.
Que ce modeste travail soit l'exaucement de vos voeux tant formulés, le fruit de vos innombrables sacrifices, bien que je ne vous en acquitterai jamais assez.
Puisse Dieu, le Très Haut, vous accorder santé, bonheur et longue vie et faire en sorte que jamais je ne vous déçoive.

À MA CHERE FEMME

Je tiens à présenter mes reconnaissances à ma chere femme Qui n'a jamais cessé de me soutenir pour que je puisse finir mes études et avoir une bonne formation et surtout être le meilleur et à qui je voudrais exprimer mes affections et mes gratitudes. Que dieu réunisse nos chemins pour un long commun serein et que ce travail soit témoignage de mon amour sincère et fidèle.

Je dedie ce travail à mes enfants, mes frères et mes soeurs, qui sont ma source d'inspiration et mon plus grand soutien , et à tout ce qui me sont plus proches.

Remerciements

Je tiens à témoigner ma reconnaissance à **DIEU** tout puissant,de ma voir donner le courage et la force de mener à terme ce projet.Qui ma ouvert les portes du savoir.

Mes chaleureux remerciements s'adressent à Monsieur le professeur **LAKSACI ALI** . Ce fut pour moi un réel plaisir de travailler sous sa direction tout au long de ces quatre années. Et me donner l'occasion de développer mes compétences en statistique fonctionnelle.je lui suis reconnaissant de m'avoir fait bénéficier tout au long de ce travail de sa compétence, de sa rigueur intellectuelle, de son efficacité certaine que je n'oublierai jamais, leurs qualités humaines, leur cordialité, leur patience infinie et leur soutien.

Je suis très honoré que Monsieur **MECHAB BOUBAKER**, maitre de conférences à l'université Djilali liabbés a accepté d'être président du jury. Je le remercie également pour la confiance qu'il m'a témoignée et pour tous ses conseils et remarques constructives.

Mes remerciements s'adressent aux Messieurs : **BENCHIKH TAWFIK**, Professeur à l'Université de Sidi Bel-Abbès ; **CHIKER-ELMEZOUAR ZOUAOUI**, Professeur à l'Université de Bechar ; **GUENDOUZI TOUFIK** Professeur à l'Université de Saida et **RAHMANI SAADIA**, Maître Conférence à l'Université de Saida pour avoir accepté de rapporter cette thèse. Je les remercie aussi pour le temps et l'attention qu'ils y'ont consacrés pour la lecture attentive de cette thèse.

Je ne pourrais finir mes remerciements sans une pensée très personnelle aux membres de ma famille, plus particulièrement mes parents, pour leur amour, leur encouragement et leur soutien inconditionnel tout au long de mes années d'étude qui m'ont mené jusqu'au doctorat. Merci à mes frères et mes soeurs pour leur soutien et leur appui moral.

Je remercie de tout mon coeur ma femme de faire partie de ma vie, sa présence, ses conseils, ses encouragements et son soutien constant et réconfortant dans les moments de doute sont pour moi très précieux.

Afin de n'oublier personne, mes vifs remerciements s'adressent à tous ceux qui m'ont aidée à la réalisation de cette modeste thèse.

Table des matières

Chapitre 1

Introduction

1.1 Résumé

Cette thèse est consacrée à l'étude d'un nouvel estimateur de l'opérateur de régression d'un variable de réponse scalaire étant donné une variable explicative fonctionnelle. Ce dernier est construit en minimisant la moyenne d'erreur relative au carre de l'operateur de régression linéaire locale. Comme résultats asymptotiques, nous établissons la convergence ponctuelle et uniforme presque complète avec les taux de cet estimateur. Une étude de Monte Carlo est réalisée pour évaluer la performance de cette estimateur

1.2 Abstract

This thesis is dedicated to the survey a new estimator of the regression operator of a scalar response variable given a functional explanatory variable. The latter is constructed by minimizing the mean squared relative error of the local linear regression operator. As asymptotic results, we establish the pointwise and the uniform almost complete consistency with rates of this estimator. A Monte Carlo study is carried out to evaluate the performance of this estimate.

List des traveaux

- Chahad, Abdelkader ; Ait-Hennani, Larbi ; Laksaci, Ali. Functional local linear estimate for functional relative-error regression. J. Stat. Theory Pract. 11 (2017), no. 4, 771–789

1.3 Literature Review

Let us introduce n pairs of random variables (X_i, Y_i) for $i = 1, \ldots, n$ that we assume drawn from the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}_+^*$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. Furthermore, we assume that the variables X and Y are connected by the following relation

$$Y = R(X) + \epsilon, \tag{1.1}$$

where R is an operator from $\mathcal{F}$ to $\mathbb{R}$ and ϵ is a random error variable independent to X.

The nonparametric estimation of the operator R is one of the most important tools to predict the relationship between Y and X. In this situation of functional covariate, there exist several nonparametric procedures allowing to estimate this operator. A popular one is the functional version of the Nadaraya-Watson estimator. The latter has been introduced by Ferraty and Vieu (2006). They obtained the almost complete consistency of this estimator. The functional version of the M-estimator of the regression operator has been studied by Chen and Zhang (2009). Burba et al. (2009) have investigated the asymptotic properties of the k-NN estimator of the regression function. Recently, Demongeot et al. (2016) have considered another estimation method based on the relative error technique. In this paper, we construct a new estimator of the regression operator. The latter is obtained by combining the ideas of the relative error regression with those of the local linear approach. Noting that the local linear approach has various nice features over the kernel method, in particular, it has a small bias compared to this last. On the other hand, it is well known that the relative error regression is more robust than the least square error regression, namely, in the presence of outliers. So, our new estimator has the nice features of both approaches.

Recall that, the local linear method was introduced in the nonparametric functional data analysis (NFDA) by Baìllo and Grané (2009). They studied the L^2-consistency of the local linear estimate of the regression function when the explanatory variable takes its values in a Hilbert space. Barrientos-Marin *et al.* (2010) have proposed an alternative fast version of the functional local linear estimate which can be used for a more general functional regressor. They proved the almost complete convergence (with rate) of the proposed estimate. Berlinet et al. (2010) constructed another local linear estimate based on the inverse of the local covariance operator of the functional explanatory variable. Zhou and Lin (2016) gave the asymptotic normality of the functional local linear regression estimate. The relative error regression has been recently introduced in NFDA by Demongeot et al. (2016). They showed that this regression model has significant advantages over the classical regression. It should be noted that both local linear estimation or relative error regression have been extensively studied in the multivariate case (see, for example, Stone (1977), Fan et al. (1996), Masry (1996), Hallin et al. (2009), Narula and Wellington (1977), Jones et al (2008), Yang et Ye (2013)). However, much less attention has been paid to the local linear estimation of the relative error regression. As far as we know, only the paper by Jones et al. (2008) provides an estimator of the relative error regression based on the multivariate local linear procedure. In this contribution, we treat the general case where the covariates are of functional nature.

The main aim of this thesis is to establish the asymptotic properties of the constructed estimator. We prove the pointwise and the uniform almost complete consistencies of this estimate. These results are obtained under some standard conditions in NFDA. The considered conditions are closely linked to the functional structure of the data as well as to the functional nature of the nonparametric model. Noting that these questions, in infinite dimension are particularly interesting, at once for the fundamental problems they formulate, but also for many applications they may allow. There exists an increasing number of situations coming from different fields of applied sciences in which the data are of functional nature (see Bosq (2000), Ramsay and Silverman (2002), Ferraty and Vieu (2006) for an overview on functional data analysis and Zhang (2014), Hsing et al. (2015), Cuevas (2014), Goia and Vieu (2016) for recent advances and references).

1.4 Some previous results

Some previous results on the iid case classical kernel case

Theorem 1 *(See, Demongeot et al. (2016)) Under assumptions*

(H1) $\mathbb{P}(X \in B(x,r)) =: \phi_x(r) > 0$ *for all* $r > 0$ *and* $\lim_{r\to 0} \phi_x(r) = 0$.

(H2) For all $(x_1, x_2) \in \mathcal{N}_x^2$*, we have :*

$$|g_\gamma(x_1) - g_\gamma(x_2)| \leq C\ d^{k_\gamma}(x_1, x_2) \ \ \textit{for } k_\gamma > 0.$$

(H3) The kernel K *is a measurable function which is supported within* $(0, 1)$ *and satisfying :*

$$0 < C_2 \leq K(\cdot) \leq C_3 < \infty.$$

(H4) The small ball probability :

$$\frac{n\phi_x(h)}{\log n} \to \infty \ \textit{ as } n \to +\infty.$$

(H5) The inverse moments of the response variable :

$$\forall m \geq 2,\ \ E[Y^{-m}|X = x] < C < \infty.$$

, we have :

$$|\widetilde{r}(x) - r(x)| = O(h^{k_1}) + O(h^{k_2}) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right). \tag{1.2}$$

Theorem 2 *(See, Demongeot et al. (2016)) Under hypotheses*

(U1) $\forall x \in S_\mathcal{F}$*,* $\forall \epsilon > 0$*,* $0 < C\phi(\epsilon) \leq \mathbb{P}\left(X \in B(x, \epsilon)\right) \leq C'\phi(\epsilon) < \infty.$

(U2) There exists $\eta > 0$*, such that :*

$$\forall x, x' \in S_\mathcal{F}^\eta,\ \ |g_\gamma(x) - g_\gamma(x')| \leq Cd^{k_\gamma}(x, x'),$$

where $S_\mathcal{F}^\eta = \{x \in \mathcal{F}, \exists x' \in S_\mathcal{F}, d(x, x') \leq \eta\}$.

(U3) The kernel K *is a bounded and Lipschitz function on its support* $[0, 1)$.

(U4) The functions ϕ *and* $\psi_{S_\mathcal{F}}$ *are such that :*

(U4a) there exists $\eta_0 > 0,\ \forall \eta < \eta_0,\ \phi'(\eta) < C$*,*

(U4b) for n *large enough, we have :*

$$\frac{(\log n)^2}{n\,\phi(h)} < \psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right) < \frac{n\,\phi(h)}{\log n},$$

(U4C) the Kolmogorov's ϵ*-entropy of* $S_{\mathcal{F}}$ *satisfies :*

$$\sum_{n=1}^{\infty} \exp\left\{(1-\beta)\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)\right\} < \infty \text{ for some } \beta > 1.$$

(U5) For any $m \geq 2$*,* $E(|Y^{-m}||X = x) < C < \infty$ *for all* $x \in S_{\mathcal{F}}$ *and* $\inf_{x \in S_{\mathcal{F}}} g_2(x) \geq C' > 0$ *.*

we have :

$$\sup_{x \in S_{\mathcal{F}}} |\widetilde{r}(x) - r(x)| = O(h^{k_1}) + O(h^{k_2}) + O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n\phi(h)}}\right). \tag{1.3}$$

Theorem 3 *(See, Demongeot et al. (2016)) Under assumptions*

(M1) The concentration property (H1) holds. Moreover, there exists a function $\beta_x(\cdot)$ *such that :*

$$\text{for all } s \in [0,1], \quad \lim_{r \to 0} \frac{\phi_x(sr)}{\phi_x(r)} = \chi_x(s).$$

(M2) For $\gamma \in \{1,2\}$*, the functions* $\Psi_\gamma(\cdot) = E\left[g_\gamma(X) - g_\gamma(x)\middle| d(x,X) = \cdot\right]$ *are derivable at* 0*.*

(M3) The kernel K *satisfies (H3) and is a differentiable function on* $]0,1[$ *where its first derivative* K' *is such that :* $-\infty < C < K'(\cdot) < C' < 0$*.*

(M4) The small ball probability satisfies :

$$n\phi_x(h) \longrightarrow \infty.$$

(M5) For $m \in \{1,2,3,4\}$, the functions $E[Y^{-m}|X = \cdot]$ are continuous in a neighborhood of x.

we have :

$$E\left[\widetilde{r}(x) - r(x)\right]^2 = B_n^2(x)h^2 + \frac{\sigma^2(x)}{n\phi_x(h)} + o(h) + o\left(\frac{1}{n\phi_x(h)}\right),$$

where

$$B_n(x) = \frac{\left(\Psi_1'(0) - r(x)\Psi_2'(0)\right)\beta_0}{\beta_1 g_2(x)}$$

and

$$\sigma^2 = \frac{\left(1 - 2r(x)E[Y^{-3}|X = x] + r^2(x)E[Y^{-4}|X = x]\right)\beta_2}{g_2^2(x)\beta_1^2}$$

with

$$\beta_0 = K(1) - \int_0^1 (sK(s))'\chi_x(s)ds, \text{ and, } \beta_j = K^j(1) - \int_0^1 (K^j)'(s)\chi_x(s)ds, \text{ for } j = 1,2.$$

Theorem 4 *(See, Demongeot et al. (2016)) Assume that (M1)-(M5) hold, then for any $x \in \mathcal{F}$, we have :*

$$\left(\frac{n\phi_x(h)}{\sigma^2(x)}\right)^{1/2} \left(\widetilde{r}(x) - r(x) - B_n(x) - o(h)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty.$$

where $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

Some previous results on the dependent case

The spatial case studied by Atoouch et al. (2017) and the considered the random filed $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N)$ satisfies the following mixing condition :

$$\begin{cases} \text{There exists a function } \varphi(t) \downarrow 0 \text{ as } t \to \infty, \text{ such that} \\ \forall E,\ E^{'} \text{ subsets of } \mathbb{N}^N \text{ with finite cardinals} \\ \alpha\left(\mathcal{B}(E), \mathcal{B}\left(E^{'}\right)\right) = \sup_{B \in \mathcal{B}(E),\, C \in \mathcal{B}\left(E^{'}\right)} \left|\mathbb{P}(B \cap C) - \mathbb{P}(B)\,\mathbb{P}(C)\right| \\ \leq s\left(\text{Card}(E), \text{Card}\left(E^{'}\right)\right) \varphi\left(\text{dist}\left(E, E^{'}\right)\right), \end{cases} \tag{1.4}$$

where $\mathcal{B}(E)$(*resp.* $\mathcal{B}\left(E^{'}\right)$) denotes the Borel σ-field generated by $(Z_{\mathbf{i}},\ \mathbf{i} \in E)$ (*resp.* $\left(Z_{\mathbf{i}},\ \mathbf{i} \in E^{'}\right)$), Card$(E)$ (*resp.* Card$\left(E^{'}\right)$) the cardinality of E (*resp.* $E^{'}$), dist$\left(E, E^{'}\right)$ the Euclidean distance between E and $E^{'}$ and $s : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that :

$$s(n, m) \leq C \min(n, m), \qquad \forall n, m \in \mathbb{N}. \tag{1.5}$$

Theorem 5 *(see, Atoouch et al. (2017)) Under assumptions*

(H1) The density f of the variable X is positif function, of class C^2 in S.

(H2) The functions $r_l(\cdot)$; $(l = 1, 2)$ is of class C^2 in S

(H3) The joint probability density $f_{\mathbf{i},\mathbf{j}}$ of $X_{\mathbf{i}}$ and $X_{\mathbf{j}}$ exists and satisfies

$$|f_{\mathbf{i},\mathbf{j}}(u, v) - f(u)f(v)| \leq C \ \textit{for some constant } C \textit{ and for all } u,\ v,\ \mathbf{i} \textit{ and } \mathbf{j}$$

(H4) The mixing coefficient defined in (1.4) satisfies

$$\sum_{i=1}^{\infty} i^{N-1}(\varphi(i))^a < \infty, \quad \textit{for some } \ 0 < a < \frac{1}{2}.$$

(H5) K is continuous lipschitz function, symmetric , with compact support.

(H6) There exists $\gamma \in (0,1)$ such that

$$\begin{cases} \sum_{\mathbf{n}} \widehat{\mathbf{n}} \,(\log \widehat{\mathbf{n}})^{-1}\varphi(p_{\mathbf{n}}) < \infty \; for \;\; p_{\mathbf{n}} = O\left(\dfrac{\widehat{\mathbf{n}}^{1-\gamma}h^d}{\log \widehat{\mathbf{n}}}\right)^{1/2N} \\ and \\ \widehat{\mathbf{n}}^{\delta} h \to \infty \quad for\ certain \;\; \delta > 0. \end{cases}$$

(H7) The inverse moments d'order $l = 1, 2$ if the response variable such that,

$$\mathbb{E}\left(\exp\left(|Y^{-l}|\right)\right) \leq C \quad and\ \forall \mathbf{i}, \mathbf{j} \qquad \mathbb{E}\left(\left|Y_{\mathbf{i}}^{-l} Y_{\mathbf{j}}^{-l}\right| \,\Big|\, X_{\mathbf{i}}, X_{\mathbf{j}}\right) \leq C'.$$

and, if $\inf_{x\in S} g_2(x) > 0$, we have :

$$\sup_{x\in S} |\widetilde{\theta}(x) - \theta(x)| = O(h^2) + O_{a.co.}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}h^d}}\right) \qquad a.co. \tag{1.6}$$

Theorem 6 *(see, Atoouch et al. (2017)) Assume that (H1)-(H5) and (H7)-(H9) hold, then we have, for any $x \in \mathcal{A}$,*

$$\left(\frac{\widehat{\mathbf{n}}h^d}{\sigma^2(x)}\right)^{1/2} \left(\widetilde{\theta}(x) - \theta(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty.$$

where

$$\sigma^2 = \frac{(g_2(x) - 2\theta(x)E[Y^{-3}|X=x] + r^2(x)E[Y^{-4}|X=x])}{g_2^2(x)} \int_{R^d} K^2(z)dz,$$

$$\mathcal{A} = \left\{x \in S,\ \left(g_2(x) - 2\theta(x)E[Y^{-3}|X=x] + r^2(x)E[Y^{-4}|X=x]\right) g_2^2(x) \neq 0\right\}$$

and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

Mechab and Laksaci have studied the quasi-associated process defined by

Definition A sequence $(Z_n)_{n\in\mathbb{N}}$ of real random vectors variables is said to be quasi-associated, if for any disjoint subsets I and J of $\mathbb{N}$ and all bounded Lipschitz functions $f : \mathbb{R}^{|I|d} \to \mathbb{R}$ and $g : \mathbb{R}^{|J|d} \to \mathbb{R}$ satisfying :

$$|Cov(f(Z_i,\, i \in I),\, g(Z_j,\, j \in J))| \leq \mathrm{Lip}(f)\mathrm{Lip}(g) \sum_{i\in I}\sum_{j\in J}\sum_{k=1}^{d}\sum_{l=1}^{d} |Cov(Z_i^k, Z_j^l)| \quad (1.7)$$

(here and in the sequel $|I|$ denotes the cardinality of a finite set I)

where Z_i^k denotes the k^{th} component of Z_i, and

$$\mathrm{Lip}(f) = \sup_{x\neq y} \frac{|f(x) - f(y)|}{\|x-y\|_1}, \text{ with } \|(x_1, \ldots, x_k)\|_1 = |x_1| + \ldots + |x_k|.$$

Theorem 7 *(see, Mechab and Laksaci (2016)) Under assumptions (H1)-(H6) and, if* $\inf_{x\in S} g_2(x) > 0$,

(H1) The density f of the variable X is positif function, of class C^2 in S and such that :

$$\sup_{|i-j|\geq 1} \|f_{(X_i,X_j)}\|_\infty < \infty.$$

(H2) The functions $r_l(\cdot) = E[Y^{-\gamma}|X = \cdot]$; $(l = 1, 2)$ is of class C^2 in S

(H3) The random pair $\{(X_i, Y_i),\ i \in \mathbb{N}\}$ is quasi-associated with covariance coefficient λ_k, $k \in \mathbb{N}$ checked

$$\exists a > 0, \quad \exists C > 0, \text{ such that } \lambda_k \leq Ce^{-ak}.$$

(H4) K is continuous lipschitz function, symmetric , with compact support.

(H5) There exists $\gamma \in (0,1)$ *and* $\xi_1, \xi_2 > 0$ *such that*

$\frac{\log n^{1/d}}{n^{(1-\gamma-\xi_2)/d}} \leq h \leq \frac{C}{(\log n)^{(1+\xi_1)/d}}.$

(H6) The inverse moments d'order $l = 1, 2$ *f the response variable such that,*

$$\mathbb{E}\left(\exp\left(|Y^{-l}|\right)\right) \leq C \quad \text{and } \forall i \neq j \ \mathbb{E}\left(\left|Y_i^{-l} Y_j^{-l}\right| \Big| X_i, X_j\right) \leq C'.$$

we have :

$$\sup_{x \in S} |\widetilde{r}(x) - r(x)| = O(h^2) + O_{a.co.}\left(\sqrt{\frac{\log n}{n^{1-\gamma} h^d}}\right) \qquad a.co. \tag{1.8}$$

Theorem 8 *(see, Mechab and Laksaci (2016)) Assume that (H1)-(H4), (H5') and (H6) hold, then for any* $x \in \widehat{R}^d$*, we have :*

$$\left(\frac{nh^d}{\sigma^2(x)}\right)^{1/2} (\widetilde{r}(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty.$$

where

$$\sigma^2 = \frac{(g_2(x) - 2r(x)E[Y^{-3}|X = x] + r^2(x)E[Y^{-4}|X = x])}{g_2^2(x)} \int_{R^d} K^2(z)dz,$$

$$\mathcal{A} = \left\{x \in S,\ \left(g_2(x) - 2r(x)E[Y^{-3}|X = x] + r^2(x)E[Y^{-4}|X = x]\right) g_2^2(x) \neq 0\right\}$$

$\xrightarrow{\mathcal{D}}$ *means the convergence in distribution and*

– *There exists* $\gamma \in (0,1)$ *and* $\xi_1, \xi_2 > 0$ *such that*

$\frac{1}{n^{2(1/2-\gamma/9-\xi_2)/d)}} \leq h \leq \frac{C}{n^{(1+\xi_1)/(d+4)}}.$

Chapitre 2

Functional local linear relative regression : Complete data case

Ce chapitre fait l'objet d'une publication au *Journal of Statistical Theory and Practice*

Functional local linear estimate for functional relative-error regression

Abdelkader Chahad, Larbi Ait-Hennani , Ali Laksaci

Laboratoire de Statistique et Processus Stochastiques,

Université de Sidi Bel Abbès,

BP 89 Sidi Bel Abbès 22000. Algeria.

chahadaek@gmail.com, alilak@yahoo.fr

Université de Lille 2

Rond-point de l'Europe, BP 557 59060 Roubaix, France.

E-mail : larbi.aithennani@univ-lille2.fr

Abstract

We present in this paper a new estimator of the regression operator of a scalar response variable given a functional explanatory variable. The latter is constructed by minimizing the mean squared relative error of the local linear regression operator. As asymptotic results, we establish the pointwise and the uniform almost complete consistencies with rates of this estimator. A Monte Carlo study is carried out to evaluate the performance of this estimate.

keyword

Functional data analysis ; Nonparametric regression ; Local linear estimate ; Kernel estimate ; Relative-error .

Subject classifications : Primary : 62G05, Secondary : 62G07, 62G08, 62G35, 62G20.

2.1 Introduction

Let us introduce n pairs of random variables (X_i, Y_i) for $i = 1, \ldots, n$ that we assume drawn from the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}_+^*$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. Furthermore, we assume that the variables X and Y are connected by the following relation

$$Y = R(X) + \epsilon, \tag{2.1}$$

where R is an operator from $\mathcal{F}$ to $\mathbb{R}$ and ϵ is a random error variable independent to X.

The nonparametric estimation of the operator R is one of the most important tool to predict the relationship between Y and X. Such a subject has taken an important place in Nonparametric Functional data analysis (NFDA). Various nonparametric techniques can be found in the literatures of NFDA. We cite for instance Ferraty and Vieu (2006) for the functional Nadaraya-Watson estimator, Attouch et al.(2010) for the nonparametric robust estimation, Burba et al.(2009) for k-nearest neighbor kernel method, Barrientos et al. (2010) for the local linear approach or Demongeot et al.(2016) for the functional relative error techniques. The main purpose of this paper is to construct a new estimator of the regression operator. The latter is obtained by combining the ideas of the relative error regression with those of the local linear approach. Noting that the local linear approach has various nice features over the kernel method, in particular, has small bias compared to this last. On the other hand it is well known the relative error regression is more robust than the least square error regression, namely, in presence of the outliers. So, our new estimator has the nice features of both approaches.

Recall that, the local linear modeling was introduced in NFDA by Baìllo and Grané (2009). They studied the L^2- consistency of the local linear estimate of the regression function when the explanatory variable takes values in a Hilbert space. We refer to Barrientos *et al.* (2010) for the almost complete convergence (with rate) of an alternative version of the local linear estimate of the nonparametric functional regression. Their simplified version has been adapted by Laksaci et al. (2013) for others nonparametric models, such as the conditional distribution function, the conditional density or the conditional mode function. We return to Berlinet et al. (2011) for another version constructed by inverting the local covariance operator of the functional explanatory variable. They obtained the convergence rate of the mean quadratic error of the constructed estimator. In parallel, the relative error regression has been recently introduced in NFDA by Demongeot et al. (2016). They shown that this regression model has significant advantages over the classical regression. It should be noted that both local linear estimation or relative error regression have been extensively studied in the multivariate case. See, for example, Stone (1977), Fan et al. (1996), Masry (1996), Hallin et al. (2009), Narula and Wellington (1977), Jones et al (2008), Yang et Ye (2013), Laksaci and Mechab (2016), Attouch et al. (2016), among others. However, much less attention has been paid to the local linear estimation of the relative error regression. As far as we know, only the paper by Jones et al (2008) provides an estimator of the relative error regression based on the multivariate local linear procedure. In this contribution we treat the general case where the regressors are of functional nature. The main aim of this paper is to establish the asymptotic proprieties of the constructed estimator. We prove the pointwise and the uniform almost complete consistency of this estimate. These results are obtained under some standard conditions in NFDA. The considered conditions are closely linked to the functional structure of the data as well as to the functional nature of the nonparametric model. It worth to noting that these questions in infinite dimension are particularly interesting, at once for the fundamental problems they formulate, but also for many applications they may allow. There exits an increasing number of situations coming from different fields of applied sciences in which the data are of functional nature (see Bosq (2000), Ramsay and Silverman (2002), Ferraty and Vieu (2006) for an overview on functional data analysis and Zhang (2014), Hsing et al. (2015), Cuevas (2014), Goia and Vieu (2016) for recent advanced and references).

The paper is organized as follows : We construct our estimate in the following Section. We study the pointwise consistency in Section 3. The uniform almost complete convergence is treated in Section 4. Some simulated data examples are reported in Section 5. All proofs are put into the Appendix.

2.2 The model and its estimate

Unlike to the multivariate case, there exists various versions of the functional local linear estimate. But, all these versions are based on two common procedures. The first one is the functional operator which is supposed smooth enough to be locally well approximated by a polynom. The second one is the use of the following least square error

$$\mathbb{E}\left[(Y - R(X))^2|X\right]$$

as a loss function to determine the estimate. However, this criterion may be unadapted to some situations. Indeed, this loss function treats all variables, in the study, as having the same weight. Thus, this approach gives irrelevant results when the data contains some outliers. In this paper we circumvent this limitation by estimating the operator R with respect the following mean squared relative error

$$\text{For} \quad Y > 0, \quad \mathbb{E}\left[\left(\frac{Y - R(X)}{Y}\right)^2 |X\right]. \tag{2.2}$$

Clearly, this criterion is a more meaningful measure of the prediction performance than the least square error, in particular, when the range of predicted values is large. Moreover, solution of (2.2) is explicitly expressed by

$$R(x) = \frac{\mathbb{E}[Y^{-1}|X = x]}{\mathbb{E}[Y^{-2}|X = x]}.$$

In this work, we adopt the fast version proposed by Barrientos *et al.* (2010) and we use the loss function (2.2) to estimate components of the linear approximation.

Specifically, for a fixed point x in $\mathcal{F}$, we suppose that

$$\forall x' \text{ in neighborhood of } x \qquad R(x') = a + b\beta(x,x') + o(\beta(x,x'))$$

and we use the loss function (2.2) to estimate a, b as follows

$$(\widehat{a},\widehat{b}) = \arg\min_{(a,b)} \sum_{i=1}^{n} \frac{(Y_i - a - b\beta(X_i,x))^2}{Y_i^2} K(h^{-1}\delta(x,X_i)) \tag{2.3}$$

where $\beta(.,.)$ is a known function from $\mathcal{F}^2$ into $\mathbb{R}$ such that, $\forall \xi \in \mathcal{F}$, $\beta(\xi,\xi) = 0$, with K is a kernel and $h = h_{K,n}$ is a sequence of positive real numbers and $\delta(.,.)$ is a function defined on $\mathcal{F} \times \mathcal{F}$ such that $d(.,.) = |\delta(.,.)|$.

Clearly, by a simple algebra, we prove that $(\widehat{a},\widehat{b})$ are solutions of (3.1) are zeros of

$$(Q'_B \Delta Q_B)\begin{pmatrix} a \\ b \end{pmatrix} - (Q'_B \Delta Y) = 0 \Rightarrow \begin{pmatrix} a \\ b \end{pmatrix} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1}.$$

where $Q'_B = \begin{pmatrix} 1 & \ldots 1 \\ B(X_1,x) & \ldots B(X_n,x) \end{pmatrix}$
$\Delta = diag(Y_1^{-2}K(h^{-1}\delta(x,X_1)), ..., Y_n^{-2}K(h^{-1}\delta(x,X_n)))$ and $Y' = (Y_1,\ldots,Y_n)$.

Thus, we get explicitly

$$\widehat{a} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\widehat{b} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Moreover, as $\beta(x,x)=0$, we can take

$$\widehat{R}(x)=\widehat{a}=\frac{\sum_{i,j=1}^{n}V_{ij}(x)Y_j}{\sum_{i,j=1}^{n}V_{ij}(x)} \tag{2.4}$$

where

$$V_{ij}(x)=\beta(X_i,x)\left(\beta(X_i,x)-\beta(X_j,x)\right)K(h^{-1}\delta(x,X_i))K(h^{-1}\delta(x,X_j))Y_i^{-2}Y_j^{-2}$$

remarkwith the convention $0/0=0$.

1) If $b=0$, then we obtain from (3.1) the same estimator as that in Demongeot et al. (2016).

2) If $\mathcal{F}=\mathbb{R}$ and $\beta(x,x')=x-x'$, then we obtain the same local linear estimate as in Jones et al. (2008).

2.3 Pointwise almost complete convergence

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. Moreover, x denotes a fixed point in $\mathcal{F}$, N_x denotes a fixed neighborhood of x and $\phi_x(r_1,r_2)=\mathbb{P}(r_2\leq\delta(X,x)\leq r_1)$ and we put $g_\gamma(u)=\mathbb{E}\left[Y^{-\gamma}|X=u\right]$, $\gamma=1,2$.

Notice that our nonparametric model will be quite general in the sense that we will just need the following assumptions

(H1) For any $r>0$, $\phi_x(r):=\phi_x(-r,r)>0$

(H2) For all $(x_1, x_2) \in \mathcal{N}_x^2$, we have

$$|g_\gamma(x_1) - g_\gamma(x_2)| \leq C\ d^{k_\gamma}(x_1, x_2) \ \text{ for } k_\gamma > 0.$$

(H3) The function $\beta(.,.)$ is such that

$$\forall x' \in \ \mathcal{F},\ C\,|\delta(x, x')| \leq |\beta(x, x')| \leq C'\,|\delta(x, x')|.$$

(H4) K is a positive, differentiable function with support $[-1, 1]$.

(H5) The functions β and ϕ_x are such that : there exists an integer n_0, such that

$$\forall n > n_0,\ -\frac{1}{\phi_x(h)}\int_{-1}^{1}\phi_x(zh, h)\frac{d}{dz}\left(z^2K(z)\right)dz > C > 0$$

and

$$h\int_{B(x,h)}\beta(u, x)dP(u) = o\left(\int_{B(x,h)}\beta^2(u, x)\,dP(u)\right)$$

where $B(x, r) = \{x' \in \mathcal{F}/|\delta(x', x)| \leq r\}$ and $dP(x)$ is the cumulative distribution of X.

(H6) The bandwidth h satisfies

$$\lim_{n\to\infty} h = 0 \qquad \text{and} \lim_{n\to\infty}\frac{\log n}{n\phi_x(h)} = 0.$$

(H7) The function $g_2(x) > C > 0$ and the inverse moments of the response variable

$$\forall m \geq 2,\ E[Y^{-m}|X = x] < C < \infty.$$

Obviously, all these conditions are very standard and are usually assumed in this context. Indeed, the conditions (H1), (H4)-(H6) are the same as those used in

Barrientos *et al.* (2010). Assumptions (H2) and (H7) are also the same as in Demongeot et al. (2016). We point out that (H2) is a regularity condition which characterizes the functional space of our model and is needed to evaluate the bias term in the asymptotic results of this paper. While (H1) is closely linked to topological structure of the functional space of the data $\mathcal{F}$.

The following theorem gives the almost-complete convergence [1] (a.co.) of $\widehat{R}(x)$.

Theorem 9 *Under assumptions (H1)-(H7), we have*

$$|\widehat{R}(x) - R(x)| = O\left(h^{b_1}\right) + O\left(\sqrt{\frac{\log n}{n\ \ \phi_x(h)}}\right),\ a.co$$

where $b_1 = \min(k_1, k_2)$

It is clear that the proof of Theorem 9 is based on the following decomposition

$$\widehat{R}(x) - R(x) = \frac{1}{\widehat{f}(x)}\left[\widehat{g}(x) - g_2(x)g_1(x)\right] + \left[g_2^2(x) - \widehat{f}(x)\right]\frac{R(x)}{\widehat{f}(x)} \tag{2.5}$$

where

$$\widehat{g}(x) = \frac{1}{n(n-1)\mathbb{E}[W_{ij}]}\sum_{i\neq j} W_{ij}(x)Y_j \text{ and } \widehat{f}(x) \ = \frac{1}{n(n-1)\mathbb{E}[W_{ij}]}\sum_{i\neq j} W_{ij}(x)$$

1. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence of real r.v.'s ; we say that z_n converges almost completely (a.co.) to zero if, and only if, $\forall\epsilon > 0$, $\sum_{n=1}^{\infty} I\!P(|z_n| > 0) < \infty$. Moreover, let $(u_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers ; we say that $z_n = O(u_n)$ a.co. if, and only if, $\exists\epsilon > 0$, $\sum_{n=1}^{\infty} I\!P(|z_n| > \epsilon u_n) < \infty$ This kind of convergence implies both almost sure convergence and convergence in probability (see Sarda and Vieu (2000) for details).

with

$$W_{ij} = \beta(X_i, x)\left(\beta(X_j, x) - \beta(X_j, x)\right) K(h^{-1}\delta(x, X_i))K(h^{-1}\delta(x, X_j)).$$

Then Theorem 9 is a direct consequence of the following Lemmas.

Lemma 1 *Under the hypotheses (H1), (H3)-(H7), we have*

$$\left|\widehat{f}(x) - \mathbb{E}[\widehat{f}(x)]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\ \ \phi_x(h)}}\right)$$

and

$$\left|\widehat{g}(x) - \mathbb{E}[\widehat{g}(x)]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\ \ \phi_x(h)}}\right).$$

Lemma 2 *Under hypotheses (H1)-(H6), we have*

$$\left|\mathbb{E}[\widehat{f}(x) - g_2^2(x)]\right| = O(h^{k_2})$$

and

$$\left|\mathbb{E}[\widehat{g}(x)] - g_2(x)g_1(x)\right| = O(h^{k_1}).$$

Corollary 1 *Under the hypotheses of Theorem 9, we obtain :*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\widehat{f}(x) < \frac{g_2^2(x)}{2}\right) < \infty.$$

2.4 Uniform almost complete convergence

In this Section, we focus on the uniform almost complete convergence of $\widehat{R}(\cdot)$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$. Noting that, the uniform consistency has a great importance in practice as well has in theory. Indeed, the uniform convergence results are indispensable tools for data-driven bandwidth choice, testing hypotheses or in bootstrapping approach. In addition, in practice, the uniform consistency allows us to make prediction, even if the data are not perfectly observed. Recall that unlike to the multivariate case, the uniform consistency is not a simple extension of the pointwise one. In functional statistic this type of consistency requires some additional tools and topological conditions (see, Ferraty and Laksaci et al. (2010) for more discussions on this subject). So, in addition to the conditions introduced in the previous section, we need the following ones.

(U1) There exists a differentiable function $\phi(.)$, such that

$$\forall x \in S_{\mathcal{F}},\ 0 < C\,\phi(h) \leq \phi_x(h) \leq C'\,\phi(h) < \infty \ \text{ and } \ \exists \eta_0 > 0,\ \forall \eta < \eta_0,\ \phi'(\eta) < C,$$

where ϕ' denotes the first derivative of ϕ.

(U2) There exists $\eta > 0$, such that

$$\forall x, x' \in S_{\mathcal{F}}^{\eta}, \qquad |g_\gamma(x) - g_\gamma(x')| \leq C d^{k_\gamma}(x, x'),$$

where $S_{\mathcal{F}}^{\eta} = \{x \in \mathcal{F}, \exists x' \in S_{\mathcal{F}}, d(x, x') \leq \eta\}$.

(U3) The function $\beta(.,.)$ is such that

$$\forall x' \in \mathcal{F}, \quad C\,d(x, x') \leq |\beta(x, x')| \leq C'\,d(x, x')$$

and

$$\forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}, \qquad |\beta(x_1, x') - \beta(x_2, x')| \leq C\,d(x_1, x_2).$$

(U4) The kernel K satisfies (H4) and, the following Lipschitz's condition

$$|K(x) - K(y)| \leq C\,\big|x - y\big|.$$

(U5) Condition (H5) is verified for all $x \in S_{\mathcal{F}}$

(U6) The subset $S_{\mathcal{F}}$ such that

$$S_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n),$$

where $x_k \in \mathcal{F}$, $r_n = O\left(\frac{\log n}{n}\right)$ and the sequence d_n satisfies

$$\frac{(\log n)^2}{n\,\phi(h)} < \log d_n < \frac{n\,\phi(h)}{\log n} \quad \text{and} \quad \sum_{n=1}^{\infty} d_n^{(1-\beta)} < \infty \text{ for some } \beta > 1.$$

(U7) For any $m \geq 2$, $E(|Y^{-m}||X = x) < C < \infty$ for all $x \in S_{\mathcal{F}}$ and $\inf_{x \in S_{\mathcal{F}}} g_2(x) \geq C' > 0$.

Once again all these conditions are usual in this context of the uniform consistency in NFDA. In particular, condition (U1), (U5) and (U6) are the same as those considered by Ferraty, Laksaci et al. (2010). It is shown in this article that these three conditions are related to the topological structure of the functional space of the data. We found in this work several examples of functional spaces for which all these functions are explicitly known.

Theorem 10 *Under assumptions (U1)-(U7) we have*

$$\sup_{x \in S_{\mathcal{F}}} \left| \widehat{R}(x) - R(x) \right| = O(h^{b_1}) + O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right). \qquad (2.6)$$

Obviously, as for Theorem 9, Theorem 10's proof can be deduced directly from the decomposition (2.5) and from the following intermediate results which correspond to the uniform versions of Lemmas 1 2 and Corollary 1. The proofs of these results are given in the Appendix.

Lemma 3 *Under assumptions (U1), (U3)-(U7), we obtain that*

$$\sup_{x\in S_{\mathcal{F}}}\left|\widehat{f}(x)-\mathbb{E}\left[\widehat{f}(x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right)$$

and

$$\sup_{x\in S_{\mathcal{F}}}|\widehat{g}(x)-\mathbb{E}[\widehat{g}(x)]| = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right).$$

Lemma 4 *Under the assumptions (U1)- (U6), we obtain that*

$$\sup_{x\in S_{\mathcal{F}}}\left|\mathbb{E}[\widehat{f}(x)-g_2^2(x)]\right| = O(h^{k_2})$$

and

$$\sup_{x\in S_{\mathcal{F}}}|\mathbb{E}[\widehat{g}(x)]-g_2(x)g_1(x)| = O(h^{k_1}).$$

Corollary 2 *Under the assumptions of Theorem 10, we have*

$$\text{There exists } C>0 \qquad \sum_{n=1}^{\infty}\mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}}\widehat{f}(x)<C\right)<\infty.$$

2.5 Appendix

In what follows, we put, for any $x\in\mathcal{F}$, and for all $i=1,\dots,n$

$$K_i(x)=K(h^{-1}\delta(x,X_i)),\quad \delta_i(x)=\delta(X_i,x)\ \text{ and }\ \beta_i(x)=\beta(X_i,x).$$

Proof of lemma 1. It is clear that

$$\widehat{f}(x)=\underbrace{\frac{n^2h^2\phi_x^2(h)}{n(n-1)\mathbb{E}[W_{12}]}}_{A_1}\left[\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(x)Y_j^{-2}}{\phi_x(h)}\right)}_{T_1}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(x)\beta_i^2(x)Y_i^{-2}}{h^2\phi_x(h)}\right)}_{T_2}\right.$$

$$- \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)Y_j^{-2}}{h\phi_x(h)}\right)}_{T_3}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(x)\beta_i(x)Y_i^{-2}}{h\phi_x(h)}\right)}_{T_3}\Bigg]$$

and

$$\widehat{g}(x) = \underbrace{\frac{n^2h^2\phi_x^2(h)}{n(n-1)\mathbb{E}[W_{12}]}}_{A_1}\Bigg[\underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(x)Y_j^{-1}}{\phi_x(h)}\right)}_{T_4}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(x)\beta_i^2(x)Y_i^{-2}}{h^2\phi_x(h)}\right)}_{T_2}$$

$$- \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)Y_j^{-1}}{h\phi_x(h)}\right)}_{T_5}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(x)\beta_i(x)Y_i^{-2}}{h\phi_x(h)}\right)}_{T_3}\Bigg].$$

Moreover, observe that, for all, $i, j = 2, \dots, 5$

$$\begin{aligned} T_iT_j - \mathbb{E}[T_iT_j] &= (T_i - \mathbb{E}[T_i])(T_j - \mathbb{E}[T_j]) + (T_j - \mathbb{E}[T_j])\mathbb{E}[T_i] \\ &\quad + (T_i - \mathbb{E}[T_i])\mathbb{E}[T_j] + \mathbb{E}[T_i]\mathbb{E}[T_j] - \mathbb{E}[T_iT_j]. \end{aligned}$$

So, the claimed result is a consequence of the following assertions

$$\sum_n \mathbb{P}\left\{|T_i - \mathbb{E}[T_i]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right\} < \infty, \qquad \text{for } i = 1, \dots, 5, \quad (2.7)$$

$$A_1 = O(1),\ \mathbb{E}[T_i] = O(1) \ \text{ and } Cov(T_i, T_j) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right) \quad \text{for } i, j = 1, \dots, 5. \quad (2.8)$$

For (2.7) we define

$$Z_i^{l,k} = \frac{1}{h^l\phi_x(h)}\left(K_i(x)Y_i^{-k}\beta_i^l(x) - \mathbb{E}\left[K_i(x)Y_i^{-k}\beta_i^l(x)\right]\right) \quad \text{for} \quad l = 0, 1, 2, \quad \text{and} \quad k = 1, 2$$

and we apply the Bernstein's exponential inequality on $Z_i^{l,k}$. For this we must evaluate asymptotically its mth order moment. Indeed, by the Newton's binomial expansion, we obtain

$$\begin{aligned}
&\mathbb{E}\left|\left(K_i(x)Y_i^{-k}\beta_i^l(x) - \mathbb{E}\left[K_i(x)Y_i^{-k}\beta_i^l(x)\right]\right)^m\right| \\
&\quad = \mathbb{E}\left|\sum_{d=0}^{m} C_m^d \left(K_i(x)Y_i^{-k}\beta_i^l(x)\right)^d \left(\mathbb{E}\left[K_i(x)Y_i^{-k}\beta_i^l(x)\right]\right)^{m-d} (-1)^{m-d}\right| \\
&\quad \leq \sum_{d=0}^{m} C_m^d \left(\mathbb{E}\left|K_i(x)Y_i^{-k}\beta_i^l(x)\right|^d\right) \left|\mathbb{E}\left[K_i(x)Y_i^{-k}\beta_i^l(x)\right]\right|^{m-d} \\
&\quad \leq \sum_{d=0}^{m} C_m^d \mathbb{E}\left|K_1^\delta(x)\beta_1^{dl}(x)\mathbb{E}[Y_1^{-dk}|X_1]\right| \left|\mathbb{E}\left[K_1(x)\beta_1^l(x)\mathbb{E}[Y_1^{-k}|X_1]\right]\right|^{m-d}
\end{aligned}$$

where $C_{k,m} = m!/(k!(m-k)!)$.

Under condition (H7) we have

$$\mathbb{E}\left[Y_1^{-kd}|X\right] = O(1), \text{ for all } d \leq m.$$

Next, using (H3) to write that

$$h^{-l}\mathbb{E}\left[K_i(x)\beta_i^l(x)\right] \leq h^{-l}\mathbb{E}\left[K_i(x)\delta_i^l(x)\right] \leq C\phi_x(h). \tag{2.9}$$

We deduce that

$$h^{-lm}\phi_x^{-m}(h)\sum_{d=0}^{m} C_m^d \mathbb{E}\left|K_1^\delta(x)\beta_1^{dl}\right| \left|\mathbb{E}\left[K_1(x)\beta_1^l\right]\right|^{m-d} \leq C\phi_x(h)^{-m+1}.$$

Therefore, for $l = 0, 1, 2,$ and $k = 1, 2$, we obtain that

$$\mathbb{E}\left|Z_i^{l,k}\right|^m = O\left((\phi_x(h))^{-m+1}\right).$$

Thus, to achieve this proof, it suffices to use the classical Bernstein's inequality (see Corollary A8 in Ferraty and Vieu (2006), page 234), with $a_n = (\phi_x(h))^{-1/2}$ to write that

$$\forall i = 1, \dots 5 \quad \mathbb{P}\left\{ |T_i - \mathbb{E}[T_i]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}} \right\} \leq C' n^{-C\eta^2}.$$

Therefore, an appropriate choice of η permits to deduce that

$$\sum_n \mathbb{P}\left\{ |T_i - \mathbb{E}[T_i]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}} \right\} < \infty, \text{ for } i = 1, 2, 3, 4, 5.$$

Now we prove (3.4). Recall that the term A_1 is the same as in Barrientos *et al.* (2010). So, it suffices to show the other terms. To do that we evaluate

$$\mathbb{E}\left[K_i(x) Y_i^{-k} \beta_i^l(x)\right], \text{ for } l = 0, 1, 2, \text{ and } k = 1, 2.$$

As previously, we conditione on X_1 to show that, for all $l = 0, 1, 2$, and $k = 1, 2$. we have

$$\mathbb{E}\left[K_i(x) Y_i^{-k} \beta_i^l(x)\right] = O(\mathbb{E}\left[K_i(x) \beta_i^l(x)\right])$$

From (2.9) we have

$$\mathbb{E}\left[K_i(x) Y_i^{-k} \beta_i^l(x)\right] = O(h^l \phi_x(h)). \tag{2.10}$$

Therefore

$$\mathbb{E}[T_i] = O(1), \qquad \text{for } i = 1, 2, 3, 4, 5.$$

Concerning the seconde part of (3.4), we precise that all $Cov(T_i, T_j)$ for $i, j = 1, 2, 3, 4, 5.$ are of order

$$\frac{1}{nh^l \phi_x^2(h)} \mathbb{E}\left[K_i(x) Y_i^{-k} \beta_i^l(x)\right], \text{ for } l = 0, 1, 2, 3, 4 \text{ and } k = 1, 2, 3, 4.$$

Once again, we use (H3) to get

$$\frac{1}{nh^l\phi_x^2(h)}\mathbb{E}\left[K_i(x)Y_i^{-k}\beta_i^l(x)\right] = O\left(\frac{1}{n\phi_x(h)}\right), \text{ for } l = 0,1,2,3,4 \text{ and } k = 1,2,3,4.$$

Hence

$$Cov(T_i, T_j) = O\left(\frac{1}{n\phi_x(h)}\right) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right) \quad \text{for} \quad i,j = 1,2,3,4,5.$$

The latter yields to the proof of the Lemma. ■

Proof of Lemma 2. Since the observations $(X_i, Y_i)_{i=1,\dots n}$ are independent and identically distributed, then

$$\begin{aligned}
\mathbb{E}[\widehat{f}(x)] - g_2^2(x) \quad &= \frac{1}{\mathbb{E}[W_{12}]}\Big(\mathbb{E}\left[\beta_1^2(x)K_1(x)Y_1^{-2}\right]\mathbb{E}\left[K_1(x)Y_1^{-2}\right] \\
&\quad -\mathbb{E}^2\left[\beta_1(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}[W_{12}]\Big) \\
&\leq \frac{1}{\mathbb{E}[W_{12}]}\mathbb{E}\left[K_1(x)Y_1^{-2}\right]\left|\mathbb{E}\left[\beta_1^2(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[\beta_1^2(x)K_1(x)\right]\right| \\
&+\frac{1}{\mathbb{E}[W_{12}]}g_2(x)\mathbb{E}\left[\beta_1^2(x)K_1(x)\right]\left|\mathbb{E}\left[K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[K_1(x)\right]\right| \\
&+\frac{1}{\mathbb{E}[W_{12}]}\mathbb{E}\left[\beta_1(x)K_1(x)Y_1^{-2}\right]\left|\mathbb{E}\left[\beta_1(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[\beta_1(x)K_1(x)\right]\right| \\
&+\frac{1}{\mathbb{E}[W_{12}]}g_2(x)\mathbb{E}\left[\beta_1(x)K_1(x)\right]\left|\mathbb{E}\left[\beta_1(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[\beta_1(x)K_1(x)\right]\right|.
\end{aligned}$$

From conditions (H2), (H3) and (H4) we have :

$$\left|\mathbb{E}\left[\beta_1^2(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[\beta_1^2(x)K_1(x)\right]\right| \leq C\mathbb{E}\left[\beta_1^2(x)K_1(x)\right]h^{k_2}$$

and

$$\left|\mathbb{E}\left[\beta_1(x)K_1(x)Y_1^{-2}\right] - g_2(x)\mathbb{E}\left[\beta_1(x)K_1(x)\right]\right| \leq C'\mathbb{E}\left[\beta_1(x)K_1(x)\right]h^{k_2}.$$

Furthermore, we use the fact that for $k, l = 0, 1, 2$

$$\mathbb{E}\left[\beta_1^l(x) K_1(x) Y_1^{-k}\right] = O\left(h^l \phi_x(h)\right)$$

to write that

$$\left|\mathbb{E}[\widehat{f}(x)] - g_2^2(x)\right| \leq \frac{Ch^2\phi_x^2(h)}{\mathbb{E}[W_{12}]} h^{k_2} \leq C' h^{k_2}.$$

By using the same arguments as above we show that

$$|\mathbb{E}[\widehat{g}(x)] - g_2(x)g_1(x)| \leq Ch^{k_1}.$$

Consequently

$$\left|\mathbb{E}[\widehat{f}(x) - g_2^2(x)]\right| = O(h^{k_2})$$

and

$$|\mathbb{E}[\widehat{g}(x)] - g_2(x)g_1(x)| = O(h^{k_1}).$$

■

Proof of Corollary 1. It is easy to remark that :

$$|\widehat{f}(x)| \leq \frac{g_2^2(x)}{2}, \quad \text{implies that} \quad |g_2^2(x) - \widehat{f}(x)| \geq \frac{g_2^2(x)}{2}.$$

So,

$$\mathbb{P}\left(|\widehat{f}(x)| \leq \frac{g_2^2(x)}{2}\right) \leq \mathbb{P}\left(|g_2^2(x) - \widehat{f}(x)| > \frac{g_2^2(x)}{2}\right).$$

Consequently,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|\widehat{f}(x)| < \frac{g_2^2(x)}{2}\right) < \infty.$$

■

Proof of Lemma 5. The proof of this lemma is based on the same decomposition as for the proof of Lemma 1 and all it remains to show the uniform version of (2.7) and (3.4). Clearly, the last equation is a direct consequence of the assumption (U1) and the evaluation obtained in Lemma 1. While the uniform version of (2.7) is based on the following decomposition

$$\sup_{x\in S_{\mathcal{F}}}|T_k(x)-\mathbb{E}[T_k(x)]| \leq \underbrace{\sup_{x\in S_{\mathcal{F}}}|T_k(x)-T_k(x_{j(x)})|}_{F_1}+\underbrace{\sup_{x\in S_{\mathcal{F}}}|T_k(x_{j(x)})-\mathbb{E}[T_k(x_{j(x)})]|}_{F_2} + \underbrace{\sup_{x\in S_{\mathcal{F}}}|\mathbb{E}[T_k(x_{j(x)})]-\mathbb{E}[T_k(x)]|}_{F_3}. \quad k=1,2,\ldots 5.$$

We have, then, to evaluate each term F_j for $j=1,2,3$.

Firstly, we treat the terms F_1 and F_3. Since K is supported within $[-1,1]$, then we can write

$$\begin{aligned} F_1 &\leq \frac{1}{nh^l\phi_x(h)}\sup_{x\in S_{\mathcal{F}}}\sum_{i=1}^n\Big|K_i(x)Y_i^{-k}\beta_i^l(x)1\!\!1_{B(x,h)}(X_i) \\ &\quad -K_i(x_{j(x)})Y_i^{-k}\beta_i^l(x_{j(x)})1\!\!1_{B(x_{j(x)},h)}(X_i)\Big| \\ &\leq \frac{C}{nh^l\phi_x(h)}\sup_{x\in S_{\mathcal{F}}}\sum_{i=1}^n K_i(x)Y_i^{-k}1\!\!1_{B(x,h)}(X_i) \\ &\quad \times\Big|\beta_i^l(x)-\beta_i^l(x_{j(x)})1\!\!1_{B(x_{j(x)},h)}(X_i)\Big| \\ &\quad +\frac{1}{nh^l\phi_x(h)}\sup_{x\in S_{\mathcal{F}}}\sum_{i=1}^n Y_i^{-k}\beta_i^l((x_{j(x)})1\!\!1_{B(x_{j(x)},h)}(X_i) \\ &\quad \times\Big|K_i(x)1\!\!1_{B(x,h)}(X_i)-K_i(x_{j(x)})\Big|. \end{aligned}$$

The Lipschitz condition on K allows to have

$$\begin{aligned} &1\!\!1_{B(x_{j(x)},h)}(X_i)\Big|K_i(x)1\!\!1_{B(x,h)}(X_i)-K_i(x_{j(x)})\Big| \\ &\leq C\epsilon 1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+C1\!\!1_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i) \end{aligned}$$

and the Lipschitz condition on β allows in turn to write

$$\begin{aligned}
&1\!\!1_{B(x,h)}(X_i)\left|\beta_i(x)-\beta_i(x_{j(x)})1\!\!1_{B(x_{j(x)},h)}(X_i)\right| \\
&\quad\le \epsilon 1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+h1\!\!1_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i) \\
&1\!\!1_{B(x,h)}(X_i)\left|\beta_i^2(x)-\beta_i^2(x_{j(x)})1\!\!1_{B(x_{j(x)},h)}(X_i)\right| \\
&\quad\le \epsilon h 1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+h^2 1\!\!1_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i).
\end{aligned}$$

Thus,

$$F_1 \le C\sup_{x\in S_{\mathcal{F}}}\left(F_{11}+F_{12}+F_{13}+F_{14}\right),$$

where

$$F_{11}=\frac{C}{n\phi(h)}\sum_{i=1}^n Y_i^{-k}1\!\!1_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i),\quad F_{12}=\tfrac{C\epsilon}{n\phi(h)}\textstyle\sum_{i=1}^n Y_i^{-k}1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i).$$

$$F_{13}=\frac{C\epsilon}{nh\phi(h)}\sum_{i=1}^n Y_i^{-k}1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i),\quad F_{14}=\tfrac{C}{n\phi(h)}\textstyle\sum_{i=1}^n Y_i^{-k}1\!\!1_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i).$$

It suffices to apply the Bernstein's inequality on

$$\Delta_i=\begin{cases}
\dfrac{Y_i^{-k}}{\phi(h)}\sup\limits_{x\in S_{\mathcal{F}}}\left[1\!\!1_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i)\right] & \text{for } F_{11}\\[2ex]
\dfrac{\epsilon Y_i^{-k}}{h\phi(h)}\sup\limits_{x\in S_{\mathcal{F}}}\left[1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i)\right] & \text{for } F_{12} \text{ and } F_{13}\\[2ex]
\dfrac{Y_i^{-k}}{\phi(h)}\sup\limits_{x\in S_{\mathcal{F}}}\left[1\!\!1_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i)\right] & \text{for } F_{14}
\end{cases}$$

to determine the almost complete limit of F_{11}, F_{12}, F_{13} and F_{14}. Clearly, under (U6), (U7) and the second part of (U1), we have for the first and the last cases

$$\mathbb{E}[\Delta_1]=O\left(\frac{\epsilon}{\phi(h)}\right)\text{ and }\mathbb{E}|\Delta_1|^m=O\left(\frac{\epsilon}{\phi^m(h)}\right)$$

and for F_{12} or F_{13} cases

$$\mathbb{E}[\Delta_1] = O\left(\frac{\epsilon}{h}\right) \text{ and } \mathbb{E}|\Delta_1|^m = O\left(\frac{\epsilon^m}{h^m \phi^{m-1}(h)}\right),$$

which implies that

$$F_{11} = O\left(\frac{\epsilon}{\phi(h)}\right) + O_{a.co.}\left(\sqrt{\frac{\epsilon \log n}{n\,\phi(h)^2}}\right), \quad F_{14} = O\left(\frac{\epsilon}{\phi(h)}\right) + O_{a.co.}\left(\sqrt{\frac{\epsilon \log n}{n\,\phi(h)^2}}\right)$$

$$F_{12} = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right) \text{ and } F_{13} = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right).$$

Hence, by assumption (U6) we get

$$F_1 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\,\phi(h)}}\right). \tag{2.11}$$

Furthermore, since

$$F_3 \leq \mathbb{E}\left[\sup_{x \in S_{\mathcal{F}}} |T_k(x) - T_k(x_{j(x)})|\right]$$

we have also

$$F_3 = O\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right).$$

Next, we treat the term F_2^k. For all $\eta > 0$, we have that

$$\begin{aligned}
&\mathbb{P}\left(F_2 > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \\
&\quad = \mathbb{P}\left(\max_{j\in\{1,\cdots,d_n\}} |T_k(x_{j(x)}) - \mathbb{E}\left[T_k(x_{j(x)})\right]| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \\
&\quad \leq d_n \max_{j\in\{1,\cdots,d_n\}} \mathbb{P}\left(|T_k(x_j) - \mathbb{E}\left[T_k(x_j)\right]| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right).
\end{aligned}$$

Set,

$$\Delta_{ki} = \frac{1}{nh^l\phi(h)}\left(K_i(x_k)Y_i^{-k}\beta_i^l(x_k) - \mathbb{E}\left[K_i(x_k)Y_i^{-k}\beta_i^l(x_k)\right]\right).$$

By using similar ideas as in the proof of Lemma 1 we show that

$$\mathbb{E}|\Delta_{ki}|^m = O\left(\phi(h)^{-m+1}\right).$$

Once again, we apply a Bernstein-type inequality to obtain directly

$$\begin{aligned}\mathbb{P}\left(|T_i(x_k) - \mathbb{E}\left[T_i(x_k)\right]| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) &= \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Delta_{lki}\right| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \\ &\leq 2\exp\{-C\eta^2\log d_n\}.\end{aligned}$$

Thus, by choosing η such that $C\eta^2 = \beta$, we get

$$d_n \max_{k\in\{1,\cdots,d_n\}} \mathbb{P}\left(|T_i(x_k) - \mathbb{E}\left[T_i(x_k)\right]| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \leq C' d_n^{1-\beta}. \tag{2.12}$$

Since $\sum_{n=1}^{\infty} d_n^{1-\beta} < \infty$, we obtain that

$$F_2 = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right).$$

■

Proof of Lemma 6. It suffices to combine the proofs of Lemma 2, to the Lipschitz's condition, uniformly on x in $S_{\mathcal{F}}$ ■

Proof of Corollary 3.

It is clear that

$$\inf_{x\in S_{\mathcal{F}}} |\widehat{f}(x)| \leq \inf_{x\in S_{\mathcal{F}}} g_2^2(x)/2 \quad \Rightarrow \quad \sup_{x\in S_{\mathcal{F}}} |\widehat{f}(x) - g_2^2(x)| \geq \inf_{x\in S_{\mathcal{F}}} g_2^2(x)/2,$$

which implies that

$$\sum_{n=1} \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} |\widehat{f}(x)| \leq \inf_{x\in S_{\mathcal{F}}} g_2^2(x)/2\right)$$
$$\leq \sum_{n=1} \mathbb{P}\left(\sup_{x\in S_{\mathcal{F}}} |\widehat{f}(x) - g_2^2(x)| \geq \inf_{x\in S_{\mathcal{F}}} g_2^2(x)/2\right) < \infty.$$

■

Chapitre 3

Functional local linear relative regression :The missing data case

3.1 The model and its estimate

In our MAR (Missing at Random) data case we introduce a Bernoulli random variable δ such that $\delta = 1$ if Y is observed and $\delta = 0$ otherwise. This consideration implies that the variables Y and δ are independent given X. Specifically,

$$\mathbb{P}(\delta = 1|X, Y) = \mathbb{P}(\delta = 1|X) = P(X).$$

The function $P(\cdot)$ is called the conditional probability of observing Y given X. In practice, this functional operator is unknown. Now, in MAR data case, the two operators a_x and b_x are obtained by the following criterion

$$(\widehat{a}, \widehat{b}) = \arg\min_{(a,b)} \sum_{i=1}^{n} \frac{(Y_i - a - b\beta(X_i, x))^2}{Y_i^2} \delta_i K(h^{-1}\delta(x, X_i)) \tag{3.1}$$

Clearly, by a simple algebra, we prove that $(\widehat{a}, \widehat{b})$ are solutions of (3.1) are zeros of

$$(Q'_B \Delta Q_B) \begin{pmatrix} a \\ b \end{pmatrix} - (Q'_B \Delta Y) = 0 \Rightarrow \begin{pmatrix} a \\ b \end{pmatrix} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1}.$$

where $Q'_B = \begin{pmatrix} 1 & \ldots 1 \\ B(X_1, x) & \ldots B(X_n, x) \end{pmatrix}$

$\Delta = diag(Y_1^{-2}\delta_1 K(h^{-1}\delta(x, X_1)), ..., Y_n^{-2}\delta_n K(h^{-1}\delta(x, X_n)))$ and $Y' = (Y_1, \ldots, Y_n)$.

Thus, we get explicitly

$$\widehat{a} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\widehat{b} = (Q'_B \Delta Y)(Q'_B \Delta Q_B)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Moreover, as $\beta(x, x) = 0$, we can take

$$\widehat{R}(x) = \widehat{a} = \frac{\sum_{i,j=1}^{n} V_{ij}(x) Y_j}{\sum_{i,j=1}^{n} V_{ij}(x)} \tag{3.2}$$

where

$$V_{ij}(x) = \beta(X_i, x) \left(\beta(X_i, x) - \beta(X_j, x)\right) \delta_i \delta_j K(h^{-1}\delta(x, X_i)) K(h^{-1}\delta(x, X_j)) Y_i^{-2} Y_j^{-2}$$

with the convention $0/0 = 0$.

Remark 1

1) If $b = 0$, then we obtain from (3.1) the same estimator as that in Demongeot et al. (2016).

2) If $\mathcal{F} = \mathbb{R}$ and $\beta(x, x') = x - x'$, then we obtain the same local linear estimate as in Jones et al. (2008).

3.2 Pointwise almost complete convergence

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. Moreover, x denotes a fixed point in $\mathcal{F}$, N_x denotes a fixed neighborhood of x and $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$ and we put $g_\gamma(u) = \mathbb{E}\left[Y^{-\gamma}|X = u\right]$, $\gamma = 1, 2$.

Notice that our nonparametric model will be quite general in the sense that we will just need the following assumptions

(H1) For any $r > 0$, $\phi_x(r) := \phi_x(-r, r) > 0$

(H2) For all $(x_1, x_2) \in \mathcal{N}_x^2$, we have

$$|g_\gamma(x_1) - g_\gamma(x_2)| \leq C\ d^{k_\gamma}(x_1, x_2) \ \text{ for } k_\gamma > 0.$$

(H3) The function $\beta(.,.)$ is such that

$$\forall x' \in\ \mathcal{F},\ C\,|\delta(x, x')| \leq |\beta(x, x')| \leq C'\,|\delta(x, x')|.$$

(H3) The function $P(\cdot)$ is continuous on $\mathcal{N}_x$ such that $P(x') > 0$, for all $x' \in \mathcal{N}_x$.

(H4) K is a positive, differentiable function with support $[-1, 1]$.

(H5) The functions β and ϕ_x are such that : there exists an integer n_0, such that

$$\forall n > n_0,\ -\frac{1}{\phi_x(h)}\int_{-1}^{1}\phi_x(zh,h)\frac{d}{dz}\left(z^2K(z)\right)dz > C > 0$$

and

$$h\int_{B(x,h)}\beta(u,x)dP(u) = o\left(\int_{B(x,h)}\beta^2(u,x)\,dP(u)\right)$$

where $B(x,r)=\{x'\in\mathcal{F}/|\delta(x',x)|\leq r\}$ and $dP(x)$ is the cumulative distribution of X.

(H6) The bandwidth h satisfies

$$\lim_{n\to\infty} h = 0 \qquad \text{and} \lim_{n\to\infty}\frac{\log n}{n\phi_x(h)} = 0.$$

(H7) The function $g_2(x) > C > 0$ and the inverse moments of the response variable

$$\forall m \geq 2,\ E[Y^{-m}|X=x] < C < \infty.$$

Obviously, all these conditions are very standard and are usually assumed in this context. Indeed, the conditions (H1), (H4)-(H6) are the same as those used in Barrientos *et al.* (2010). Assumptions (H2) and (H7) are also the same as in Demongeot et al. (2016). We point out that (H2) is a regularity condition which characterizes the functional space of our model and is needed to evaluate the bias term in the asymptotic results of this paper. While (H1) is closely linked to topological structure of the functional space of the data $\mathcal{F}$.

The following theorem gives the almost-complete convergence [1] (a.co.) of $\widehat{R}(x)$.

Theorem 11 *Under assumptions (H1)-(H6), we have*

$$|\widehat{R}(x) - R(x)| = O\left(h^b\right) + O\left(\sqrt{\frac{\log n}{n\ \ \phi_x(h)}}\right), \ a.co.$$

Proof of Theorem 11. We consider the following decomposition :

$$\widehat{R}(x) - R(x) = \widehat{B}(x) + \frac{\widehat{M}(x)}{\widehat{f}_D(x)} + \frac{\widehat{Q}(x)}{\widehat{f}_D(x)}$$

where

$$\begin{aligned}\widehat{Q}(x) &:= \left(\widehat{f}_N(x) - \mathbb{E}[\widehat{f}_N(x)]\right) - R(x)\left(\widehat{f}_D(x) - \mathbb{E}[\widehat{f}_D(x)]\right)\\ \widehat{B}(x) &:= \frac{\mathbb{E}[\widehat{f}_N(x)]}{\mathbb{E}[\widehat{f}_D(x)]} - R(x) \quad \text{and} \quad \widehat{M}(x) := -\widehat{B}(x)\left(\widehat{R}_D(x) - \mathbb{E}[\widehat{f}_D(x)]\right)\end{aligned}$$

with

$$\widehat{f}_N(x) = \frac{1}{n\,h^2\phi_x(h)}\sum_{i,j}^{n} W_{ij}(x)Y_j \text{ and } \widehat{f}_D(x) = \frac{1}{nh^2\phi_x(h)}\sum_{i,j}^{n} W_{ij}(x).$$

Thus, the proof of Theorem 11 is based on the following intermediate results, for which the proofs are given in the Appendix.

Lemma 5 *Under the hypotheses of Theorem 11, we have that :*

$$\left|\widehat{f}_D(x) - \mathbb{E}[\widehat{f}_D(x)]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right).$$

1. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence of real r.v.'s; we say that z_n converges almost completely (a.co.) to zero if, and only if, $\forall\epsilon > 0$, $\sum_{n=1}^{\infty} I\!P(|z_n| > 0) < \infty$. Moreover, let $(u_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers; we say that $z_n = O(u_n)$ a.co. if, and only if, $\exists\epsilon > 0$, $\sum_{n=1}^{\infty} I\!P(|z_n| > \epsilon u_n) < \infty$ This kind of convergence implies both almost sure convergence and convergence in probability (see Sarda and Vieu (2000) for details).

and

$$|\widehat{f}_N(x) - \mathbb{E}[\widehat{f}_N(x)]| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right).$$

Corollary 3 *Under the hypotheses of Lemma 5, there exists a positive real C such that*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\widehat{f}_D(x) < C\right) < \infty.$$

Lemma 6 *Under assumptions (H1), (H2) and (H4) we have*

$$\left|\widehat{B}(x)\right| = O(h^b).$$

■

3.3 Uniform almost complete convergence

In this Section, we focus on the uniform almost complete convergence of $\widehat{R}(\cdot)$ on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$. Noting that, the uniform consistency has a great importance in practice as well has in theory. Indeed, the uniform convergence results are indispensable tools for data-driven bandwidth choice, testing hypotheses or in bootstrapping approach. In addition, in practice, the uniform consistency allows us to make prediction, even if the data are not perfectly observed. Recall that unlike to the multivariate case, the uniform consistency is not a simple extension of the pointwise one. In functional statistic this type of consistency requires some additional tools and topological conditions (see, Ferraty and Laksaci et al. (2010)

for more discussions on this subject). So, in addition to the conditions introduced in the previous section, we need the following ones.

(U1) There exists a differentiable function $\phi(.)$, such that

$$\forall x \in S_{\mathcal{F}},\ 0 < C\,\phi(h) \leq \phi_x(h) \leq C'\,\phi(h) < \infty \ \text{ and } \ \exists \eta_0 > 0,\ \forall \eta < \eta_0,\ \phi'(\eta) < C,$$

where ϕ' denotes the first derivative of ϕ.

(U2) There exists $\eta > 0$, such that

$$\forall x, x' \in S_{\mathcal{F}}^{\eta}, \qquad |g_\gamma(x) - g_\gamma(x')| \leq C d^{k_\gamma}(x, x'),$$

where $S_{\mathcal{F}}^{\eta} = \{x \in \mathcal{F}, \exists x' \in S_{\mathcal{F}}, d(x, x') \leq \eta\}$.

(U3) The function $\beta(.,.)$ is such that

$$\forall x' \in \ \mathcal{F}, \quad C\,d(x, x') \leq |\beta(x, x')| \leq C'\,d(x, x')$$

and

$$\forall (x_1, x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}, \qquad |\beta(x_1, x') - \beta(x_2, x')| \leq C\,d(x_1, x_2).$$

(U4) The kernel K satisfies (H4) and, the following Lipschitz's condition

$$|K(x) - K(y)| \leq C\,\big|x - y\big|.$$

(U5) Condition (H5) is verified for all $x \in S_{\mathcal{F}}$

(U6) The subset $S_{\mathcal{F}}$ such that

$$S_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n),$$

where $x_k \in \mathcal{F}$, $r_n = O\left(\frac{\log n}{n}\right)$ and the sequence d_n satisfies

$$\frac{(\log n)^2}{n\,\phi(h)} < \log d_n < \frac{n\,\phi(h)}{\log n} \quad \text{and} \quad \sum_{n=1}^{\infty} d_n^{(1-\beta)} < \infty \text{ for some } \beta > 1.$$

(U7) For any $m \geq 2$, $E(|Y^{-m}||X = x) < C < \infty$ for all $x \in S_{\mathcal{F}}$ and $\inf_{x\in S_{\mathcal{F}}} g_2(x) \geq C' > 0$.

(U8) The function $P(\cdot)$ is continuous on $S_{\mathcal{F}}$ and satisfies

$$0 < C < \inf_{x\in S_{\mathcal{F}}} P(x) < \sup_{x\in S_{\mathcal{F}}} P(x) < C' < \infty.$$

Theorem 12 *Under assumptions (U1)-(U8), we have*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{R}(x) - R(x)| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right). \qquad (3.3)$$

Proof of Theorem 12. It is clear that, the Theorem 12's proof can be deduced directly from the following uniform versions of the previous lemmas.

Lemma 7 *Under the hypotheses of Theorem 12, we have*

$$\sup_{x\in S_{\mathcal{F}}} \left|\widehat{f}_D(x) - \mathbb{E}[\widehat{f}_D(x)]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

and

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{f}_N(x) - \mathbb{E}[\widehat{f}_N(x)]| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right).$$

Corollary 4 *Under the hypotheses of Lemma 5, there exists a positive real C such that :*

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in S_{\mathcal{F}}} \widehat{f}_D(x) < C\right) < \infty.$$

Lemma 8 *Under the assumptions (H1), (H2) and (H4) we have*

$$\sup_{x\in S_{\mathcal{F}}} \left|\widehat{B}(x)\right| = O(h^b).$$

■

3.4 Appendix

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. Moreover, we put, for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$

$$K_i(x) = K(h^{-1}d(x, X_i)) \text{ and } \beta_i(x) = \beta(X_i, x).$$

Proof of lemma 5. Similarly to Barrientos et al. (2010), we write

$$\begin{aligned}
\widehat{f}_D(x) &= \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\delta_j Y_j^{-2}K_j(x)}{\phi_x(h)}\right)}_{E_2(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i^{-2}K_i(x)\beta_i^2(x)}{h^2\phi_x(h)}\right)}_{E_3(x)}\\
&- \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\delta_j Y_j^{-2}K_j(x)\beta_j(x)}{h\phi_x(h)}\right)}_{E_4(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i^{-2}K_i(x)\beta_i(x)}{h\phi_x(h)}\right)}_{E_5(x)}
\end{aligned}$$

and

$$\begin{aligned}
\widehat{f}_N(x) &= \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\delta_j Y_j^{-1}K_j(x)}{\phi_x(h)}\right)}_{T_2(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i^{-2}K_i(x)\beta_i^2(x)}{h^2\phi_x(h)}\right)}_{T_3(x)}\\
&- \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\delta_j Y_j^{-1}K_j(x)\beta_j(x)}{h\phi_x(h)}\right)}_{T_4(x)}\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i^{-2}K_i(x)\beta_i(x)}{h\phi_x(h)}\right)}_{T_5(x)}.
\end{aligned}$$

It follows that

$$\widehat{f}_D(x)-\mathbb{E}[\widehat{f}_D(x)] = (E_2(x)E_3(x) - \mathbb{E}[E_2(x)E_3(x)])-(E_4(x)E_5(x) - \mathbb{E}[E_4(x)E_5(x)])$$

$$\widehat{f}_N(x)-\mathbb{E}[\widehat{f}_N(x)] = (T_2(x)T_3(x) - \mathbb{E}[T_2(x)T_3(x)])-(T_4(x)T_5(x) - \mathbb{E}[T_4(x)T_5(x)])\,.$$

So, by a simple algebra, we write for all $i \neq j$

$$\begin{aligned} E_i(x)E_j(x) - \mathbb{E}[E_i(x)E_j(x)] &= (E_i(x) - \mathbb{E}[E_i(x)])\,(E_j(x) - \mathbb{E}[E_j(x)]) \\ &\quad + (E_j(x) - \mathbb{E}[E_j(x)])\,\mathbb{E}[E_i(x)] \\ &\quad + (E_i(x) - \mathbb{E}[E_i(x)])\,\mathbb{E}[E_j(x)] \\ &\quad + \mathbb{E}[E_i(x)]\mathbb{E}[E_j(x)] - \mathbb{E}[E_i(x)E_j(x)] \end{aligned}$$

and

$$\begin{aligned} T_i(x)T_j(x) - \mathbb{E}[T_i(x)T_j(x)] &= (T_i(x) - \mathbb{E}[T_i(x)])\,(T_j(x) - \mathbb{E}[T_j(x)]) \\ &\quad + (T_j(x) - \mathbb{E}[T_j(x)])\,\mathbb{E}[T_i(x)] \\ &\quad + (T_i(x) - \mathbb{E}[T_i(x)])\,\mathbb{E}[T_j(x)] \\ &\quad + \mathbb{E}[T_i(x)]\mathbb{E}[T_j(x)] - \mathbb{E}[T_i(x)T_j(x)] \end{aligned}$$

So, all it remains to show is that

$$\sum_n \mathbb{P}\left\{|E_i(x) - \mathbb{E}[E_i(x)]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right\} < \infty \quad \text{for} \quad i = 2,3,4,5,$$

$$\sum_n \mathbb{P}\left\{|T_j(x) - \mathbb{E}[T_j(x)]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right\} < \infty \quad \text{for} \quad j = 2,3,4,5,$$

$$\mathbb{E}[E_i(x)] = O(1) \qquad \text{and} \qquad \mathbb{E}[T_i(x)] = O(1) \quad \text{for} \quad i = 2,3,4,5,$$

and

$$Cov(E_2(x), E_3(x)) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right),\ Cov(E_4(x), E_5(x)) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right),$$

$$Cov(T_2(x), T_3(x)) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right) \text{ and } Cov(T_4(x), T_5(x)) = o\left(\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right).$$

■

Proof of the first claimed result : In order to obtain both convergence rates we apply the unbounded Bernstein's exponential inequality (see Corollary A8 in Ferraty and Vieu (2006), p. 234). We precise that, the latter is based on the asymptotic evaluation of mth order moments of the following random variables

$$Z_i^{l,k} = \frac{1}{h^l\phi_x(h)}\left(\delta_i Y_i^{-k} K_i(x)\beta_i^l(x) - \mathbb{E}\left[\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right]\right)$$

for $l = 0, 1, 2$, and $k = 1, 2$.

Notice that, by the Newton's binomial expansion, we obtain :

$$\begin{aligned}
&\mathbb{E}\left|\left(\delta_i Y_i^{-k} K_i(x)\beta_i^l(x) - \mathbb{E}\left[\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right]\right)^m\right| \\
&\quad = \mathbb{E}\left|\sum_{d=0}^{m} C_m^d \left(\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right)^d \left(\mathbb{E}\left[\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right]\right)^{m-d} (-1)^{m-d}\right| \\
&\quad \leq \sum_{d=0}^{m} C_m^d \left(\mathbb{E}\left|\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right|^d\right) \left|\mathbb{E}\left[\delta_i Y_i^{-k} K_i(x)\beta_i^l(x)\right]\right|^{m-d} \\
&\quad \leq \sum_{d=0}^{m} C_m^d \mathbb{E}\left|\delta_1 K_1^d \beta_1^{dl}(x) Y_i^{-dk}\right| \left|\mathbb{E}\left[\delta_1 K_1(x)\beta_1^l(x) Y_1^{-k}\right]\right|^{m-d}
\end{aligned}$$

where $C_{k,m} = m!/(k!(m-k)!)$.

Since the variables δ and Y are independent given X, then under Assumption (H6) we have for all $d \leq m$

$$\mathbb{E}\left[\delta Y^{-dk}|X\right] = (P(x) + o(1))\mathbb{E}\left[Y^{-dk}|X\right] \leq C.$$

Now, from Barrientos et al. (2010)

$$h^{-lm}\phi_x^{-m}(h)\sum_{d=0}^{m} C_m^d \mathbb{E}\left|K_1(x)\beta_1^{dl}(x)\right|\left|\mathbb{E}\left[K_1(x)\beta_1^{l}(x)\right]\right|^{m-d} \leq C\phi_x(h)^{-m+1}.$$

Therefore, for $l = 0, 1, 2$, and $k = 0, 1$, we obtain

$$\mathbb{E}\left|Z_i^{l,k}\right|^m = O\left((\phi_x(h))^{-m+1}\right).$$

Consequently, it suffices to apply the Corollary A8 in Ferraty and Vieu (2006), for $a_n = (\phi_x(h))^{-1/2}$, to conclude

$$\sum_n \mathbb{P}\left\{|E_i(x) - \mathbb{E}[E_i(x)]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right\} < \infty \quad \text{for} \quad i = 2, 3, 4, 5,$$

$$\sum_n \mathbb{P}\left\{|T_j(x) - \mathbb{E}[T_j(x)]| > \eta\sqrt{\frac{\log n}{n\,\phi_x(h)}}\right\} < \infty \quad \text{for} \quad j = 2, 3, 4, 5,$$

■

Proof of the second claimed result. The proof of these results are based on the fact that the observations (X_i, δ_i, Y_i) for $i = 1, \ldots, n$ are identically distributed.

Therefore

$$\begin{cases} \mathbb{E}[E_2(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-2} K_1(x)]}{\phi_x(h)}, \\ \\ \mathbb{E}[E_3(x)] = \mathbb{E}[T_3(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-2} K_1(x)\beta_1^2(x)]}{h^2\phi_x(h)}, \\ \\ \mathbb{E}[E_4(x)] = \mathbb{E}[E_5(x)] = \mathbb{E}[T_5(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-2} K_1(x)\beta_1(x)]}{h\phi_x(h)}, \\ \\ \mathbb{E}[T_2(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-1} K_1(x) Y_1]}{\phi_x(h)} \quad \text{and} \quad \mathbb{E}[T_4(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-1} K_1(x) Y_1 \beta_1(x)]}{h\phi_x(h)}. \end{cases} \tag{3.4}$$

Thus, everything is based on on the evaluation of the following quantities

$$\mathbb{E}\left[\delta_i Y_i^{-k} K_i^j(x)\beta_i^l(x)\right] \text{ for } l = 0, 1, 2, \; j = 1, 2 \text{ and } k = 1, 2.$$

Once again we use the fact that the variables δ and Y are conditionally independent with respect to the functional variable X. Therefore, for all $l = 0, 1, 2$, and $k = 0, 1$, we have

$$\mathbb{E}\left[\delta_i Y_i^{-k} K_i^j(x)\beta_i^l(x)\right] = O(\mathbb{E}\left[Y_i^{-k} K_i(x)\beta_i^l(x)\right]) = O(h^l\phi_x(h)). \tag{3.5}$$

■

Proof of the last claimed result. Similarly to the previous case it suffices to evaluate

$$\begin{cases} \mathbb{E}[E_2(x)E_3(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-4} K_1^2 \beta_1^2(x)(x)]}{nh^2\phi_x^2(h)}, \\ \\ \mathbb{E}[E_4(x)E_5(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-4} K_1^2(x) \beta_1^2(x)]}{nh^2\phi_x^2(h)}, \\ \\ \mathbb{E}[T_2(x)T_3(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-3} K_1^2(x) \beta_1^2(x) Y_1]}{nh^2\phi_x^2(h)}, \\ \\ \mathbb{E}[T_4(x)T_5(x)] = \dfrac{\mathbb{E}[\delta_1 Y_1^{-3} K_1^2(x) \beta_1^2(x) Y_1]}{nh^2\phi_x^2(h)}. \end{cases}$$

Using the same arguments as for proving (3.5) we show that

$$\mathbb{E}\left[\delta_i Y_i^{-k} K_i^2(x) Y_i^k \beta_i^l(x)\right] = O(h^l \phi_x(h)).$$

It follows that

$$\begin{cases} \mathbb{E}[E_2(x)E_3(x)] = O\left(\dfrac{1}{n\phi_x(h)}\right), \ \mathbb{E}[E_4(x)E_5(x)] = O\left(\dfrac{1}{n\phi_x(h)}\right), \\ \\ \mathbb{E}[T_2(x)T_3(x)] = O\left(\dfrac{1}{n\phi_x(h)}\right), \ \mathbb{E}[T_4(x)T_5(x)] = O\left(\dfrac{1}{n\phi_x(h)}\right). \end{cases}$$

Now, by combining this last to (3.4), we get

$$Cov(E_2(x), E_3(x)) = o\left(\sqrt{\frac{1}{n\,\phi_x(h)}}\right), \qquad Cov(E_4(x), E_5(x)) = O\left(\sqrt{\frac{1}{n\,\phi_x(h)}}\right), \tag{3.6}$$

$$Cov(T_2(x), T_3(x)) = O\left(\sqrt{\frac{1}{n\,\phi_x(h)}}\right) \quad \text{and} \quad Cov(T_4(x), T_5(x)) = O\left(\sqrt{\frac{1}{n\,\phi_x(h)}}\right).$$

which completes the proof of this lemma.

■

Proof of Lemma 6. We write

$$\widehat{B}(x) = \frac{\mathbb{E}\left[\widehat{f}_N(x)\right] - R(x)\mathbb{E}\left[\widehat{f}_D(x)\right]}{\mathbb{E}\left[\widehat{f}_D(x)\right]} = \frac{\mathbb{E}\left[\widehat{f}_N(x)\right] - R(x)\mathbb{E}\left[\widehat{f}_D(x)\right] + g_2(x) - g_2(x)}{\mathbb{E}\left[\widehat{f}_D(x)\right]}.$$

Thus all it remain show that

$$\left|\mathbb{E}[\widehat{f}_D(x) - g_1(x)]\right| = O(h^{k_2})$$

and

$$|\mathbb{E}[\widehat{f}_N(x)] - g_2(x)| = O(h^{k_1}).$$

The proof is based on the same arguments as those used in the forst chapter. It suffices to observe that

$$\left|\mathbb{E}\left[\beta_1(x)K_1(x)Y_1^{-\gamma}\right] - g_\gamma(x)\mathbb{E}\left[\beta_1(x)P(X_1)K_1(x)\right]\right| \le C'\mathbb{E}\left[\beta_1(x)K_1(x)\right]h^\gamma.$$

Furthermore, we use the fact that for $\gamma, l = 0, 1, 2$

$$\mathbb{E}\left[\beta_1^l(x)K_1(x)Y_1^{-\gamma}\right] = O\left(h^l\phi_x(h)\right)$$

to prove the claimed result ■

Proof of Lemma 7. The proof of this lemma is based on same ideas as for Lemma 5, since it suffices to prove the uniformity of its claimed results. Concerning (3.4) and (3.6), we use the fact that $\phi_x(h) > C\phi(h)$ to give the uniform limit. Now, all it remains to show is that

$$\sup_{x\in E_{\mathcal{F}}} |E_k(x) - \mathbb{E}[E_k(x)]| = O\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right), \text{ a.co., for } k = 2, 3, 4, 5. \qquad (3.7)$$

$$\sup_{x\in E_{\mathcal{F}}} |T_k(x) - \mathbb{E}[T_k(x)]| = O\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right), \text{ a.co., for } k = 2,3,4,5 \tag{3.8}$$

The proof of (3.5) follows the same ideas as in Ferraty et al. (2010). Indeed, we define $j(x) = \arg\min_{j\in\{1,2,\dots,d_n\}} |\varrho(x, x_k)|$, and we consider the following decomposition

$$\sup_{x\in S_{\mathcal{F}}} |E_k(x) - \mathbb{E}[E_k(x)]| \leq \underbrace{\sup_{x\in S_{\mathcal{F}}} |E_k(x) - E_k(x_{j(x)})|}_{F_1} + \underbrace{\sup_{x\in S_{\mathcal{F}}} |E_k(x_{j(x)}) - \mathbb{E}[E_k(x_{j(x)})]|}_{F_2} + \underbrace{\sup_{x\in E_{\mathcal{F}}} |\mathbb{E}[E_k(x_{j(x)})] - \mathbb{E}[E_k(x)]|}_{F_3}.$$

and

$$\sup_{x\in E_{\mathcal{F}}} |T_i(x) - \mathbb{E}[T_i(x)]| \leq \underbrace{\sup_{x\in E_{\mathcal{F}}} \left|T_i(x) - T_i(x_{j(x)})\right|}_{E_1} + \underbrace{\sup_{x\in S_{\mathcal{F}}} \left|T_i(x_{j(x)}) - \mathbb{E}[T_i(x_{j(x)})]\right|}_{E_3} + \underbrace{\sup_{x\in S_{\mathcal{F}}} \left|\mathbb{E}[T_i(x_{j(x)})] - \mathbb{E}[T_i(x)]\right|}_{E_2}.$$

Concerning the first term F_1 we have

$$\begin{aligned} F_1 \leq\ & \frac{C(k-2)}{nh^l\phi_x(h)} \sup_{x\in E_{\mathcal{F}}} \sum_{i=1}^{n} \delta_i Y_i^{-k} K_i(x) 1\!\!1_{B(x,h)}(X_i) \times \left|\beta_i^l(x) - \beta_i^l(x_{j(x)}) 1\!\!1_{B(x_{j(x)},h)}(X_i)\right| \\ & + \frac{1}{nh^l\phi_x(h)} \sup_{x\in E_{\mathcal{F}}} \sum_{i=1}^{n} \delta_i \beta_i^l((x_{j(x)}) 1\!\!1_{B(x_{j(x)},h)}(X_i) Y_i^{-k} \times \left|K_i(x) 1\!\!1_{B(x,h)}(X_i) - K_i(x_{j(x)})\right|. \end{aligned}$$

Under, the Lipschitz condition, on the kernek K, we obtain

$$\begin{aligned} & 1\!\!1_{B(x_{j(x)},h)}(X_i) \left|K_i(x) 1\!\!1_{B(x,h)}(X_i) - K_i(x_{j(x)})\right| \\ & \qquad \leq C\epsilon 1\!\!1_{B(x,h)\cap B(x_{j(x)},h)}(X_i) + C 1\!\!1_{B(x_{j(x)},h)\cap \overline{B(x,h)}}(X_i). \end{aligned}$$

Furthermore, by the Lipschitz condition on β, we obtain

$$\begin{aligned}
&\mathbb{1}_{B(x,h)}(X_i)\left|\beta_i(x)-\beta_i(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h)}(X_i)\right| \\
&\qquad\leq \epsilon\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+h\mathbb{1}_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i) \\
&\mathbb{1}_{B(x,h)}(X_i)\left|\beta_i^2(x)-\beta_i^2(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h)}(X_i)\right| \\
&\qquad\leq \epsilon h\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+h^2\mathbb{1}_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i)
\end{aligned}$$

which allows to write that

$$\begin{aligned}
&\mathbb{1}_{B(x,h)}(X_i)\left|\beta_i^l(x)-\beta_i^l(x_{j(x)})\mathbb{1}_{B(x_{j(x)},h)}(X_i)\right| \\
&\qquad\leq \epsilon h^{k-3}\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i)+h^l\mathbb{1}_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i).
\end{aligned}$$

Therefore

$$F_1 \leq C\sup_{x\in E_{\mathcal{F}}}\left(F_{11}^k+F_{12}+F_{13}^k+F_{14}\right),$$

where

$$\begin{aligned}
F_{11}^k &= \frac{C}{n\phi(h)}\sum_{i=1}^n Y_i^{-k}\mathbb{1}_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i), \\
F_{12} &= \frac{C\epsilon}{n\phi(h)}\sum_{i=1}^n Y_i^{-k}\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i). \\
F_{13}^k &= \frac{C\epsilon}{nh\phi(h)}\sum_{i=1}^n Y_i^{-k}\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i). \\
F_{14} &= \frac{C}{n\phi(h)}\sum_{i=1}^n \mathbb{1}_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i).
\end{aligned}$$

Now, we apply a standard inequality for sums of bounded random variables (cf.

Corollary A.8 in Ferraty and Vieu, 2006) with Z_i is identified such that :

$$Z_i = \begin{cases} \frac{1}{\phi(h)} Y_i^{-k} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x,h)\cap\overline{B(x_{j(x)},h)}}(X_i)\right] & \text{for } F_{11}^k \\ \\ \frac{\epsilon}{h\phi(h)} Y_i^{-k} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x,h)\cap B(x_{j(x)},h)}(X_i)\right] & \text{for } F_{12} \text{ and } F_{13}^k \\ \\ \frac{1}{\phi(h)} Y_i^{-k} Y_i^{-k} \sup_{x\in S_{\mathcal{F}}} \left[\mathbb{1}_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i)\right] & \text{for } F_{14} \end{cases}$$

So, under the second part of (U1), we have

$$F_1 \;=\; O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right). \tag{3.9}$$

Concerning F_3, we use the fact that

$$F_3 \leq \mathbb{E}\left[\sup_{x\in S_{\mathcal{F}}} |E_k(x) - E_k(x_{j(x)})|\right]$$

to get

$$F_3 \;=\; O\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right).$$

By the same arguments, we obtain

$$E_1 = O_{a.co.}\left(\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \text{ and } E_2 = O\left(\sqrt{\frac{\log d_n}{n\,\phi(h)}}\right). \tag{3.10}$$

Therefore, for $i = 2, 3, 4, 5$, and $C\eta^2 = \beta$, we have

$$d_n \max_{k\in\{1,\dots,d_n\}} \mathbb{P}\left(|E_i(x_k) - \mathbb{E}\left[E_i(x_k)\right]| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right)$$
$$= I\!\!P\left(\frac{1}{n}\left|\sum_{i=1}^{n} \Delta_{lki}\right| > \eta\sqrt{\frac{\log d_n}{n\phi(h)}}\right) \leq d_n^{1-\beta}.$$

and

$$d_n \max_{k\in\{1,\dots,d_n\}} \mathbb{P}\left(|T_i(x_k) - \mathbb{E}T_i(x_k)| > \eta\sqrt{\frac{\log d_n}{n\ \phi(h)}}\right) \leq C' d_n^{1-\beta}.$$

The proof of this Lemma is now finished.

■

Proof of Corollary 4. It easy to show that if

$$\inf_{x\in S_{\mathcal{F}}} \widehat{f}_D(x) \leq \frac{1}{2} \Rightarrow \exists x \in S_{\mathcal{F}},\ \text{such that } 1 - \widehat{f}_D(x) \geq \frac{1}{2} \Rightarrow \sup_{x\in S_{\mathcal{F}}} |1 - \widehat{f}_D(x)| \geq \frac{1}{2}.$$

From Lemma 5, we get

$$\mathbb{P}\left(\inf_{x\in E_{\mathcal{F}}} \widehat{f}_D(x) \leq \frac{1}{2}\right) \leq \mathbb{P}\left(\sup_{x\in S_{\mathcal{F}}} |1 - \widehat{f}_D(x)| > \frac{1}{2}\right).$$

It follows that

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in E_{\mathcal{F}}} |\widehat{f}_D(x)| < \frac{1}{2}\right) < \infty.$$

■

Proof of lemma 8. The proof follows the same lines allowing to show Lemma 6 combined with the uniformity, of x in $E_{\mathcal{F}}$, of the Lipschitz assumption. That completes the proof of this Lemma.

■

Chapitre 4

Dissections and Applications

4.1 Monte Carlo study

In this section we carry out a numerical simulation to evaluate the performance of the proposed method for finite samples data. The aims of this study are :

- To show how we can implement easily and rapidly our estimator (local linear relative error regression (L.L.R.E.R.)) in practice.
- To compare the efficiency of the developed estimator to other regression models such as the classical regression (C.R.), the local linear regression (L.L.R) and to the local constant relative-error regression (L.C.R.E.R.).

Recall that the C.R. model was introduced in NDFA by Ferraty and Vieu (2006) and its estimator is defined by

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

While, for L.L.R. model, we use the estimator of Barrientos *et al.* (2010) defined by

$$\frac{\sum_{i,j=1}^{n} W_{ij}Y_j}{\sum_{i,j=1}^{n} W_{ij}} \quad \text{where} \quad W_{ij} = \beta(X_i,x)\,(\beta(X_i,x)-\beta(X_j,x))\,K(h^{-1}\delta(x,X_i))K(h^{-1}\delta(x,X_j)).$$

The function L.C.R.E.R. is more recent than other regression models. It was introduced by Demongeot et al. (2016) and is defined by

$$\frac{\sum_{i=1}^{n} Y_i^{-1}K(h^{-1}d(x,X_i))}{\sum_{i=1}^{n} Y_i^{-2}K(h^{-1}d(x,X_i))}.$$

In this illustration example, the random variables $(X_i, Y_i)_{i=1,\dots,n}$ are generated according to the following formula

$$Y_i = R(X_i) + \varepsilon_i, \quad i = 1,\dots n, \quad \varepsilon_i \rightsquigarrow \exp(.5)$$

where the explanatory curves are defined by

$$X_i(t) = \sin(4(b_i - t)\pi) + a_i t^2\,, \;\; \forall t \in (0,1) \text{ and } i = 1,\dots,n$$

with b_i is distributed as $\mathcal{N}(0,1)$, while the n random variables a_i's are generated according to a $\mathcal{N}(4,3)$ distribution. All the curves X_i's are discretized on the same grid generated from 100 equispaced measurements in $(0,1)$

The operator R, is defined by

$$R(X_i) = \exp\left\{-\int_0^1 \frac{dt}{1+X_i^2(t)}\right\} \text{ for } i = 1,\dots,n. \tag{4.1}$$

For practical purposes, we select the smoothing parameters, for the four regressions models, by the local cross-validation method on the number of nearest neighbors

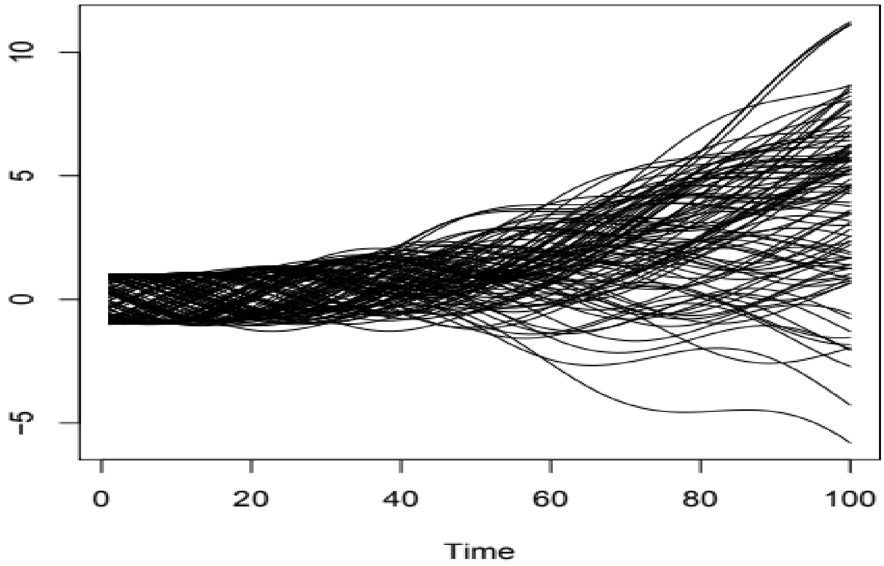

FIGURE 4.1 – A sample of 100 curves

and we use the quadratic kernel of support $[-1, 1]$ defined by

$$K(x) = \frac{3}{4}(1 - x^2) \text{ if } x \in [-1, 1].$$

Notice that the infinite dimensional structure of the data is an interesting source of information, which brings many opportunities for all statistical analysis. Thus, the choice of the parameters of the topological structure is an important point for insuring a good behavior of the functional regression analysis. We refer to Ferraty and Vieu (2006) for more discussions on the choice of these parameters. Nevertheless, here, the curves X_i are smooth enough to consider the following types of functions

$$\delta(x, x') = \left(\int_0^1 (x^{(i)}(t) - x'^{(i)}(t))^2 dt \right)^{1/2} \text{ and } \beta(x, x') = \int_0^1 \theta(t)(x^{(j)}(t) - x'^{(j)}(t))dt$$

where $x^{(i)}$ denotes the ith derivative of the curve x and θ is the eigenfunction of the empirical covariance operator $\frac{1}{n}\sum_{i=1}^{n}(X_i^{(j)} - \overline{X^{(j)}})^t ((X_i^{(j)} - \overline{X^{(j)}}))$ associated to the q-greatest eigenvalues.

In this simulation study, we have worked with several values of i, q and j, but, for the sake of shortness, we present only the results of the case where $i = 2$, $j = 1$ and $q = 3$. In order to test the rapidity of our estimate we compare the four regression models over various sample size $n = 100, 200, 500, 1000$. Furthermore, as discussed in the introduction, the main advantage of the relative-error regression is its resistance to the presence of outliers. So, we emphasize this feature by comparing these regression models in both cases (absence and presence of outliers). More precisely, we artificially introduced outliers by multiplying, by 100, $k\%$ of the response variables Y. In the following table we report the Mean Absolute Error (MAE-error) defined by

$$\frac{1}{n}\sum_{i=1}^{n}|Y_i - m(X_i)|$$

where $m(\cdot)$ means C.R., L.L.R., L.C.R.E.R. or L.L.R.E.R.

n	k%	C.R.	L.L.R.	L.C.R.E.R.	L.L.R.E.R
n=100	0%	1.54	1.05	1.18	0.96
	5%	14.90	9.56	4.54	3.40
	10%	171.58	139.20	15.55	13.47
	20%	784.23	908.39	22.75	17.16
n=200	0%	0.67	0.80	0.72	0.53
	5%	12.05	7.22	3.22	2.02
	10%	114.78	103.18	10.19	7.54
	20%	756.68	603.89	13.45	11.97
n=500	0%	0.42	0.59	0.55	0.44
	5%	6.78	4.21	1.52	1.33
	10%	90.39	84.25	12.03	6.88
	20%	276.62	154.38	09.65	8.78
n=1000	0%	0.33	0.22	0.21	0.20
	5%	3.52	2.25	0.67	0.11
	10%	80.81	58.37	7.56	2.38
	20%	188.38	133.71	8.95	6.96

Table 1 The MAE-error of the estimates .

It appears clearly that the efficiency of these functional varies with respect to k. However, the two L.C.R.E.R and L.L.R.E.R models are more stable than C.R. and L.L.R.. In the sense that the MAE of C.R. and L.L.R. increase substantially with respect to values of k, whereas the variability of this error in L.C.R.E.R and L.L.R.E.R models is very low. We point out that the error for the L.L.R.E.R and L.L.R. is smaller than for C.R. and L.C.R.E.R in the most cases.

n	C.R.	L.L.R.	L.C.R.E.R.	L.L.R.E.R
n=100	0.18	0.41	0.30	0.56
n=200	0.83	1.27	1.03	1.42
n=500	1.59	2.16	1.87	2.63
n=1000	2.25	3.12	2.48	3.48

Table 2 The computational time without perturbation.

In the Table 2 we compare the computational time (in seconds) for the four regression models. It is clear that the difference is not very significative. All these models are faster even if the sample sizes is large.

4.2 Conclusion and prospects

In this contribution we have modeled the co-variability between a functional variable X and a scalar variable Y by minimizing the mean squared relative error. The main feature of this loss function is that it takes into account the size of each observation, unlike to the least square loss function for which e all observations have the same weight. From a theoretical point of view, our approach has the same asymptotic properties as the least square regression. In particular it has the same (pointwise and uniform) almost complete convergence rate. On the other hand, from a practical point of view our approach has more advantage than the classical regression. Typically, the local linear relative error regression is more efficiency than other competitive models, such as the kernel relative regression, the local linear regression or the local constant regression. While the robustness of this model is due to the considered loss function, the superiority in precision is justified by the small bias of the local linear approach. It should be noted that the theoretical quantification of the gain in the bias term requires some additional tools and conditions. Precisely, it is obtained by the determination of the exact asymptotic mean squared error. This question is a natural prospect of the present work. Furthermore, the approach developed in this paper generates other interesting perspectives. Specifically, it can be adapted to nonparametric as well as parametric or semiparametric regression models such as the linear model, the k-NN method or the partial linear modeling.

Bibliographie

[1] Baìllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response, *Journal of Multivariate Analysis*, **100**, Pages 102–111.

[2] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally modelled regression and functional data. *J. of Nonparametric Statistics*, **22**, No. 5, Pages 617–632.

[3] Biau, G. (2003). Spatial kernel density estimation. *Math. Methods Statist.***12**,371–390.

[4] Bosq, D., Lecoutre, J.P (1987) . Théorie de l'estimation fonnctionnelle. *Economica* .

[5] Bosq, D. (2000). *Linear Processes in Function Spaces : Theory and applications.* Lecture Notes in Statistics, **149**, Springer.

[6] Cardot, H., Ferraty, F., Sarda, P (1999). Linear Functional Model. *Statistic and Probability*,Letters, **45**,11–22 .

[7] Cardot, H. and Sarda, P (2008). Varying-coefficient functional linear regression models. *Comm. Statist. Theory Methods* , **37** p 3186–3203.

[8] Clarkson Douglas B., Fraley C, Gu C, Ramsay and Ji (2005) . S+Functional Data Analysis . *Springer* .

Cressie, N. A. C. (1991). Statistics for Spatial Data. *Wiley, New York.*

[9] Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *Pub. Inst. Stat. Univ. Paris*, **3**, Pages 27–42.

[10] Dabo-Niang, S. et Laksaci, A (2007). Propriétés asymptotiques d'un estimateur Ã noyau du mode conditionnel pour variable explicative fonctionnelle.(French) *Ann. I.S.U.P* **51**, 27-42.

[11] De Gooijer, J. and Gannoun, A. (2000). Nonparametric conditional predictive regions for time series. *Comput. Statist. Data Anal.*, **33**, Pages 259–257.

[12] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2010) Local linear estimation of the conditional density for functional data. *C. R., Math., Acad. Sci. Paris*, **348**, Pages 931-934.

[13] Demongeot, J., Laksaci, A., Madani, F. and Rachdi,M. (2011). Functional data : Local linear estimation of the conditional density and its application. *Technical report,*, page 21.

[14] Demongeot,J. Hamie,A. Laksaci,A. Rachdi,M. (2016). Relative-error prediction in nonparametric functional statistics : Theory and practice . J. Multivariate Anal. 146, 261–268, .

[15] El Methni, M. and Rachdi, M. (2011). Local weighted average estimation of the regression operator for functional data. *Commun. Stat., Theory and Methods*, Volume '0, Pages "3141–3153.

[16] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data.

J. Nonparametr. Stat., **20**, Pages 3–18.

[17] Fan, J. (1992). Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, Pages 998–1004.

[18] Fan, J. and Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *Annal of Statistics.***20**,2008–2036.

[19] Fan, J. and Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting : Variable bandwidth and spatial adaptation. *J. Roy. Statist. Soc. Ser.*B,**57** ,371–394.

[20] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications.* London, Chapman & Hall.

[21] Fan, J. and Yao, Q. (2003). *Nolinear Time Series : Nonparametric and Parametric Methods.* Springer-Verlag, New York.

[22] Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P.(2010). Rate of uniform consistency for nonparametric estimates with functional variables. *Journal of statistical planning and inference*, **140**, Pages 335–352.

[23] Ferraty, F., Laksaci, A. and Vieu, P. (2005). Functional times series prediction via conditional mode. *C. R., Math., Acad. Sci. Paris*, **340**, Pages 389–392.

[24] Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Stat. Inference Stoch. Process.,* **9**, Pages 47–76.

[25] Ferraty .F and Vieu .P (2002). The functional nonparametric model and application to spectrometric data. *Computational Statistics.*

[26] Ferraty .F and Vieu .P (2003) . Curve discrimination : A nonparametric functional approach. *Computational Statistics and Data Analysis***44**,545–564..

[27] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis. Theory and Practice.* Springer Series in Statistics. New York.

[28] Francisco-Fernándeza and Vilar-Fernàndez, J. M (2001). Local Polynomial Regression Estimation With Correlated Errors. *Communications in Statistics - Theory and Methods .*

[29] Hall, P., Wolef, R. C. L. and Yao, Q. (1999). Methods for estimating a conditional distribution function. *J. Am. Statist. Assoc.,* **94**, Pages 154-163.

[30] Hallin,M. Lu,Z.Tran,L T (2004). Local Linear Spatial Regression. *The Annals of Statistics* **6**, 2469–2500.

[31] Hafen, R. P. (2010) *Local regression models : Advancements, applications, and new methods.* Thesis (Ph.D.)

[32] R. A. Irizarry (2001) Local Regression With Meaningful Parameters Local Regression With Meaningful Parameters pages 72-79 *The American Statistician* Volume 55, Issue 1, 2001 pages 72-79

[33] Laksaci, A. Convergence en moyenne quadratique de l'estimateur Ã noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Pub. Inst.Stat. Univ. Paris,***3**, Pages 69–80.

[34] Loader, C. (2004) *Smoothing : local regression techniques.* Handbook of computational statistics, 539–563, Springer, Berlin,

[35] Lu, Z. and Chen, X. (2004). Spatial kernel regression estimation : Weak consistency. *Statist.Probab. Lett.***68** ,125–136.

[36] Marron, J.S. and M.P. Wand (1992). Exact mean integrated squared error. *Ann. Statist.* **20**, Pages 1919-1932.

[37] Masry, E. and Fan, J. (1997). Local polynomial estimation of regression functions for mixing processes. *Scand. J. Statist.***24**, 165–179.

[38] Mechab,W and Laksaci,A. Nonparametric relative regression for associated random variables . Metron 74, 75–97 (2016).

[39] Ouassou, I. and Rachdi, M. (2010). Stein type estimation of the regression operator for functional data. *Advances and Applications in Statistical Sciences*, **1**, No 2, Pages 233–250.

[40] E. Rafajfowicza and R. Schwabeb (2003) Experimental Design for (Semi-)Local Regression Experimental Design for (Semi-)Local Regression *Communications in Statistics - Theory and Methods* Volume 32, Issue 5, pages 1035-1055

[41] Ramsay J, Silverman B (1997). Functional data analysis. *Springer* .

[42] Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis. Methods and case studies.* Springer Series in Statistics. New York.

[43] Ripley, B. (1981). Spatial Statistics. *Wiley, New York.*

[44] Tony ,T C and Peter ,H (2006). Prediction in functional linear regression. *The Annals of Statistics* .

[45] Vieu, Ph (1991). Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.* **39**, 324-347.

[46] Vidaurre, D. Bielza, C. Larranaga, (2012) Pedro Lazy lasso for local regression. *Comput. Statist.* 27 531 -550.

[47] Welsh, A. H. ; Yee, T. W. (2006) Local regression for vector responses. *J. Statist. Plann. Inference* 136 ? 3007–3031.

[48] K. Xiuli ; Braun, W. J. Stafford, James E (2011). Local regression when the responses are interval-censored. it J. Stat. Comput. Simul. 81, 1247–1279.

[49] Yu,K and Jones,M.C (1998). Local linear quantile regression . *Journal of the American Statistical Association*,**93**,228–237.