Ministère de l'Enseignement Supérieur et la Recherche Scientifique
Université Djillali Liabès de Sidi Bel Abbès
Faculté des Sciences
Département des Mathématiques

THÉSE DE DOCTORAT

Discipline : Mathématiques

Option : Statistique

Présentée par
Hachemi Nawel

L'estimation Semi-paramétrique de la fonction de répartition conditionnelle à variable explicative fonctionnelle

Composition du Jury

Président	Mr.A.Gheriballah	Professeur. Université de S.B.A
Encadreur	Mr. Laksaci Ali	Professeur. Université de S.B.A
Examinateurs	Mr. Attouch. M	Maître de conférence A. Université de S.B.A
	Mr. A. Kandouci	Maître de conférence A. Université de S.B.A
	Mr. T. Guendouzi	Maître de conférence A. Université de S.B.A

Je dédie ce travail à toute bonne volonté humaine
et à mes parents

Remerciement

Je remercie Dieu, le tout puissant de m'avoir donné la force et la patience pour achever ce travail et de continuer mes études.

Je tiens à remercie monsieur le professeur Ali Laksaci pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral, pour ces multiples conseils et pour toutes les heures qu'il a consacrées à diriger cette recherche. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et son respect sans faille des délais serrés de relecture des documents que je lui ai adressés.

Je voudrais aussi remercier vivement monsieur le Professeur Gheriballah, je tiens à exprimer ma profond gratitude pour l'honneur singulier qu'il m'accorder en acceptant d'être président de ce jury.

Je tiens également à remercie monsieur Attouch qui s'est naturellement associé à cette équipe et qui, malgré ses nombreuses occupations, à accepté d'être le jury de cette thèse.

Pour l'intérêt qu'il a bien voulu accorder à mes recherches en acceptant de participer au jury, je tiens à exprimer mes vifs remerciements à monsieur Kandouci abdeldjabbar.

Je désire remercier vivement monsieur Guendouzi Toufik pour la confiance d'avoir accepté de faire partie de ce jury.

Je dois un grand merci à monsieur Cheiker Mazouar pour leur intérêt qu'ils ont porté à mon travail.

J'adresse mes remerciements à tout ceux qui ont participé de prés où de loin à bonne réalisation de la soutenance.

Table des matières

Résumé

Dans cette thèse, nous nous proposons d'étudier l'estimation semi-paramé trique de la fonction de répartition conditionnelle comme outil réducteur de la dimension. On va mettre l'accent sur l'effet de la dimension en estimation fonctionnelle et on a constate que la dimension à un effet inverse sur la vitesse de convergence. En effet, on suppose qu'on dispose d'une variable aléatoire réelle (réponse), notée Y et d'une variable fonctionnelle (explicative), notée X. Plus précisèment, Nous nous intéressons au problème de la prévision à partir d'un espace semi-métrique dans un espace de dimension infinie, et nous cherchons à développer des alternatives à la méthode linéaire locale qui est plus générale par rapport à la méthode du noyau.

Tout d'abord, nous considérons que les observations sont indépendants et identiquement distribuées i.i.d.[1] *Dans ce contexte, on a d'une part se place dans un cadre fonctionnel et l'autre part on introduit des hypothèses moins restrictives que celle utilisent habituellement dans le cadre réelle ou vectorielle. On fixe comme objectif la convergence presque complète en précisant sa vitesse de convergence.*

Ensuite, on généralise ces derniers résultats au cas dépendant sous des conditions un plus fort pour eviter le terme de covariance, cependant, la vitesse de convergence sera lente par rapport au cas indépendant, considérons le cas où les observations sont fortement mélangentes en donnant le taux de la convergence.

1. indépendants et identiquement distribuées.

Notre étude est liée à l'effet de la dimension en estimation fonctionnelle. Ainsi, l'utilisation de l'espace semi-métrique donne une forte concentration de la mesure de probabilité de la variable fonctionnelle sur des petites boules. Plus précisément, des hypothèses portent sur les probabilités de petites boules nous permettent de proposer une solution originale au problème du fléau de la dimension et ainsi de généraliser à la dimension infinie de nombreux résultats existant dans le cas multivarié.

Summary

In this thesis, we study the problem of a semiparametric estimation of the conditional distribution function as a tool reduction of the dimension. We will focus on the effect of the dimension of functional estimation and it was found that a dimension are inverse effect on the rate of convergence. Indeed, we assume that we have a real random variable (named response variable) noted Y *and a functional variable (explanatory variable) noted* X. *We are interested the problem of prediction from a semimetric space and we search to develop by a local linear method which is more general than the kernel method.*

Firstly, we consider a sequence of identically distributed observations. In this context, we have on the one hand a place of functional setting and the other's are introduced by a less restrictive assumptions than that usually used in the real context. We focus on the almost complete convergence rate of the estimator.

Then, we generalize the previous result in the dependant case under a stronger condition to avoid the covariance term. However, the rate of convergence will be slow compared to the independent case and we consider the observations are mixing by giving the convergence rate.

O study is related to the effect of the functional dimension. Then, the use of the semimetric space give a high concentration of the probability measure of the rate of convergence. This allows to increase the rate of convergence while preserving the nature of the functional data. More precisely, these ideas are used to give a statistical solution to curse of dimension

and to generalize to infinite dimension many asymptotic results existing in the multivariate case.

Liste des travaux

Publication

1. Laksaci, A. Hachemi, N. (2012) : Note on the functional linear estimate of conditional cumulative distribution function. Journal of Probability ans Statistical Science ; 10(2), 153-160.

Liste des notations

Ensembles, Fonctions, Nombres

$(\Omega, \mathcal{A}, \mathbb{P})$	L'espace probabilité où les variables aléatoires sont considérées.
$\mathbb{P}$	La mesure attachée à l'espace probabilité $(\Omega, \mathcal{A})$.
E, $\mathcal{F}$ $\mathcal{H}$	Les espaces fonctionnelles.
X	La variable explicative.
Y	La variable de réponse.
ε	La variable aléatoire centrée et réduite indépendant de x.
$d(.,.)$	La distance entre les vecteurs de l'espace semi-métrique.
(E, d)	L'espace fonctionnel et sa semi-métrique.
$\beta(.,.)$	Fonction connue définie sur $\mathcal{F}$ dans R
$\delta(.,.)$	Le diamètre de sous-ensembles $\mathcal{A}$ sur l'espace (E, d).
$\mathsf{1}_A(.)$	La fonction indicatrice qui vaut 1 sur l'ensemble $\mathcal{A}$ et 0 ailleurs.
$r(.)$	La fonction de régression.
$\mathbb{P}(./\chi = x)$	Mésuré la probabilité conditionnel.
$f^x(.)$	La densité conditionnel de Y sachant x.
$F^x(.)$	La fonction de répartition conditionnel de Y sachant X.
θ	Le mode conditionnel.
R, R	Les ensembles des nombres réelles.
$\mathsf{R}^+, \mathsf{R}^+$	Les ensembles des nombres réelles positive.
Z, Z , $\mathcal{Z}$	Les ensembles des entiers.
h	Le paramètre de lissage.
$B(\chi, h)$	La boule de centre χ et de rayon h de l'espace (E, d).
$\phi_\chi(h)$	Mesuré la boule $B(\chi, h)$ par rapport à la loi de probabilité de χ.
K	La fonction du noyau.

Variables aléatoires

$(X_i, Y_i)_{i=1}^n$	L'échantillon de n variables aléatoires.
$\mathsf{E}(Y/\chi = x)$	Moyenne conditionnellle de Y par rapport à X (réalisation de $\mathsf{E}(Y/X)$)
$Var(x)$	La Variance de X
$\sigma(x)$	L'écart-type (réalisation de $Var(Y/X)$)
$Cov(X, Y)$	Covariance de X et Y

Symboles

Soient $(A_n)_{n \quad 1}$ et $(B_n)_{n \quad 1}$ deux suites de variables aléatoires telles que pour tout $n \geq 1$, on a $B_n \neq 0$.

$A_n = O(B_n), n \to \infty$	$A_n/B_n \to C$ où C est une constante avec A_n *et* B_n sont deux suites aléatoires.
$A_n = o(B_n), n \to \infty$	$A_n/B_n \to 0$ avec A_n, B_n sont deux suites aléatoires.
$(a.co)$	La convergence presque complète
$0_{a.co}$	Le taux de la convergence presque complète.
$i.i.d$	Indépendants et identiquement distribuées
$c.d.f$	Fonction de répartition conditionnelle.
NW	Nadaraya Waston
$P.L$	Polynôme Locaux
$L.L$	Linéaire locale
$r.v$	Variable aléatoire

Chapitre 1

Introduction

Le problème qui suit, peut-être le plus usuellement étudié, est celui de la relation d'une variable par rapport à une où plusieurs autres. Cette dernière a été largement étudiée pour des variables réelles ou vectorielles, mais la relation se produit aussi évidemment avec des variables fonctionnelles. En outre, il généralise les résultats de dimension finie. La statistique fonctionnelle a un développement très important dans les dernières années, car une dimensionnalité élevée implique un mauvais résultat. L'étude des données fonctionnelles posent des questions cruciales en statistique, car il y a effectivement un grand nombre dans laquel les collectées sont des courbes. En effet, les progrés informatiques à la fois en termes de mémoires et de leurs capacités de calcul, nous permet de faire des grands ensembles. En particulier, pour un seul phénomène nous pouvons observer un très grand nombre de variables. Ces axes de recherche concernent les techniques statistiques classiques telles la modélisation paramétrique, semi-paramétrique et non-paramétrique.

1.1 Approche paramétrique

La plupart des méthodes statistiques élémentaires connues sont paramétriques. On dit que le problème est paramétrique si la famille de départ est paramétrée par un vecteur de dimension finie ; dans ce cas sa résolution se résume à l'identification de ce paramètre. L'estimation de ce dernier permet alors de retenir l'une des distributions de cette classe. Ainsi cette approche nécessite donc des hypothèses restrictives. D'une manière générale, si les hypothèses supplémentaires sont correctes, ces méthodes peuvent produire des estimations plus exactes et précises. Autrement dit, les formules paramétriques sont souvent plus simples à écrire et plus rapides à calculer. Par contre, si ces hypothèses sont incorrectes, les méthodes paramétriques peuvent-être très trompeuses, parfois cette approche peut être mal adaptée à la réalité des données. Pour cette raison, il ne sont pas souvent considérés comme robustes.

1.2 Approche non-paramétrique

L'approche non-paramétrique a été développée afin de pallier les problèmes de l'hypothèse de la modélisation paramétrique, il s'agit d'estimer un vecteur de dimension infinie. L'avantage de cette méthode est qu'elle ne nécessite pas d'hypothèses sur la nature de la distribution et la robustesse à la présence des points aberrants. Lorsque les données n'ont aucune interprétation numérique claire, comme les hypothèses non-paramétrique sont beaucoup plus large que les méthodes paramétriques correspondantes, de nombreux travaux ont été menés pour cette estimation aussi bien dans un cadre théorique que sur le plan des applications.

1.3 Approche semi-paramétrique

L'estimation semi-paramétrique est devenue un sujet très important au cours des dernières années. Ces modèles permettent d'obtenir un compromis entre un modèle paramétrique qui est généralement trop restrictif et un modèle non-paramétrique où la vitesse de convergence des estimateurs se dégrade vite en présence d'un grand nombre de variables explicatives. Il s'agit d'estimer un vecteur de dimension finie mais la famille de départ n'est pas paramétrée par un vecteur de dimension finie. Cette considération est due en premier lieu aux problèmes dûs à la mauvaise spécification de certains modèles. Aborder un problème de mauvaise spécification de manière semi-paramétrique consiste à ne pas spécifier la forme fonctionnelle de cer-

taines composantes du modèle. Cette approche complète celle des modèles non-paramétriques, qui ne peuvent pas être utiles dans des échantillons de petite taille ou avec un grand nombre de variables. Le deuxième inconvénient est lié au manque des moyens pour quantifier l'effet de chaque variable explicative. Cette approche alternative est naturellement fournie pour pallier ces inconvénients. Dans ce domaine, différents types de modèles ont été déjà étudiés dans la littérature : parmi les plus célèbres, on peut citer les modèles additifs, les modèles partiellement linéaires ou encore les modèles à direction révélatrice unique. Ces derniers modèles représentent une généralisation naturelle des modèles linéaires généralisés. D'un point de vue du lissage non-paramétrique multidimensionnel, les modèles à indices offrent des techniques pour réduire la dimension permettant d'éviter les vitesses de la convergence trop lentes des estimateurs non-paramétrique.
Dans ce cas, chacun de ces trois approches de prévision conduit à l'estimation de certains opérateurs, l'espérance conditionnelle (pour la régression fonctionnelle), la fonction de répartition conditionnelle (par la médiane fonctionnelle) et la densité conditionnelle (pour le mode fonctionnel). Nous faisons remarquer que ces paramètres conditionnels sont largement étudiés en dimension finie. Nous présentons ici deux méthodes d'estimation générale et particulière :

- La méthode du noyau
- La méthode des polynômes locaux

1.4 Estimation de la fonction de régression

La régression est l'un des outils les plus fréquemment utilisés pour les études statistiques. Dans sa forme la plus simple, ce dernier permet aux chercheurs d'analyser les relations d'une variable par rapport à une où plusieurs autres. Pendant longtemps, la régression d'une variable aléatoire y sur le vecteur de variables aléatoires X désignait la moyenne conditionnelle de y sachant x. Aujourd'hui, le terme de régression désigne tout élément de la distribution conditionnelle de y sachant x considérée comme une fonction de x. Les performances de cette méthode dans la pratique dépend de la forme des processus générateurs des données et la façon dont elle se rapporte à la méthode de régression utilisé. Dans la prévision, ces modèles sont souvent utiles même si elles ne peuvent pas fonctionner de manière optimale. Cependant, dans de nombreuses applications et en particulier avec des petits effets où des questions de causalité base sur des données d'observations, les méthodes de la régression peuvent donner des résultats trompeux.

1.4.1 Méthode du Noyau

L'estimation de la fonction de la régression par la méthode du noyau est une technique pour la prévision qui a été largement développée dans la statistique. Les premiers résultats ont été obtenues par Nadaraya et Waston ([108, 131], 1964), ils ont établi la convergence uniforme, Stone en ([129],1977) a étudié la convergence en norme L_p lorsque les observations sont indépendantes et identiquement distribuées. Plus de références sur ce sujet sont données par Colomb en ([25], 1981, [26], 1984, [27], 1985) qui a démontré la convergence-presque complète de l'estimation non-paramétrique, Györfi et al. ((([80], 1989) ont introduit l'estimation non-paramétrique pour des séries de temporelles, Roussas ([123], 1989) a étudié la normalité asymptotique pour un processus Markovien. Hardle ([83], 1990) a fait une application sur la régression non-paramétrique, Yoshihara ([135], 1994), Bosq ([16, 17], 1996, 1998) ont étudié cette méthode pour la prévision de la statistique non-paramétrique pour des processus stochastique. Nous citons par exemple Hardle et al. ([89], 1993), Hristache et al. ([92], 2001), ils ont estimé le coefficient à indice simple par une méthode directe. Ensuite, ils ont établi la convergence en probabilité avec prévision de ses estimateurs. Délécroix et al. ([36], 2003) ont étudié l'esimation à indice simple pour établir des propriétés asymptotiques. Dans le cas fonctionnel, les premiers résultats de l'estimation non-paramétriques ont été étudié par Ferraty et View ([61, 62], 2000, 2002), ils ont établi la convergence presque-complète dans le cas indépendant, ces derniers résultats ont été généralisée au cas α-mélange en ([60], 2004), la convergence en norme L_p dans le cas indépendant (voir Dabo-Niang et Rhomari ([31], 2002). Pour le modèle à indice fonctionnel, Ferraty et al. ([67], 2003) ont obtenu des propriétés asymptotiques lorsque l'indice est connu. Ils ont établi la convergence presque complète dans le cas indépendant. Leurs résultats ont été généralisés au cas dépendent par Ait Saidi et al. ([5],2005). Deheuvels and Mason ([38], 2004)) ont obtenu des intervalles de confiance pour l'estimateur à noyau de type Nadaraya-Waston. Masry en ([106], 2005) a montré la normalité asymptotique de l'estimateur de Ferray et al ([60], 2004), la convergence en moyenne quadratique a été étudiée par Ferraty et al ([68], 2007), ces derniers résultats ont été utilisés par Rachdi et Vieu ([111], 2007) pour déterminer le paramètre de lissage par la méthode de la validation croisée. Ferraty et view ([65], 2006) ont établit la convergence presque complète de la fonction de régression. Les résultats sur l'intégrabilité uniforme ont été établis par Delsol ((([44], 2007). Une étude générale est faite par Ait Saïdi et al. ([1], 2008) lorsque l'indice fonctionnel est inconnu, ils ont proposé un estimateur de ce paramètre, basé sur la technique de validation croisée. Dabo-Niang et Serge ([32], 2010) ont utilisé un modèle semi-paramétrique fonctionnel où la variable réponse est à valeurs réelles est expliquée par la somme d'une combinaison linéaire des composantes supposées inconnues. Bouraine et al. ([12], 2010) ont utilisé le modèle de Ait Saidi et al. ([1], 2008) pour estimer la fonction de la régression à indice multi-fonctionnel par la méthode de validation croisée. On se refère à Ferraty et Vieu ([70], 2011) ainsi qu'à Delsol ([47], 2011) qui ont étudié le modèle non-paramétrique de la variable aléatoire fonctionnelle dans le cas de mélange fort.

1.4.2 Méthode des Polynômes locaux

Au XXe siècle, la régression polynômiale a joué un rôle important dans le développement de l'analyse de régression avec un plus grand effet sur les questions de l'inférence. Cette méthode d'ordre p est plus générale par rapport à la méthode du noyau. Gergonne ([78], 1815) a été le premier qui a abordé la régression polynômiale par la méthode des moindres carrées. Pour cette raison, la régression polynômiale est considérée comme un cas particulier de la régression linéaire multiple. Cette méthode est naturellement plus complexe mais elle a des meilleures propriétés statistiques. L'ajustement de cette méthode et en particulier l'estimation linéaire locale ont des grandes propriétés attirrantes par à rapport à l'estimateur de Nadaraya Watson ([108, 131], 1964). L'idée maîtresse est de considérer le problème de la régression sous l'angle des moindres carrés. En particulier, l'estimation linéaire locale de la régression a été développée lorsque la variable explicative est univariée ou multivarié (voir Wand et Jones ([132], 1995) pour un aperçu sur ce sujet). Les propriétés statistiques de la régression polynômiale local pour des données dépendantes ont été étudies dans les travaux récents de Masry et Fan ([105], 1997), Masry ([103], [104], 1996a, b), Hardle et Tsybakov ([88], 1997), Hardle et al. ([85], 1998), Opsömer ([99], 1997), et Vilar-Fernández (([49], 1998), ([50], 2000)). Martins-Filho et Yao ([102], 2009) ont établit la normalité asymptotique de cet estimateur. Très récemment, les modèles de la régression linéaire local ont été étudié lorsque la variable explicative est fonctionnelle (voir Aneiros-Pérez, Cao, et Vilar-Fernàndez ([1], 2008) ; Boj, Delicado, et Fortina ([14],2008) ; Baíllo et Grané ([10], 2009)), nous nous référons à García-Soidán et al. ([74], 2006), Hertgartner et al ([82], 2002), Cheng et al. ([24], 2002). Barrientos-Margin ([47], 2010) a montré la convergence presque complète. Chouaf et al. ([29], 2012) ont adressé l'estimateur linéaire local fonctionnel de la régression non-paramétrique spatiale.

1.5 La fonction de répartition conditionnelle

La fonction de répartition conditionnelle est une technique statistique qui permet une meilleure compréhension de la relation entre une variable de réponse et un ensemble de variables, en comparant avec les méthodes de la régression usuelles . L'inconvénient de la fonction de régression est qu'elle ne fournit qu'une valeur prédictive sans nous renseigner sur la probabilité que la réalisation soit proche de celle-ci (c'est à dire sans décrire la distribution de la variable conditionnée) et elle est sensible aux valeurs abérrantes et peut se montrer inapropriée dans certains cas, comme lorsque la distribution est multimodale ou fortement asymétrique. En revanche, la fonction de répartition conditionnelle donne la probabilité que la variable d'intérêt appartienne à un intervalle donné.

1.5.1 Méthode du Noyau

Plusieurs auteurs ont introduit cette estimation comme une étude préliminaire du quantile conditionnel. Stone ([129], 1977) semble être le premier qui a étudié l'inverse de cette estimation pour obtenir le quantile conditionnel, il a obtenu la convergence en probabilité. L'estimation non-paramétrique du quantile conditionnel et ses dérivées partielles a été obtenue par Schlee ([127], 1982). Stute ([126], 1986) a établit la convergence presque sure de la fonction de répartition empirique. Samanta ([125],1989) a démontré la normalité asymptotique et la convergence uniforme dans le cas indépendant. Roussas ([113], 1991) a élaboré un résultat sur la convergence presque pour un processus de Markov. Tandis que Avec cette même idée, Ducharme, Gannoun, Guertin et Jéquier ([37], 1995) ont obtenu la normalité asymptotique de l'estimation du quantile conditionnel. Yu et Jones ([134],1998) ont proposé une méthode du double noyau pour étudier la convergence en probabilité et la convergence en moyenne quadratique. Honda ([91], 2000) a montré la normalité asymptotique et la convergence uniforme pour un processus α-mélange. Gannoun et al. ([77], 2003) pour la prévision non-paramétrique de la médiane conditionnelle et du quantile. Dans le cas fonctionnel, les premiers résultats ont été développés par Ferraty et View ([65], 2006), ils ont établi la convergence uniforme presque complète dans le cas indépendant et identiquement distribuée et offrent des nouvelles perspectives dans de nombreux domaines d'applications lorsque les variables explicatives sont évaluées dans un espace de dimension infinie. Le cas de α-mélange a été étudié par Ferraty et al. ([4], 2005). Citons par exemple, Ezzahriouri et Ould-Said ([50],2005-2006) qui ont étudié la normalité asymptotique de cet estimateur dans les deux cas indépendant et dépendant. Nous revenons à Cardot et al. ([23], 2004) pour une approche linéaire du quantile conditionnel. La convergence en norme L_p sous l'hypothèse de concentration de la mesure de probabilité du quantile conditionnel a été obtenu par Dabo-Niang et Laksaci Ali ([33],2012).

1.5.2 Méthode des Polynômes locaux

L'utilisation de la fonction de répartition conditionnelle a été étudiée par plusieurs auteurs. Il est bien connu que le biais asymptotique de l'estimateur de Nadaraya-Waston a une mauvaise forme et pour surmonter ce problème, on a utilisé cette méthode alternative qui dépend des techniques décrits par une fonction polynomiale local dans Fan et Gijbels ([58], 1996). Rappelons que dans la modélisation linéaire local de la fonction de répartition conditionnelle dans le cas multivarié, il faut utiliser un développement de Taylor. Cependant, dans le cadre éventuellement fini, cette technique n'est plus utilisable et ils l'ont développé par une nouvelle technique. Pour cela, il existe plusieurs façons d'étendre les idées linéaires locals à la dimension infinie. Min et Mouvid ([107], 2000) ont étudie l'estimateur linéaire local de la fonction de répartition conditionnelle mais cette méthode implique des calculs lourds. Cet estimateur a été moins étudié et a l'avantage d'offrir des meilleurs résultats en termes de l'érreur en moyenne quadratique et de corriger automatiquement les effets de bords. Ils ont étudié les propriétés asymptotiques comme la convergence uniforme de l'estimateur linéaire local. Ce problème a été récemment abordé par yu et jones ([134], 1998) via une application du double noyau de l'approche linéaire, l'approche linéaire local de Fan, Yao, et Tong ([56], 1996). Sandie Ferrigno, Gilles R. Ducharme ([124], 2007) ont proposé un critère du choix optimal pour utiliser la fenêtre d'ajustement pour l'estimation polynômiale local de la fonction de répartition conditionnelle. Le résultat qui a été obtenu par Ali Laksaci et al.([5], 2013) concernant la modélisation spatiale de l'estimateur linéaire local de l'estimation de la fonction de répartition conditionnelle dans le cas ou la variable explicative est fonctionnelle.

1.6 La densité conditionnelle

la littérature statistique regorge d'expression pour décrire la densité conditionnelle comme outil pour analyser la dépendance entre les variables aléatoires. Cette estimation a un grand intérêt au cours des vingts ans derniers. Autant que la symétrie des fonctions de la densité conditionnelles n'ont jamais été abordé dans la littérature, cette technique a une grande importance pour de nombreux scientifiques où les connaissances en moyennes conditionnelles, obtenues par des méthodes de régression, n'ont pas suffisantes pour tirer des conclusions valables sur le problème à la main. Il existe des variétés d'estimateurs de la densité conditionnelle, mais la plupart d'entre eux manquent de généralité où facilité d'interprétation. On sait que, sous certaines hypothèses de differentiabilité, la densité conditionnelle peut être obtenue en dérivant la fonction de répartition conditionnelle ;

$$\hat{f}_Y^x(\chi, y) = \frac{\partial}{\partial y} \hat{F}_Y^\chi(\chi, y)$$

Tel que

$$\hat{F}_Y^\chi(\chi, y) = \mathsf{P}(Y \leq y/\chi = x)$$

1.6.1 Méthode du Noyau

L'estimateur à noyau d'une densité est l'un des estimateurs les plus étudiés et les plus performants. Le premier résultat de la limite centrale de l'estimateur de processus de markov a été obtenu par Roussas ([114, 115], 1967, 1969), Rosenblatt ([116, 117],1970, 1971) a étudié les propriétés asymptotiques indépendantes. La convergence presque-sûre a été obtenue par Bosq ([14], 1971 ,[15],1973), chahboun ([18], 1984) et Délécroix ([22], 1975). Délécroix et al. ([36], 2003) ont utilisé le principe de maximum de vraisemblance sur la densité conditionnelle pour estimer l'indice du modèle. Hyndman et al. ([97], 1996) ont montré l'erreur de la moyenne quadratique et ils ont permet à réduire le biais de l'estimation pour certains modèles. Ferrety et al. ([66], 2006) et Ali laksaci ([6], 2007) ont étudié la moyenne quadratique de la densité conditionnelle pour des variables fonctionnelles. Hall et al., ([81],2004), Fan et Yim, ([73],2004) également ont étudié les propriétés et le choix de certains estimateurs. Nous renvoyons à Youndjé ([54], 1996) et ([55], 2002) pour une présentation assez générale sur l'estimation de la densité conditionnelle, laquelle donne des conditions nécessaires de la convergence ponctuelle ou uniforme en probabilité, des conditions suffisantes de la convergence en moyenne quadratique pour le noyau de convolution. Récemment, Attaoui et al. ([9], 2011) ont établié la convergence presque complète lorsque la variable de réponse Y est scalaire sur une variable aléatoire Hilbertien X. Ils ont supposé que la densité conditionnelle $\widehat{f^x}(t)$ existe sur un structure d'indice fonctionnel unique θ.

1.6.2 Méthode des Polynômes locaux

L'estimation de la densité conditionnelle par la méthode des polynômes locaux a été proposé par Fan, Yao et Tong ([56], 1996). Une correction de biais a été proposé par Hyndman, Bashtannyk et Grunwald ([97], 1996), voir également Fan et Yao ([64], 2003). Bashtannyk et Hyndman et al. ([98], 2001), Fan et Yim ([73], 2004) et Hall, Racine et Li ([81], 2004). Dans d'autres documents, ils ont utilisé la densité conditionnelle en tant que contribution à d'autres problèmes, y compris les Robinson ([112], 1991), Tjostheim ([130], 1994), Polonik et Yao ([110], 2000) et Hyndman and Yao ([96], 2002). Ali Laksaci et al. ([42], 2010) ont estimé la densité conditionnelle et sous certaines conditions, ils ont établit des propriétés asymptotiques de cet estimateur telles que la convergence presque complète et la convergence en moyenne quadratique dans le cas où les observations sont indépendantes et identiquement distribuées.

1.7 Prévision du mode conditionnel

L'estimation du mode conditionnel et inconditionnel de la densité a une longue histoire dans la littérature statistique. L'estimateur du mode conditionnel $\hat{\theta}$ est comme suit :

$$\hat{f}^x(\hat{\theta}) = \sup_{y 2 S} \hat{f}^x(y) \tag{1.1}$$

Où S est un ensemble compact choisi de telle sorte que le mode θ de f^x est supposé être défini de manière unique. L'estimation du mode conditionnel est très abondante lorsque la variable explicative est à valeurs dans un espace de dimension finie. Citons par exemple, Pearson ([109],1962) qui était le premier ayant considéré le problème de l'estimation du mode de probabilité univariée, Quintela del Rio et Vieu ([4], 1997), Ould-Saïd ([50], 1997), et Louani et Ould-Saïd ([39],1999). Concernant le cas fonctionnel, Gasser et al. ([75], 1998) ont proposé un estimateur du mode de la distribution des courbes aléatoires, Ezzahrioui et Ould-Said ([51], 2006a, 2006b) ont étudié la normalité asymptotique de l'estimateur du noyau du mode conditionnel dans les deux cas indépendant et de mélange fort. La monographie de Ferraty et Vieu

([68],2006b) présente une importante collection d'outils statistiques pour la prévision non-paramétrique fonctionnel. Récemment, Dabo-Niang et Laksaci ([43],2007) ont démontré la convergence en norme L_p dans le cas indépendant, cette vitesse de convergence est étroitement liée à la propriété de concentration de la mesure de probabilité de la variable explicative sur des petites boules (voir Dabo-Niang et Rhomari ([31],2002)). Ce résultat nous permet d'établir une solution originale au problème de la dimension et de généraliser les résultats asymptotiques existent dans le cas multivarié à la dimension infinie. Damongeot et al. ([41], 2011a), Damongeot et al ; ([42], 2011b) et Ferraty et al. ([69], 2010) ont proposé un estimateur du mode conditionnel basé sur la densité conditionnel fonctionnelle. Ezzahrioui et Ould-Said ([52],2005) ont étudié la normalité asymptotique pour le mode conditionnel dans les deux cas indépendant et dépendant. Ali laksaci et al.([5],2013) ont développé l'approche linéaire local pour estimer le mode conditionnel $\widehat{\theta}$ de la relation précédente (1.1).

1.8 L'analyse sur des espaces fonctionnels

Depuis plusieurs années, de nombreux staticiens ont développés des applications permettant le traitement des variables aléatoires fonctionnelles. L'analyse fonctionnelle est un domaine des mathématiques qui s'est développé dans la première moitié du 20 ème siècle grâce en particulier aux travaux de M.Fréchet, S. Banach, D. Hilbert. L'étude de la stabilité amène naturellement à considérer des espaces munis de la topologie définie par des normes, des semi-normes où des distances. Le passage de la dimension infinie n'est pas toujours facile car on perd une partie de l'intuition géométrique. Alors que sur un espace vectoriel de dimension finie il y a une seule topologie "raisonnable", sur un espace de dimension infinie on doit souvent considérer plusieurs topologies simultanément. L'analyse fonctionnelle est un point de vue (intuitivement) géométrique, (formellement) algébrique, sur la théorie des fonctions qui sont considérées comme des points d'un espace de dimension infinie. La statistique pour ces données ont étudiés les observations qui ne sont pas des variables réelles où vectorielles mais des courbes aléatoires et a connu un développement très important ces dernières années. Cette branche a été étudiée par des données qui, d'une part leur structure et le fait qu'elles soient collectées sur des grilles très fines, peuvent être assimilées à des courbes où à des surfaces. Compte tenu des capacités actuelles, des appareils de mesure et de stockage informatique. Cette spécialité a donné lieu à une littérature abondante depuis une trentaine d'années à cause du grand nombre des travaux, indique qu'elle est un des centres d'intérêt de la statistique actuelle. Cette thèse répond à une question importante qui est la dégradation de la dimension ainsi cette notion est étroitement liée à la rareté des données dans un espace de grande dimension. Nos travaux peuvent être considérés comme une extension de la combinaison entre la modélisation statistique semi-paramétrique et la variable aléatoire fonctionnelle par l'utilisation de la méthode linéaire locale. Il y a des questions intéressantes qui se pose :

- Que peut-on dire sur l'augmentation de la dimension lorsqu'on nous travaillons avec des données fonctionnelles ?
- Puisque la méthode des polynômes locaux est alternative par rapport à la méthode du noyau dans le cas finie ; comment cette idée est adaptés dans l'espace de dimension infinie ?

1.9 L'effet de la dimension infinie via l'espace semi-métrique

L'utilité des métriques et des distances est de pouvoir comparer les ressemblances et les différences entre deux vecteurs. Cette opération est importante, par exemple, dans le domaine de la classification. Il est plus probable que deux vecteurs semblables soient dans une même classe que deux vecteurs dissemblables. Du effet, l'utilisation des métriques pour mesurer des distances est presque toujours une étape essentielle à la classification automatique. Dans le cas d'une distance, on cherche habituellement les éléments les plus proches, c'est-à-dire qu'on cherche la distance minimale. Mais avant d'aller plus loin, il est nécessaire de préciser pourquoi nous préférons utiliser une semi-métrique plutôt qu'une métrique. En mathématiques, la notion d'espace semi-métrique est une généralisation de celle d'espace métrique dans laquelle on n'impose pas l'inégalité triangulaire. Dans les traductions de textes russes, le terme "semi-métrique" est parfois remplacé par "symétrique". Lorsque l'on se pose la question de savoir si deux objets sont plus où moins proche, la première idée consiste à essayer de construire une distance entre eux ? Dans le cas de dimension infinie, il est bien connu que toutes les métriques ne sont pas équivalentes. L'espace semi-métrique est liée à la grande dimension des données qui définie à partir des données fonctionnelles sur un espace de dimension plus petite. Cela permet de réduire la dimension des données et ainsi d'augmenter la vitesse de convergence tout préservant la nature des données fonctionnelle. Dans le cas de dimension infinie, l'augmentation de la vitesse est liée à la forte concentration de la mesure de probabilité de la variable explicative fonctionnelle.

1.10 Contribution de la thèse

Ce travail est consacré à des progrès récents concernant à la fois : la modélisation statistique semi-paramétrique et une variable fonctionnelle. Rappelons que ce type est une solution originale pour diminuer la dimension par rapport à l'augmentation de la vitesse de convergence. Dans cette section, nous donnons un résumé structurel de cette thèse. Une fois de plus à penser comment étendre ces idées de dimension infinie, la méthode des polynômes locaux devrait recevoir une attention particulière, tous ces résultats sont des extensions au cadre fonctionnel. Il s'agit tout d'abord pour le rôle de la dimension dont le bon choix va permettre aux estimateurs des polynômes locaux d'atteindre des vitesses de convergence optimales puis du problème de la dimension dont on verra qu'elle dégrade sensiblement. Il est important de souligner le fait que la méthode du noyau est en réalité un cas particulier de la méthode des polynôme locaux. En effet, certains problèmes ; qui comporte la méthode du noyau peuvent être diminués car il y a des différences dans le biais aux bornes des données de ces méthodes. Il y a plusieurs façons d'accorder la création de la prédiction et l'un des plus populaires est certainement la méthode de la régression qui est basée sur l'espérance conditionnelle pour des raisons de robustesse (en fonction du comportement de la fonction de la distribution conditionnelle) et deux alternatives sont la médiane conditionnelle et le mode conditionnel. Ces trois prévisions ont été largement étudiées dans des situations où la variable explicative est évalué dans un espace de dimension infinie.

1.11 Plan de la thèse

Cette thèse est organisé comme suit : Le premier est un chapitre introductif, qui présente une étude bibliographique des problème liés à l'analyse statistique des variables fonctionnelles ainsi qu'a l'estimation des paramétres conditionnels que ce soit dans le cadre de dimension finie ou infinie. Ensuite, nous exposons les résultats obtenus dans la règression, la densité, la fonction de répartition ainsi que le mode conditionnel tout en fournissant et discutant les hypothèses qui ont permis d'obtenir ces résultats à l'estimation par la méthode du noyau et polynome locaux. On présente notre modèle linéaire local de la fonction de répartition conditionnelle et les résultats obtenus dans le cas indépendant et dépendant dans la dernière partie.

Le deuxième chapitre est consacré à l'étude d'une variable aléatoire réelle Y sachant une variable aléatoire fonctionnelle X à valeur dans un espace semi-métrique. Cette relation est basée sur une estimation par linéaire locale. On constate que la dimension à un effet inverse sur la vitesse de convergence et il y a une dégradation de la vitesse de convergence par rapport à l'augmentation de la dimension. Nous établissons sous certains conditions, la convergence presque complète ainsi la vitesse de convergence de la fonction de répartition conditionnelle dans le cas où les observations sont indépendants et identiquements distribuées.

Ensuite, Dans le même contexte fonctionnel, nous considérons le cas dépendant, c'est à dire, le cas où les observations sont fortement mélangeant. Nous établissons la vitesse de convergence presque complète de l'estimateur construit.

Tandis que dans le chapitre 4, nous exposons quelques conclusions et des perspectives de recherches permettant de généraliser les résultats de cette thèse.

Le dernier chapitre est consacré á l'introduction des définitions et outils techniques utilisés pour l'élaboration de nos résultas. En particulier, nous rappelons la définition de processus de mélange, les définitions de différents modes de convergence,$\cdots$ On trouvera, aussi dans ce chapitre un nombre très important d'outils nécessaires pour l'élaboration des nos résultats.

1.12 Présentation du modèle linéaire locale de la fonction de répartition conditionnelle

Soit $(X_1, Y_1), \cdots, (X_n, Y_n)$ un n-échantillons aléatoires des données a deux variables que nous disposons pour estimer la fonction de répartition conditionnelle $F^x(y) = \mathbb{E}(\mathbf{1}_{Y \quad y}/X = x)$. Le but principal se trouve à travers la liaison entre la variable réponse réelle Y et la variable explicative fonctionnelle X dans un espace de dimension infinie $\mathcal{F}$. Pour pallier ces problèmes, plusieurs méthodes telles que la méthode des polynômes locaux, noyau, ... ont été supposées dans la Littérature. L'ajustement des polynômes locaux d'ordre p est noté par : $(L_p)(p)$. Le choix de cette méthode à cause de leur flexibilité et de leur performance asymptotique, l'approche la plus populaire est de combiner le concept ou $p \in \{0, 1\}$, on trouve l'estimateur de Nadaraya-Waston lorsque $p = 0$. Plus précisément, comme dans le cas réel, l'estimateur peut-être présenté comme la solution du problème de la minimisation suivante :

$$\min_{F} \sum_{i=1}^{n} (\mathbf{1}_{Y_i \quad y} - F^x(y))^2 K(h \quad ^1(\delta(x, \chi_i)))$$

K est un noyau, $\delta(.,.)$ localise un élément de $\mathcal{F}$ par rapport à l'autre et h est un paramètre de lissage (suites de nombres réelles positifs). Dans ce cas, $F^x(.)$ est une constante fonctionnelle local car elle se rapproche de l'indicatrice $\mathbf{1}_{fY \quad yg}$ par rapport à une constante. Par contre, si $p = 1$ on trouve l'estimateur linéaire locale qui est obtenu lorsque (a, b) minimise le problème suivant :

$$\min_{(a,b)2\mathsf{E}^2} \sum_{i=1}^{n} (\mathbf{1}_{Y_i \quad y} - (a + b\beta(x, \chi_i)^2 K(h \quad ^1(\delta(x, \chi_i))) \tag{1.2}$$

$\beta(.,.)$ est une fonction connue définit dans $\mathcal{F}$ vers R telle que $\forall \xi \in \mathcal{H}, \beta(\xi, \xi) = 0$ et peut prendre aussi des valeurs négatives. Une fois que la valeur $\hat{a}$ de a

minimise l'équation précédente, on prend $\hat{F}^x_{LL}(y) = \hat{a}$ comme l'estimateur linéaire locale de $F^x(y)$. Par suite,

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left[\mathbb{1}_{\{Y_i \leq y\}} - a - b\beta(x,\chi_i)\right]^T K_i \left[\mathbb{1}_{\{Y_i \leq y\}} - a - b\beta(x,\chi_i,x)\right]$$

Ou T désigne la transposition pour un vecteur ou une matrice. Il y a des différents choix de l'idée linéaire locale lorsque la variable est évaluée dans un espace de dimension infinie. L'estimateur $\hat{F}^x$ de F^x est la solution de $\hat{a}$ qui est comme suit :

$$\hat{F}(y/x) = e_1^T (Q_\beta^T K Q_\beta)^{-1} (Q_\beta^T K \mathbb{1}_{\{Y \leq y\}}) \tag{1.3}$$

Tels que les notations matricielles associées sont par définition

$$Q_\beta^T = \begin{pmatrix} 1 & \cdots & 1 \\ \beta_1 & \cdots & \beta_n \end{pmatrix}, \qquad \text{et} \qquad \mathbb{1}_{\{Y \leq y\}} = \begin{pmatrix} \mathbb{1}_{\{Y_1 \leq y\}} \\ \vdots \\ \mathbb{1}_{\{Y_n \leq y\}} \end{pmatrix}$$

K désigne la matrice diagonale $n \times n$ de poids

$$K = \operatorname{diag}\{K(\delta(x,\chi_1)), \cdots, K(\delta(x,\chi_n))\}$$

Plus précisement, l'estimateur linéaire locale de la fonction de répartition conditionnelle est comme suit :

$$\hat{F}^x(y) = \hat{a} + \hat{b}\beta(x, X_i)$$

Tel que $\hat{a}$ et $\hat{b}$ sont des solutions de (1.2). L'avantage principal est que l'on obtient l'expression suivante :

$$\hat{F}^x(y) = \frac{\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij} \mathbb{1}_{\{Y_j \leq y\}}}{\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}}$$

Tel que ; $W_{ij} = \beta_i(\beta_i - \beta_j) K_i K_j$ ou $K_i = K(h^{-1}(\delta(x,\chi_i)))$ et $\beta_i = \beta(x,\chi_i)$.

1.13 Résultats de la thèse

1.13.1 La convergence presque complète

Cas i.i.d

Dans cette section, nous considérons que les observations sont indépendantes et identiquement distribuées. Sous des hypothèses de régularité, on a établit la convergence presque complète de l'estimateur construit de la fonction de répartition conditionnelle par la méthode linéaire locale dans un espace de dimension infinie par des données fonctionnelles.

Théorème 1.13.1. *([7], 2012) Pour tout $x \in \mathcal{F}$, et sous les conditions du théorème (2.1), on a*

$$\left|\hat{F}^x(y) - F^x(y)\right| = \circ(h_K^b) + \circ\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad p.co\,[1]$$

Cas de α mélange

On généralise ces derniers résultats au cas dépendant. Nous gardons les mêmes hypothèses ainsi les même notations du cas indépendant et on ajoute les conditions suivantes :

– La suite $(X_i, Y_j)_{i=1,\ldots,n}$ est $\alpha-$mélangeante, c'est à dire il existe deux constantes $c \in \mathbb{R}^+$ et $a \in \mathbb{R}^+$ telles que les coefficients de mélange vérifiant :

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

Et sur la concentration mixtes des variables fonctionnelles
Pour tout i, j

$$\sup_{i\neq j} \mathbb{P}((X_i, X_j) \in B(x,h) \times B(x,h)) \leq C\phi_x(h)^{a/(a-1)} \quad \text{avec} \quad a > 2.$$

– Sur le paramètre de lissage h_K :

il existe une constante $\eta > 0$ tel que : $\begin{cases} \lim\limits_{n\to+\infty} h_K = 0 & \lim\limits_{n\to+\infty} \dfrac{\log n}{n\phi_x(x)} = 0 \\ \text{et} & \exists\eta > 0, \quad \phi_x(h) \geq n^{\frac{3-a}{a+1}+\eta} \end{cases}$

1. la convergence presque complète.

Théorème 1.13.2. *Sous les conditions du théorème (2.1,[7], 2012) ainsi sur le coefficient de mélange $\alpha(n)$ et sur le paramètre de lissage (h) :*

$$|\hat{F}^x(y) - F^x(y)| = \circ(h^b) + \circ\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad p.co.$$

Bibliographie

[1] Ait Saïdi, A., Ferraty, F., Kassa, R. et Vieu, P. (2008) Cross-validated estimations in the single functional index model. (soumis)

[2] Aneiros-Pérez,g., Cao, R., Vilar-Fernández, J.M. (2008), Functional Methods for Time Series Prediction : A Nonparametric ApproachS. Proceedings of IAS2008, Pacifico, Yokohama, Japan, pp. 91-100.

[3] A. Quintela del Rio, P. Vieu (1997), A nonparametric conditional mode estimate, Nonparametric Statistic **8**, 253-266.

[4] Ait Saïdi, A., Ferraty, F., Kassa, R. (2005). Single functional index model for a time series. *Rev. Roumaine Math. Pures Appl.*, **50**, (4) 321-330.

[5] Laksaci. A, Rachdi. M, Rahmani. S, (2013).s Spaciale modelization : Local linear estimation of the conditional distribution for functional data.

[6] Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. Ann. I.S.U.P. **51**, 69-80

[7] Laksaci. A. Hachemi, N. (2012) : Note on the functional linear estimate of conditional cumulative distribution function. Journal of Probability ans Statistical Science ; **10(2)**, 153-160.

[8] Aneiros-Pèrez Vieu (2006), ph. Semi-functional partial linear regression. Statistics et Probability Letters **76**, 1102-1110.

[9] Attaoui, S., Laksaci, A. & Ould-Saïd, E. (2011). A note on the conditional density estimate in the single functional index model. Statist. Probab. Lett. **81**, 45-53.

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

[11] 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.

[12] Bouraine, M., Saidi, Ahmed A., Ferraty, F., Vieu, P., 2010. Cross-validation estimation in the multi-functional index model. Romanian Journal of Pure and Applied Mathematics 55, 355-367.

[13] Bosq, D., Lecoutre, J.P. Théorie de l'estimation fonctionnelle (in french). Economica (1987).

[14] Bosq, D. (1971). Contribution à la Théorie de l'Estimation Fonctionnelle. PhD thesis, Paris VI.

[15] Bosq, D. (1973). Sur l'estimation de la densité d'un processus stationnaire et mélangeant. C. R. Acad. Sci. Paris Sér. A-B, 277 :A535-A538.

[16] Bosq, D. (1996), Nonparametric Statistics for Stochastic Processes. Estimation and Prediction, Lecture Notes in Statistics 110, New-York : Springer-Verlag.

[17] Bosq, D. *Nonparametric Statistics for Stochastic Process. Estimation and Prediction* 2^{nd} *edition,* Lecture Notes in Statistics, 110, Spring-Verlag New York (1998).

[18] Bosq, D. (1997), Parametric rates of nonparametric estimators and predictors for continuous time processes, *Ann. of Stat., à paraître*

[19] Bosq, D. (2000). *Linear Processes in Function Spaces*. Lecture Notes in Statistics 149, Springer-Verlag, Berlin.

[20] Boj, E., Delicado, P., and Fortiana, J. (2008), 'Local Linear Functional Based on Weighted Distance-Based Regression', in *Functional and Operatorial Statistics*, eds. S. Dabo-Niang and F. Ferraty , Heidelberg : Physica-Verlag, pp. 57-64.

[21] Battacharya, P., Gangopadhyay, A., (1990), Kernel and nearestneighbor estimation of a conditional quantile. Ann. Statist., 18, 1400-1415.

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

[23] Cardot, H., Crambes, C. and Sarda, P. (2004) Spline estimation of conditional quantities for functional covariates. C. R. *Acad. Sci., Paris.* 339, (2) 141-144.

[24] Cheng,M-Y and Hall, P. (2002). Error Dependent Smoothing Rules in Local Linear Regression. *Statistica Sinica,* 12, 429-447.

[25] Collomb, G. (1981) Estimation non-paramétrique de la régrèssion : revue bibliographique. *International Statistacal Review,* 49, 75-93.

[26] Collomb, G. (1984) Propriétés de convergence presque complète du prédicteur à noyau. (French) [Almost complete convergence properties of kernel predictors] *Z. Wahrsch. Verw. Gebiete* 66 (3) 441-460.

[27] Collomb, G. (1985) Nonparametric regression : an up-to-date bibiliography. *Statistics* 16 (2) 309-324.

[28] Chahboun, M. (1984). *Estimation des densités de la loi et les probabilités de transition par les méthodes du noyau, des δ-suites et des fonctions orthogonales pour un processus de Markov stationnaire et mélangeant.* PhD thesis, Université de Rouen.

[29] Chouaf, A., Laksaci, A., 2012. On the functional local linear estimate for spatial regression. Stat. Risk. Model. 29 (3), 189-214.

[30] Caroll, R. J., Fan, J., Gijbels, I., Wand, M. P. (1997). Generalized partially linear singleindex models. J. *Amer. Statist. Assoc.*, 92, 477-489.

[31] Dabo-Niang S and Rhomari N (2002). *Nonparametric regression estimation when the regressor takes its values in a metric space.* Technical Report 9, Pre´publications L.S.T.A.

[32] Dabo-Niang S, and Serge Guillas.Functional semiparametric partially linear model with autoregressive errors.*J. Multivariate Analysis 101(2) :307-315 (2010)*

[33] Dabo-Niang Sophie, LAKSACI Ali ; (2012). *Nonparametric Quantile Regression Estimation for Functional Dependent Data.* vol. 41, n°7-9, pp. 1254-1268.

[34] Delecroix, M. (1975). *Sur l'estimation des densités marginales et de transition d'un processus stationnaire et mélangeant.* PhD thesis, Université de Lille.

[35] Delecroix, M., Hristache, M. (1999). M-estimateurs semi-paramétriques dans les modèles à direction révélatrice unique. *Bull. Belg. Math. Soc. Simon Stevin*, 6, 161-185.

[36] Delecroix, M., Härdle, W. and Hristache, M. (2003). Efficient estimation in conditional single-index regression. J. Multivariate Anal. 86 213-226.

[37] Ducharme, G., Gannoun, A., Guertin, M., Jéquier, J., (1995), Reference values obtaines by kernel based estimation of quantile regressions. 51, 11056-1116.

[38] Deheuvels P., Mason D.M., (2004). General Asymptotic bands based on Kernel-type function estimators, Statist. Inf. for Stoch. Proc. 7, 225-277.

[39] D. Louani, E. Ould Saïd, Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis, J. Nonparametric Stat. 11 (4) (1999) 413-442.

[40] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M., 2010. Estimation locale linéaire de la densité conditionnelle pour des données fonctionnelles. C. R. Acad. Sci., Paris I 348 (15-16), 931-934.

[41] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M., 2011a. Functional data : local linear estimation of the conditional density and its application. Statistics 47 (1), 26-44.

[42] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M., 2011b. A fast functional locally modeled conditional density and mode for functional time-series. In : Recent Advances in Functional Data Analysis and Related Topics. In : Contributions to Statistics, Physica-Verlag, Springer , pp. 85-90.

[43] Dabo-Niang, S., Laksaci, A., 2007. Nonparametric estimation of the conditional mode when the regressor is functional (Estimation non-paramétrique du mode conditionnel pour variable explicative fonctionnelle) (French. English summary). C. R., Math., Acad. Scc. Paris 344 (1), 49-52.

[44] Delsol, L. (2007) Régression sur variables fonctionnelle : Estimation, Tests de structure et Application. Thèse de Doctorat, Université de Toulouse.

[45]] Delsol, L. (2007a) CLT and L^p errors in nonparametric functional regression C. R. Math. Acad. Sci. 345 (7) 411-414.

[46] Delsol, L. (2007b) Régrèssion non-paramétrique fonctionnelle : Expressions asymptotiques des moments Annales de l'I.S.U.P. LI (3) 43-67.

[47] Delsol, L. (2011). Nonparametric methods for á-mixing functional random variables. In The Oxford Handbook of Functional Data Analysis (Ed. F. Ferraty and Y. Romain). Oxford University Press

[48] Dony, J., Einmahl, U. and Mason, D., 2006. Uniform in bandwidth consistency of local polynomial regression function estimators. Austrian Journal of Statistics, 35, 105-120.

[49] Efromovich, S. (2007). Optimal nonparametric estimation of the density of regression errors with finite support. *Ann. Inst. Statist. Math.*, 59(4) :617-654.

[50] E. Ould-Saïd, A note on ergodic processes prediction via estimation of the conditional mode function, Scand. J. Statist. 24 (2) (1997) 231-239.

[51] Ezzahrioui, M., Ould-Saïd, E., 2006. On the asymptotic properties of a nonparametric estimator of the conditional mode for functional dependent data. Technical report of LMPA, No. 277. Univ. du Littoral Côté d'Opale.

[52] Ezzahrioui, M., Ould-Saïd, E. Asymptotic normality of nonparametric estimators of the conditional mode function for functional data. Preprint (2005).

[53] Ezzahrioui, M., Ould-Saïdi, E. Asymptotic normality of the kernel estimators of the conditional quantile in the normed space. Preprint (2005).

[54] Élie Youndjé. Propriétés de convergence de l'estimateur à noyau de la densité conditionnelle. Rev. Roumaine Math. Pures Appl., 41(7-8) :535-566, 1996.

[55] Élie Youndjé and Martin T. Wells. Least squares cross-validation for the kernel deconvolution density estimator. C. R. Math. Acad. Sci. Paris, 334(6) :509-513, 2002.

[56] Fan, Jianqing, Qiwei Yao, and Howell Tong (1996) : "Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems," *Biometrika*, 83, 189-206.

[57] Fan, Jianqing and Qiwei Yao (1998) : "Efficient estimation of conditional variance functions in stochastics regression, *Biometrice*, 85, 645-660.

[58] J. Fan, I. Gijbels, Local Polynomial Modelling and its Applications, Chapman & Hall, London, 1996.

[59] Ferraty, F., Vieu, Ph., (2004), Nonparametric models for functional data , with application in regression, time-series prediction and curve discrimination. J. Nonparametric Stat. 16, No.1-2, 111-125.

[60] Ferraty, F., Laksaci, A. and Vieu, P., (2004). Estimating some chacteristics of the conditional distribution in nonparametric functional models. Statist. Inf. for Stoch. Proc., 9 47-76.

[61] Ferraty, F., Vieu, Ph., Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. C. R. Acad. Sci., Paris, Ser. I, Math. 330, No.2, (2000), 139-142.

[62] Ferraty, F., Vieu, Ph., (2002), The functional nonparametric model and applications to spectrometric data. Comput. Stat. 17, No.4, 545-564.

[63] Fan, J., 1992. Desigh-adaptive nonparametric regression. Journal of the American Statistical Association, 87, 998-1004.

[64] Fan, Jianqing and Qiwei Yao (2003) : Nonlinear Time Series : Nonparametric and Parametric Methods. New York : Springer-Verlag

[65] Ferraty, F., Vieu, P., 2006. Nonparametric functional data analysis. In : Theory and Practice. In : Springer Series in Statistics, New York.

[66] Ferraty, F., Laksaci, A., Vieu, P., 2006. Estimation of some characteristics of the conditional distribution in nonparametric functional models. Stat. Inference Stoch. Process. 9, 47-76.

[67] Ferraty, F., Peuch, A., Vieu, P. (2003). Modèle à indice fonctionnel simple. C.R. Acad. Sci. Paris, Ser. I, 336, 1025-1028.

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

[69] 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.

[70] F. Ferraty, A. Laksaci, A. Tadj, P. Vieu (2011). Kernel regression with functional response. *Electronic J. Statistics*, 5, 159-171.

[71] F. Ferraty, A. Quintela-del-Rio, P. Vieu (2011). Analysis of time of occurence of earthquakes : a functional data approach. *Mathematical Geosciences*, 43, 695-719.

[72] Ferraty, F., Vieu, P. (2011). Kernel regression estimation for functional data. In the Oxford Handbook of Functional Data Analysis (Ed. F. Ferraty and Y. Romain). Oxford University Press

[73] Fan, J. and Yim, T. H. (2004). A crossvalidation method for estimating conditional densities. *Biometrika*, 91(4) :819-834.

[74] García-Soidaán, P. (2006) Asymptotic normality of the Nadaraya-Waston semivariogram estimators. *To appear in TEST.*

[75] Gasser, T., Hall, P., Presnell, B., 1998. Nonparametric estimation of the mode of a distribution of random curves. J. R. Stat. Soc., Ser. B, Stat. Methodol. 60 (4), 681-691.

[76] Gannoun, A., (1990). Estimation non-paramétrique de la médiane conditionnelle. C.R. Acad. Sci., Paris, Sér. 5, 295-298.

[77] Gannoun, A., Saracco, J., Yu, K., (2003). Nonparametric prediction by conditional median and quantiles. J. Stat. Plann. Inference, 117, No.2, 207-223.

[78] Gergonne, J. D. (November 1974) [1815]. "The application of the method of least squares to the interpolation of sequences". Historia Mathematica (Translated by Ralph St. John and S. M. Stigler from the 1815 French ed.) 1 (4) : 439-447. doi :10.1016/0315-0860(74)90034-2

[79] Györfi, L. Recent results on nonparametric regression estimate and multiple classification. *Problems Control Inform. Theory, 10, 43-52(1981)*

[80] Györfi, L., Härdle, W., Sarda, P., Vieu, P. *Nonparametric curve estimation from time series.* Lecture Notes in Statistics, 60. Spring-Verlag, Berlin (1989).

[81] Hall et al., 2004 Hall, P., Racine, J., and Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *J. Amer. Statist. Assoc.*, 99(468) :1015-1026.

[82] Hertgarner, N., Wegkamp, M., and Martzner-Lober, E (2002) Bandwidth Selection for Local Linear Regression Smoothing. *Journal of Royal Statitical Society. Serie B Stat. Methodologgy,* 64, 791-804.

[83] Härdle, W. Applied nonparametric regression. Cambridge Univ. Press, UK (1990).

[84] Härdle, W., Hall, P., Ichumira, H. (1993). Optimal smoothing in single-index models, *Ann. Statist.*, 21, 157-178.

[85] Hardle, W., Tsybakov, A., Yang L. (1998). "Nonparametric vector autoregression," *J. of Statistical Planning and Inference*, 68, 221-245.

[86] Horowitz, J., Härdle, W. (1996). Direct semiparametric estimation of single-index models. *J. Amer. Statist. Assoc.*, 91, 1632-1640.

[87] Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct estimation of the index coefficient in the single-index model. *Ann. Statist.*, 29, 595-623.

[88] Härdle, W. and Tsybakov, A. (1997). Local polynomial est

[89] Härdle, W., Hall, P., Ichumira, H. (1993). Optimal smoothing in single-index models, Ann. Statist., 21, 157-178.

[90] Hedli-Griche, S. (2008) Estimation de l'opérateur de régrèssion pour des données fonctionnelles et des erreurs corrélées. *PhD Thesis.*

[91] Honda, T., (2000), Nonparametric estimation of a conditional quantile for $\alpha-$mixing processes. Ann. Inst. Stat. Math. 52, No. 3, 459-470.

[92] Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct estimation of the index coefficient in the single-index model. *Ann. Statist.*, 29, 595-623.

[93] Horowitz, J., Härdle, W. (1996). Direct semiparametric estimation of single-index models. J. Amer. Statist. Assoc., 91, 1632-1640.

[94] Hristache. M, A. Juditsky, V. Spokoiny, Direct estimation of the index coefficient in the single-index model, Ann. Statist. 29 (2001) 595-623.

[95] Hyndman, Rob J., D.M. Bashtannyk and G.K. Grunwald (1996) : "Estimating and visualizing conditional densities," Journal of Computational and Graphical Statistics, 5, 315-336.

[96] Hyndman, Rob J. and Qiwei Yao (2002) : "Nonparametric estimation and symmetry tests for conditional density functions," Nonparametric Statistics, 14, 259-278.

[97] Hyndman, R. J., Bashtannyk, D. M., and Grunwald, G. K. (1996). Estimating and visualizing conditional densities. J. *Comput. Graph. Statist.*, 5(4) :315-336.

[98] Bashtannyk, D.M. and Rob J. Hyndman (2001) : "Bandwidth selection for kernel conditional density estimation, Computational Statistics and Data Analysis, 36, 279-298.

[99] J.-D. Opsomer, On the existence and asymptotic properties of backfitting estimators. Preprint 96-12, Department of Statistics, Iowa State University. , 1997.

[100] Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. Ann. I.S.U.P., 51(3) :69-80 (2008).

[101] Ling, M., Norris, J.M., Kelman, M., Ward, M.P., 2012. Risk factors for death from canine parvoviral-related disease in Australia. Vet. Microbiol. 158, 280-290.

[102] Martins-Filho, C., Yao, F., 2009. Nonparametric regression estimation with general parametric error covariance. Journal of Multivariate Analysis 100, 309-333.

[103] Masry, E. (1996a). "Multivariate regression estimation : local polynomial fitting for time series," Stoch. Proc. Appl., 65, 81-101.

[104] Masry, E. (1996b). "Multivariate local polynomial regression for time series : uniform strong consistency and rates," J. Time Ser. Anal., 17, 571-599.

[105] Masry, E., Fan, J., 1997. Local polynomial estimation of regression functions for mixing processes. Scandinavian Journal of Statistics 24, 1965-1979.

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

[107] M. Mint EL Mouvid (2000). Sur l'estimateur linéaire local de la fonction de répartition conditionnelle. Thèse Université Montpellier 2.

[108] Nadaraya, E.A. On estimating regression Teor. Veroyatn. Primen. Teor. Veroyatn. Primen. 9, (1964), 157-159.

[109] Parzen, E. (1962). On estimation of a probability density function and mode. Annals of Mathematical Statistics 33, 1065-1076.

[110] Polonik, W. and Qiwei Yao (2000) : "Conditional minimum volume preditive regions for stochastic processes," Journal of the American Statistical Association, 95, 509-519.

[111] Rachdi, M. et Vieu, P. (2007) Nonparametric regression for functional data : automatic smoothing parameter selection. *J. Statist. Plann. Inference* 137 (9) 2784-2801.

[112] Robinson, Peter M. (1991) : "Consistent nonparametric entropy-based testing," Review of Economic Studies, 58, 437-453.

[113] Roussas, G. G. (1991). *Kernal estimates under association : strong uniform consistency.* Statist. Probab. Lett. 12, 215-224.

[114] Roussas, G. (1967). Nonparametric estimation in Markov processes. *Ann. Inst. Statist. Math.* 21, 73-87.

[115] Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. *Ann. Math. Statist.* 40, 1386-1400.

[116] Rosenblatt, M. (1970). Densiy estimates and Markov processes. In *Nonparametric Techniques in Statistical Inference* (M.L. Puri, ed.) Cambridge University Press, Cambridge, 199-210.

[117] Rosenblatt, M. (1971). Curce estimates. *Ann. Math. Statist.* 42, 1815-1842.

[118] Ramsay, J., Silverman, B.W. Functional Data Analysis. Springer-Verlag, New York (1997).

[119] Ramsay, J., Silverman, B.W. Applied functional data analysis ; Methods and case studies. Springer-Verlag, New York (2002).

[120] Ramsay, J., Silverman, B.W. Functional Data Analysis. 2nd Edition. Springer-Verlag, New York (2005).

[121] Ramsay, J. O., Silverman, B. W. (2005). Applied functional data analysis. Methods and case studies. Second edition, Springer-Verlag.

[122] Roussas, G. G., (1968). On some properties of nonparametric estimates of probability density functions. Bull. Soc. Math. Grèce (N.S.), 9, 29-43.

[123] Roussas, George G. Consistent regression estimation with fixed design points under dependence conditions. Stat. Probab. Lett. 8, No.1, (1989), 41-50.

[124] Sandie Ferrigno, Gilles R. Ducharme. (2007), Un choix de fenêtre optimal en estimation polynomiale locale de la fonction de répartition conditionnelle C. R. Acad. Sci. Paris, Ser. I 346 (2008) 83-86.

[125] Samanta, M., (1989), Nonparametric estimation of conditional quantiles. Stat. Probab . Lett. 7, No.5, 407-412.

[126] Stute, W., (1986). On Almost Sure Convergence of Conditional Empirical Distribution Functions. The Annals of Probability, 14, 891-901.

[127] Schlee, W., (1982), Estimation non-paramétrique du alpha-quantile conditionnel et ses dérivées partielles. 7, 32-47.

[128] Stute, W., (1986). On Almost Sure Convergence of Conditional Empirical Distribution Functions. The Annals of Probability, 14, 891-901.

[129] Stone, C.J. (1977). Consistent nonparametric regression, The Annals of Statistics, 5, 595-645.

[130] Tjostheim, D. (1994) : "Non-linear time series : A selective review," Scandinavian Journal of Statistics, 21, 97-130.

[131] Waston, G.S. (1964). Smooth regression analysis. *Sankhyà Ser.* A 26 359-372.

[132] Wand, M. P., and Jones, M. C. (1995). Kernel Smoothing. London : Chapman and Hall.

[133] Xia, Y., Tong, H., Li, W. K. (1999). On extended partially linear single-index models. *Biometrika*, 86, 363-410.

[134] Yu, K. Jones, M.C. 1998. Local linear quantile regression. Journal of the American Statistical Association, 93, 228-237.

[135] Yoshihara, K. Weakly dependent stochastic sequences and their applications. IV : Curve estimation based on weakly dependent data. Sanseido (1994).

Chapitre 2

Note on the functional linear estimate of conditional cumulative distribution function

This article is being published in the journal probability and statistical science to the local linear estimate of conditional distribution function.

Note on the functional linear estimate of conditional cumulative distribution function

Ali Laksaci
Laboratoire de Mathématiques,
Université de Sidi Bel Abbès,
BP 89 Sidi Bel Abbès 22000. Algérie.
alilak@yahoo.fr

Nawal Hachemi
Département de Mathématiques,
Centre universitaire Moulay Tahar Saida. Algérie 20000.
hacnawel@yahoo.fr

Abstract In this paper we introduce a new nonparametric estimation of the conditional cumulative distribution of a scalar response variable Y given a functional random variable X. Our estimate is based on the local linear approach. We prove under general conditions, the almost complete convergence (with rate) of the construct estimate.

Mathematics Subject Classification : 62G05, 62G99, 62G20.
Keywords : Local linear estimation, Conditional cumulative distribution, Functional random variables, Semi-metrics space, small balls probability.

2.1 Introduction

Recently, the modelization statistic of functional data have received a growing attention. This great consideration is motivated by the progress of the informatics tools permits the recuperation of large data sets, available essentially by real time monitoring, and computers can manage such databases. Functional data statistic (see Bosq ([3], 2000), Ramsay and Silverman, ([12], 2002), for the parametric model, Ferraty and Vieu, ([10], 2006), for the nonparametric case) can help to analyze such data sets. In this work, we consider the problem of the estimation of the cumulative distribution function of real random variable Y conditioned by functional random variable X by using a local linear approach. Noting that, the estimation of the conditional cumulative distribution function has great importance. It is involved in many applications, such as reliability or in survival analysis. Moreover, in prediction problem's, there are several prediction tools in nonparametric statistic, such the conditional mode, the conditional median or the conditional quantiles, are based on the preliminary estimate of this nonparametric model. The literature on this topic is quite important, (see, for instance, Roussas ([13], 1968) and Stute ([14], 1986), Deheuvels and Mason ([5], 2004) for non functional case and Ferraty and al. ([9], 2006), Laksaci and Maref ([11], 2009), Ferraty and Laksaci et al. ([8], 2010) for the functional case). However, a most of the existing literatures, in nonparametric functional data analysis, use the Nadaraya-Watson techniques as estimation method, but the latter suffer from some drawbacks, namely, in the bias term. In the nonfunctional case, the local polynomial fitting has been recognized to have superior bias properties than the kernel method (see. Fan and Gijbels ([7], 1996) for an extensive discussion on the comparison between the both methods). Recently, this estimation method has been considered for functional statistics. More precisely, the first result in this topic were obtained by Baìllo and Grané ([4], 2009). They studied the local linear estimate of the regression function when the explanatory variable takes values in a Hilbert space. The general case where the regressors are not Hilbertiann has been considered by Barrientos-Marin et al. ([1], 2010). In this work, the authors obtained the almost complete convergence (with rate) of the proposed estimate. We return to Boj et al. ([2], 2010) for an other alternative version for the functional local linear modeling. More recently, Demongeot et al. ([6], 2011) consider the local polynomial modeling of the conditional density function when the explanatory variable is functional. The aim of this paper, is to study, under general conditions, the asymptotic proprieties of the local linear estimate. More precisely, we prove the almost complete convergence (with rate) of this estimate. It should be noted that, the advantage of this mode is that implies the convergence in probability and the almost sure consistency. We note that, as usually in a functional nonparametric context, the rates of convergence obtained depend on the regularity of the model and the concentration properties on the probability measure of the underlying explanatory variable. It should be noted that, our hypotheses and results unify the both cases of finite or infinite dimension of the regressors. It is worth to note that, to the best of our knowledge the problem of the convergence almost complete or the almost sure of the conditional distribution function estimate by using local method has not been address in the multivariate case. The few existing results in the multivariate case concerns only the convergence in quadratic error. So, we can say that these asymptotic results are also new in finite dimensional case. The paper is organized as follows : In Section 2 we provide the notations and the local linear estimate of our model. We state our main results in Section 3. Finally, Section 4 is devoted to the technical proofs of the results of this paper.

2.2 Model and its estimate

Consider n pairs of random variables (X_i, Y_i) for $i = 1, \cdots, n$ that we assume drawn from the pair (X, Y) which is valued in $\mathcal{F} \times \mathsf{R}$, where $\mathcal{F}$ is a semi-metric space equipped with a semi-metric d. For $x \in \mathcal{F}$, we assume that there exists a regular version of the conditional probability of Y given $X = x$ that is absolutely continuous with respect to Lebesgue measure on R. Our purpose is to study the local linear estimate of the conditional cumulative distribution function, (cdf) of Y given $X = x$, denoted by $F^x(y) = \mathsf{E}[\mathsf{1}_{Y \quad y}/X = x]$. In this functional context, we use the idea's of Barenties-Martin et al. ([1], 2010) for the nonparametric functional regression, to define a functional local linear estimate of $F^x(y)$, based on the minimization with respect to (w.r.t.) (a,b) of

$$\sum_{i=1}^{n}\left(1_{\{Y_i \le y\}} - a - b\beta(X_i, x)\right)^2 K(h^{-1}\delta(x, X_i)) \tag{2.1}$$

With $\beta(.,.)$ and $\delta(.,.)$ are known functions from $\mathcal{F}$ into R, K is a kernel and $h_K = h_{K,n}$ is chosen as a sequence of positive real numbers. More precisely, the functional local linear estimate of F^x is defined by :

$$\widehat{F}^{x_0}(y) = \widehat{a} + \widehat{b}\beta(x_0, x)$$

With $\hat{a}$ and $\hat{b}$ are solutions of (2.1). However, if the bi-functional operator β such that, $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$, the quantity $\widehat{F}^x(y)$ is explicitly defined by

$$\widehat{F}^x(y) = \frac{\displaystyle\sum_{j=1}^{n}\sum_{i=1}^{n} W_{ij}(x) 1_{Y_j \le y}}{\displaystyle\sum_{j=1}^{n}\sum_{i=1}^{n} W_{ij}(x)} \tag{2.2}$$

Where

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

with the convention $0/0 = 0$.

2.3 Main Results

In the following x (resp. y) will be a fixed point in $\mathcal{F}$ (resp. R), N_x will denote a fixed neighborhood of x, and $\phi_x(r_2, r_1) = P(r_1 \le \delta \le r_2)$. Moreover, we put, for all $i = 1, \cdots, n$:

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

Without loss of generality, we assume that our nonparametric model satisfies the following conditions :

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

H2 The conditional distribution function F^x is such that : there exist $b > 0$, $\forall (x_1, x_2) \in N_x \times N_x$

$$|F^{x_1}(y) - F^{x_2}(y)| \leq C \left| \delta(x_1, x_2) \right|^b,$$

Where C is a positive constant depending on x.

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

$$\forall x' \in \mathcal{F}, \quad C_1 |\delta(x', x)| \leq |\beta(x, x')| \leq C_2 |\delta(x, x')|, \quad \text{Where} \quad C_1 > 0, C_2 > 0.$$

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

H5 The bandwidth h satisfies : there exists an integer n_0, such that

$$\forall n > n_0, -\frac{1}{\phi_x(h)} \int_{-1}^{1} \phi_x(zh,, h) \frac{d}{dz}(z^2 K(z)) dz > C_3 > 0, \quad \lim_{n \to \infty} \frac{\log n}{n \phi_x(h)} = 0$$

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.

We recall that the conditions (H1), (H3) (H4) and (H5) are used and commented by Barrientos-Marin and al. ([1], 2010). While, assumptions (H2) is a simple 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. The following theorem gives the almost-complete convergence

Theorem 2.1. *Under assumptions $(H1) - (H5)$, we have that :*

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

Remark that, the proof of Theorem (2.1) is a direct consequence of the decomposition :

$$\begin{aligned} \widehat{F}^x(y) - F^x(y) &= \frac{1}{\widehat{F}_D^x} \left\{ \left(\widehat{F}_N^x(y) - \mathsf{E}(\widehat{F}_N^x(y)) - (F^x(y) - \mathsf{E}(\widehat{F}_N^x(y))) \right) \right\} \\ &+ \frac{F^x(y)}{\widehat{F}_D^x} \left(1 - \widehat{F}_D^x \right) \end{aligned}$$

Where

$$\widehat{F}_N^x(y) = \frac{1}{n(n-1)\mathsf{E}[W_{12}(x)]} \sum_{i\neq j} W_{ij}(x)\mathbb{1}_{Y_j \le y}$$

and

$$\widehat{F}_D^x(y) = \frac{1}{n(n-1)\mathsf{E}[W_{12}(x)]} \sum_{i\neq j} W_{ij}$$

and of Lemmas (2.3.1), (2.3.2) and (2.3.3) bellow, for which the proofs are given in the Appendix.

Lemma 2.3.1. *([1], 2010) Under assumptions (H1), (H3), (H4) and (H5), we have that :*

$$1 - \widehat{F}_D^x = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co.$$

and

$$\exists \delta < 0, \quad \text{such that} \quad \sum_{i=1}^{\infty} \mathsf{P}\left(\widehat{F}_D^x < \delta\right) < \infty.$$

Lemma 2.3.2. *Under assumptions (H1), (H2) and (H4), we obtain :*

$$\left| F^x(y) - \mathsf{E}[\widehat{F}_N^x(y)] \right| = O(h^b).$$

Lemma 2.3.3. *Under assumptions of Theorem (2.1), we get :*

$$\left| \widehat{F}_N^x(y) - \mathsf{E}[\widehat{F}_N^x(y) \right| = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co.$$

2.4 Appendix

In what follows, when no confusion is possible, we will denote by C and C' some strictly positive generic constants.

Proof of the lemma 2.3.2 : Since the pairs (X_i, Y_i) are identically distributed, then from assumption $(H4)$ we obtain :

$$\mathsf{E}[\widehat{F}_N^x(y)] = \frac{1}{\mathsf{E}[W_{12}]} \mathsf{E}[W_{12}(x)[\mathsf{E}[\mathbb{1}_{Y \le y}/X]]].$$

First note that $W_{12}(x) = W_{12}(x)\mathbb{1}_{B(x,h)}(X_1)$ under $(H4)$. So, we deduce from $(H4)$ that

$$F^x(y) - \mathsf{E}[\hat{F}_N^x(y)] = \frac{1}{\mathsf{E}(W_{12}(x))} \mathsf{E}[W_{12}(x)\mathbb{1}_{B(x,h)}(X_1)(F^x(y) - F^{X_1}(y))]$$

Then by $(H2)$ we have

$$\mathbf{1}_{B(x,h)}(X_1)|F^x(y) - F^{X_1}(y)| \leq Ch^b,$$

this last inequality yield the proof,

Proof of lemma 2.3.3. Similarly to the proof of Lemma 2 "Barrientos et al. (2010,[1])" we consider the following decomposition

$$\hat{F}_N^x(y) = T_1(T_2T_3 - T_4T_5) \tag{2.3}$$

where

$$T_1 = \frac{n^2h^2\varphi_x(h)^2}{n(n-1)\mathsf{E}(W_{12})} \quad T_2 = \frac{1}{n}\sum_{j=1}^{n}\frac{K_j\mathbf{1}_{Y_j\ \ y}}{\varphi_x(h)} \quad T_3 = \frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i^2}{h^2\varphi_x(h)}$$

$$T_4 = \frac{1}{n}\sum_{j=1}^{n}\frac{K_j\beta_j\mathbf{1}_{Y_j\ \ y}}{h\varphi_x(h)}, \quad T_5 = \frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i}{h\varphi_x(h)}$$

Thus, we can write

$$\hat{F}_N^x(y) - \mathsf{E}(\hat{F}_N^x(y)) = T_1\left((T_2T_3 - \mathsf{E}(T_2T_3)) - (T_4T_5 - \mathsf{E}(T_4T_5))\right)$$

Observe that :

$$\begin{aligned} T_2T_3 - \mathsf{E}(T_2T_3) &= (T_2 - \mathsf{E}(T_2))(T_3 - \mathsf{E}(T_3)) + (T_3 - \mathsf{E}(T_3))\mathsf{E}(T_2) \\ &+ (T_2 - \mathsf{E}(T_2))\mathsf{E}(T_3) + \mathsf{E}(T_2)\mathsf{E}(T_3) - \mathsf{E}(T_2T_3) \end{aligned}$$

and in the same way :

$$\begin{aligned} T_4T_5 - \mathsf{E}(T_4T_5) &= (T_4 - \mathsf{E}(T_4))(T_5 - \mathsf{E}(T_5)) + (T_5 - \mathsf{E}(T_5))\mathsf{E}(T_4) \\ &+ (T_4 - \mathsf{E}(T_4)\mathsf{E}(T_5)) + \mathsf{E}(T_4)\mathsf{E}(T_5) - \mathsf{E}(T_4T_5) \end{aligned}$$

So, the claimed result can be derived from the following assertions

$$T_i - \mathsf{E}(T_i) = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad \text{for} \quad i = 2,3,4,5 \tag{2.4}$$

$$T_1 = O(1) \quad \text{and} \quad \mathsf{E}(T_l) = O(1), \quad \text{for} \quad l = 2,3,4 \tag{2.5}$$

and

$$Cov(T_2, T_3) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \tag{2.6}$$

$$\text{and} \quad Cov(T_4, T_5) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \tag{2.7}$$

Note first that (2.5) was already stated by Barrientos and al. ([1],2010). Therefore, it suffices to prove (2.4), (2.6) and (2.7) to finish the proof of the Lemma.

Concerning (2.4), we put

$$S_{l,k} - \mathsf{E}(S_{l,k}) = \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} Z_i^{l,k} \quad \text{for} \quad l = 0,1,2 \quad \text{and} \quad \text{k=0,1}$$

Where $Z_i^{l,k} = \dfrac{1}{h^l \phi_x(h)}(K_i(x)\mathbf{1}^k_{Y_i \ y}\beta_i^l(x) - \mathsf{E}(K_i(x)\mathbf{1}^k_{Y_i \ y}\beta_i^l(x)))$, for $l = 0,1,2$ and $k = 0,1$. Thus, it remains to check that

$$S_{l,k} - \mathsf{E}(S_{l,k}) = 0_{a.co}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \quad \text{for} \quad \text{l=0,1,2} \quad \text{and} \quad k = 0,1$$

Under (H3), we have $(K_i(x)\beta_i^l(x))/h^l < C$, then

$$\left|Z_i^{l,k}\right| \leq \frac{C}{\phi_x(h)} \quad \text{and} \quad \mathsf{E}(Z_i^{l,k^2}) \leq \frac{C^0}{\phi_x(h)}$$

So, Thus, the use of the classical Bernstein's inequality (Uspensky [15] 1937, p205) allows us to write for all $\eta \in (0, \frac{C'}{C})$

$$\mathsf{P}\left\{|S_{l,k} - \mathsf{E}(S_{l,k})| > \eta\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} \leq C^0 n^{\ \ C\eta^2}$$

Finally, an appropriate choice of η permits to deduce that :

$$\mathsf{P}\left\{|S_{l,k} - \mathsf{E}(S_{l,k})| > \eta\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} \leq C^0 n^{\ 1 \ \ \gamma}, \text{for} \quad \text{l=0,1,2,} \quad \text{and} \quad k = 0,1$$

Now, we proceed in proving the results of (2.6) and (2.7). Since the pairs (X_i, Y_i), $i = 1, \cdots, n$ are identically distributed, we obtain that :

$$\begin{cases} Cov(T_2, T_3) & = & \dfrac{1}{nh^2\phi_x^2(h)}[\mathsf{E}(K_1^2\mathbf{1}_{Y_1 \ y}\beta_1^2) - \mathsf{E}(K_1\mathbf{1}_{Y_1 \ y})\mathsf{E}(K_1\beta_1^2))] \\ \text{and} \quad Cov(T_4, T_5) & = & \dfrac{1}{nh^2\phi_x^2(h)}[\mathsf{E}(K_1^2\mathbf{1}_{Y_1 \ y}\beta_1^2) - \mathsf{E}(K_1\beta_1\mathbf{1}_{Y_1 \ y})\mathsf{E}(K_1\beta_1))] \end{cases}$$

So, for both quantities, we have to evaluate :

$$\mathsf{E}(K_i(x)\mathbf{1}^k_{Y_i \quad y}\beta^l_i(x)), \quad \text{for} \quad l = 0,1,2 \quad \text{and} \quad k = 0,1$$

For all $l = 0,1,2$, and $k = 0,1$, we have :

$$\mathsf{E}(K_i(x)\mathbf{1}^k_{Y_i \quad y}\beta^l_i(x)) = o(\mathsf{E}(K_i(x)\beta^l_i(x)))$$

And by Lemma 3 in ([1], 2010), we obtain that :

$$\mathsf{E}(K_i(x)\mathbf{1}^k_{Y_i \quad y}\beta^l_i(x)) = o(h^l\phi_x(h)))$$

Which implies that :

$$Cov(T_2,T_3) = O\left(\frac{1}{n\phi_x(h)}\right) = o\left(\frac{\log n}{n\phi_x(h)}\right)$$

$$\text{and} \quad Cov(T_4,T_5) = O\left(\frac{1}{n\phi_x(h)}\right) = o\left(\frac{\log n}{n\phi_x(h)}\right)$$

Acknowledgements

The authors thank the anonymous referees immensely for a careful reading of the paper

Bibliographie

[1] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *Journal of Nonparametric Statistics,* 22, No.5, Pages 617-632.

[2] Boj, E. Delicado, P. and Fortiana, J. (2010). Distance-based local linear regression for functional predictors Computational Statistics and Data Analysis, 54, 429-437

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

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

[5] Deheuvels P., Mason D.M., (2004). General Asymptotic bands based on Kernel-type function estimators, Statist. Inf. for Stoch. Proc. 7, 225-277. MR2111291.

[6] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2011). Functional data : Local linear estimation of the conditional density and its application *Statistics*, in presse.

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

[8] 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.

[9] Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Journal of Stat.Inference Stoch. Process.*, 9, No. 1, 47-76.

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

[11] Laksaci, M. and Maref, F. (2009). Conditional cumulative distribution estimation and its applications *Journal of probability and statistical sciences*, 13, No 2, Pages 47-56.

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

[13] Roussas, G. G., (1968). On some properties of nonparametric estimates of probability density functions. Bull. Soc. Math. Grèce (N.S.), 9, Pages 29-43.

[14] Stute, W., (1986). On Almost Sure Convergence of Conditional Empirical Distribution Functions. The Annals of Probability, 14, 891-901.

[15] Uspensky, J. V. (1937). *Introduction to Mathematical Probability*. McGraw-Hill, New York.

Chapitre 3

Semiparametric locally modelling of the conditional distribution function for functional dependent Data

This chapter deals with the almost convergence complete of the functional linear estimate of conditional distribution function when the observations are dependent..

Semiparametric locally modelling of the conditional distribution function for functional dependent Data

Ali Laksaci
Laboratoire de Mathématiques,
Université de Sidi Bel-Abbès,
Bp 89 Sidi Bel Abbès 22000. Algérie
alilak@yahoo.fr

Nawal Hachemi
Département de Mathématiques,
Université Moulay Tahar, Saida. 20000
hacnawel@yahoo.fr

Abstract

The main goal of this paper is to study the local linear estimate of the conditional distribution function when the explanatory variable X is valued in the functional space $\mathcal{F}$ of eventually infinite dimension and when the observations $(X_i, Y_i)_{(i=1,\cdots,n)}$ are strongly mixing. Under some general conditions, we establish the almost complete convergence (with rate) of the estimate

Keyboard : Mixing Data, Functional Data Analysis, Local Linear Estimation, Semi-metric space, Small balls Probability, Conditional distribution function.

3.1 Introduction

Let $(X_1, Y_1)_{i=1, \quad ,n}$ be a stationary $\alpha-$mixing process valued in $\mathcal{F} \times \mathcal{R}$ where $\mathcal{F}$ is a semi-metric vector space, $d(.,.)$ denoting the semi-metric. In the remaining of the paper x is fixed in $\mathcal{F}$ and N_x denotes a neighborhood of x. We suppose that exists a regular version of the conditional probability of Y given $X = x$ exists for any $x \in N_x$. In this paper, we generalize the result of Laksaci Ali and Hachemi Nawal ([2], 2012) of the cumulative conditional distribution function "$F^x(y)$" by using the local linear approach when the explanatory variable is functional and the observations are dependent. We estimate the conditional distribution function by $\hat{a}$ which is obtained by minimizing the following quantity :

$$\sum_{i=1}^{n} \left(\mathbb{1}_{Y_i \quad y} - a - b\beta(X_i, x)\right)^2 K(h_K^{\ 1}\delta(X_i, x))$$

Where $\beta(.,.)$ and $\delta(.,.)$ are a known functions from $\mathcal{F}$ into $\mathcal{R}$ such that ; $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$ and $d(.,.) = |\delta(.,.)|$. The function $K(.)$ is a positive kernel and $h_K = h_{K,n}$ is a sequence of positive real numbers. Our purpose is to estimate the conditional distribution function $\hat{F}^x$ of F^x as follows :

$$\hat{F}^x(y) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x)\mathbb{1}_{Y_j \quad y}}{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(x)}$$

Where

$$W_{ij}(x) = \beta(X_i, x)(\beta(X_i, x) - \beta(X_j, x))K(h_K^{\ 1}\delta(x, X_i))K(h_K^{\ 1}\delta(x, X_j))$$

with the convention $0/0 = 0$.

This method is a popular extension that does not suffer from the bias of the widely of Nadaraya Waston ([18], 1964) which behaves some how badly. Local polynomial regression (see Clevand ([6], 1979) among others) uses weighted least squares (WLS) regression to fit a p^{th} degree polynomial ($p \neq 0$) to data. Hastie and Loader ([13], 1993) showed that local polynomial regression addresses the boundary problem of potentially inflated biais and variance. The local linear smoother (Fan 1992, 1993) is obtained via fitting a data set locally with a linear function. This idea's has been developed in the regression context for univariate or multivariate explanatory variables (See Wand and Jones ([22], 1995)) for an overview on this topic). Therefore, it has recently been by Yu and Jones ([23], 1998) via an application of the double kernel, linear approach of Fan, Yao and Tong ([12],1996), Fan and Gijbels ([9], 1996), Li and Racine ([16], 2007) recently established the pointwise asymptotic distribution (central limit theorem) for the local linear estimator of a nonparametric regression function. There has been rich literature on the uniform convergence rates for the local linear estimator under mixing conditions, see Masry ([17], 1996), Fan and Yao ([8], 2003) and Hansen ([15], 2008) for example. Very recently, local linear regression models have been investigating when the explanatory variable is a function predictor (Aneiros-Pérez, Cao, and Vilar-Fernandez ([1], 2008)) ; Boj, Delicado, and Fortiana ([5], 2008) ; and Baíllo and Grané ([4], 2009) for more details) and Lu and Linton ([16], 2007) established the pointwise asymptotic distribution for the local linear estimators. Barrientos ([3], 2010) study the fast functional locally modeling for the regression. Indeed, the study of statistical models adapted to such type of infinite dimensional data has been several works in the recent literature, the effects of this space will be seen through the concentration properties on small balls of the probability measure of this function variable. This paper is related to the conditional distribution function in a situation when the explanatory are functional nature feature in the sense that it is assumed to take its values in some abstract semi-metric space. More precisely, we focus for the conditional distribution function of a real random variable given a functional random variable under mixing assumptions, we establish the almost complete convergence of this estimate. This paper is organized as follows : In section 2, we put the notations of dependence and hypotheses using in the sense, then, show its almost complete convergence. The appendix contains all technical details of the result.

3.2 Notation and Assumptions

In order to characterize the model of dependence between the pairs $(X_i, Y_i)_{(i=1, \quad ,n)}$, we will use the notion of α mixing sequence presented in the literature.

Let us recall the following definition. Let $\mathcal{F}_i^k(Z)$ denotes the $\sigma-$algebra generated by $\{Z_j, i \leq j \leq k\}$.

Definition 3.2.1. *([19], 1994) Let $\{Z_i, i = 1, 2, \cdots\}$ denote a sequence of random variable. Given a positive integer n, set :*

$$\alpha(n) = \sup\{|\mathsf{P}(A \cap B) - \mathsf{P}(A)\mathsf{P}(B)| : A \in \mathcal{F}_1^k(Z), B \in \mathcal{F}_{k+n}^{1}(Z), k \in \mathsf{N} \ \}$$

The sequence is said to be $\alpha-$mixing (Strong mixing) if the mixing coefficient $\alpha(n) \to 0$ as $n \to \infty$.

3.3 Hypotheses

In this paper, x denotes a fixed point in $\mathcal{F}$, N_x notes a fixed neighborhood of x such as

$$B(x, r) = \{x^0 \in \mathcal{F} / |\delta(x, x^0)| \leq r\} \quad \text{and} \quad \phi_x(r_1, r_2) = \mathsf{P}(r_2 \leq \delta(X, x) \leq r_1)$$

Then, we consider the following hypotheses :

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

H2 For the conditional distribution function F^x :

$$\exists b > 0, \quad \forall (x_1, x_2) \in N_x \times N_x, \quad |F^{x_1}(y) - F^{x_2}(y)| \leq C_x |\delta(x_1, x_2)|^b$$

Where C_x is a positive constant depending on x.

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

$$\forall x^0 \in \mathcal{F}, \quad C_2|\delta(x, x^0)| \leq |\beta(x^0, x)| \leq C_2|\delta(x, x^0)|, \quad \text{where} \quad C_1 > 0, \quad C_2 > 0.$$

H4 The sequence $(X_i, Y_i)_{(i=1, \quad ,n)}$ is an α-mixing sequence with mixing coefficient $\alpha(n)$ such as

$$\exists a > 0, \quad \exists c > 0, \quad \forall n \in \mathsf{N}, \quad \alpha(n) \leq cn^{\quad a}$$

And

$$\sup_{i\neq j} \mathsf{P}((X_i, X_j) \in B(x,h) \times B(x,h)) \leq C\phi_x(h)^{a/(a-1)}, \quad \text{with} \quad a > 2.$$

H5 K is a positive, differentiable function with support $[-1,1]$.

H6 The bandwidth h satisfies : there exists an anteger n_0, such that

$$\exists n_0, \forall n > n_0, -\frac{1}{\varphi_x(h)} \int_{-1}^{+1} \left(\varphi_x(uh,h) \left(\frac{d}{du} u^2 K(u) \right) \right) du > C > 0$$

and

$$h_K \int_{B(z,h_K)} \beta(u,x) d\mathsf{P}(u) = 0 \left(\int_{B(z,h_K)} \beta^2(u,x) dP(u) \right)$$

where $B(x,r) = \{x' \in \mathcal{F}/d(x',x) \leq r\}$ and $d\mathsf{P}(x)$ the cumulative function

H7 $$\begin{cases} \lim_{n\to+\infty} h_K = 0 & \lim_{n\to+\infty} \frac{\log n}{n\phi_x(x)} = 0 \\ \text{and} & \exists \eta > 0, \quad \phi_x(h) \geq n^{\frac{3-a}{a+1}+\eta} \end{cases}$$

- **Commentaries** : (H1), (H3), (H5) and (H6) is the same as ([2], 2010)

- (H1) Is needed to deal with the functional nonparametric of our model by controlling the concentration properties of the probability measure of the variable X.
- (H2) is a simple regularity condition which characterizes the functional space of our model and needed to evaluate the bias term in the asymptotic results of this paper.
- (H3) is a mild regularity condition permits to control the shape of locating function β
- (H4) is a standard choice for the mixing coefficient in time series and the second measure the local dependence of the observations.

The assumptions $(H4)$ and $(H7)$ are added to avoid the covariance term. However, the speed of convergence will be slow compared to the speed of independent case. Well, we establish the almost complete convergence with the same precision, but, under stronger conditions. The other conditions are standard technical and it is imposed only for the brevity of proofs.

Theorem 3.1. *Under the conditions (H1) -(H7), we have*

$$|\hat{F}^x(y) - F^x(y)| = o(h^{b_1}) + o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co.$$

The demonstration is based on the following decomposition

$$\begin{aligned}\hat{F}^x(y) - F^x(y) &= \frac{1}{\hat{F}_D^x}\left\{(\hat{F}_N^x(y) - \mathsf{E}(\hat{F}_N^x(y)) - (F^x(y) - \mathsf{E}(\hat{F}_N^x(y))\right\} \\ &- \frac{F^x(y)}{\hat{F}_D^x}\left\{1 - \hat{F}_D^x\right\}\end{aligned}$$

Where

$$\hat{F}_N^x(y) = \frac{1}{n(n-1)\mathsf{E}(W_{12})}\sum_{j=1}^{n}\sum_{i=1}^{n} W_{ij}\mathsf{1}_{Y_j \quad y}$$

And

$$\hat{F}_D^x = \frac{1}{n(n-1)\mathsf{E}(W_{12})}\sum_{j=1}^{n}\sum_{i=1}^{n} W_{ij}$$

So, the proof is a consequence of the following lemmas and corollary. For $x \in \mathcal{F}$, $\mathsf{E}(\hat{F}_D) = 1$

Lemma 3.3.1. *If the hypotheses (H1), (H3) -(H7) are satisfied, then we have*

$$\hat{F}_D^x - 1 = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co \tag{3.1}$$

And

$$\sum_{i=1}^{n} \mathsf{P}(\hat{F}_D^x(y) < \epsilon) < \infty \tag{3.2}$$

Lemma 3.3.2. *Under the assumptions (H1), (H2) and (H4) , we obtain :*

$$|F^x(y) - \mathsf{E}(\hat{F}_N(y))| = o(h_K^b) \tag{3.3}$$

Lemma 3.3.3. *Suppose that the conditions* $(H1)$, $(H3) - (H7)$*of Theorem are satisfied. Then, we have*

$$|\hat{F}_N^x(y) - \mathsf{E}(\hat{F}_N^x(y))| = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.co \tag{3.4}$$

3.4 Appendix

The proof is essentially based on the same techniques used for the independent case ([7], 2012). Recall that the assumptions of comments have no influence on the party through the convergence rate, under this conclusion we will evaluate only the part of dispersion.

Proof of lemma 3.3.1. We start by writing

$$\hat{F}_D^x = \underbrace{\frac{n^2 h_K^2 \phi_x(h)^2}{n(n-1)\mathsf{E}(W_{12})}}_{B_1} \left[\begin{array}{c} \underbrace{\left(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j}{\phi_x(h)}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i^2}{h_K^2\phi_x(h)}\right)}_{B_2B_3} \\ -\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i\beta_i}{h_K\phi_x(h)}\right)^2}_{B_4} \end{array} \right]$$

Observe that

$$\hat{F}_D^x - \mathsf{E}(\hat{F}_D^x) = B_1(B_2B_3 - \mathsf{E}(B_2B_3)) - (B_4^2 - \mathsf{E}(B_4^2)) \tag{3.5}$$

It suffices to follow step by step the proof of the lemma

$$\begin{aligned} B_2B_3 - \mathsf{E}(B_2B_3) &= (B_2 - \mathsf{E}(B_2))(B_3 - \mathsf{E}(B_3)) + (B_3 - \mathsf{E}(B_3)\mathsf{E}(B_2) \\ &+ (B_2 - \mathsf{E}(B_2))\mathsf{E}(B_3) + \mathsf{E}(B_2)\mathsf{E}(B_3) - \mathsf{E}(B_2B_3) \end{aligned}$$

And

$$B_4^2 - \mathsf{E}(B_4^2) = (B_4 - \mathsf{E}(B_4))^2 + 2(B_4 - \mathsf{E}(B_4))\mathsf{E}(B_4) + \mathsf{E}^2(B_4) - \mathsf{E}(B_4^2)$$

So, we prove the following

$$\sum_n \mathsf{P}\left\{|B_l - \mathsf{E}(B_l)| > \epsilon\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} < \infty, \quad \text{for} \quad l = 2,3,4. \tag{3.6}$$

$$B_1 = O(1) \quad \text{and} \quad \mathsf{E}(B_l) = O(1) \quad \text{for} \quad l = 2,3,4, \tag{3.7}$$

$$Cov(B_2, B_3) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \tag{3.8}$$

$$Var(B_4) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right) \tag{3.9}$$

It is shown that Barrientos and al. ([2], 2010) that

$$B_1 = O(1) \quad \text{and} \quad \mathsf{E}(B_l) = O(1) \quad \text{for} \quad l = 2, 3, 4.$$

For (3.6), let us

$$\Delta_i^k = \frac{1}{h_K^k}\left(K_i(x)\beta_i^k - \mathsf{E}(K_i(x)\beta_i^k)\right) \quad \text{for} \quad k = 0, 1, 2$$

It is easy to see that

$$B_{k+2} - \mathsf{E}(B_{k+2}) = \frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\Delta_i^k \quad \text{for} \quad k = 0, 1, 2.$$

The Fuk-Nagaev inequality (Rio (1999, [20]), formula) allows one to get, for all ϵ and $r \geq 1$:

$$\begin{aligned}
\mathsf{P}\{|B_{k+2} - \mathsf{E}(B_{k+2})| > \epsilon\} &= \mathsf{P}\left(\left|\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\Delta_i^k\right| > \epsilon\right), \quad \text{k=0,1,2} \\
&\leq \mathsf{P}\left(\left|\sum_{i=1}^{n}\Delta_i^k\right| > \epsilon n\phi_x(h)\right) \\
&\leq C\left(\frac{n}{r}\left(\frac{r}{n\epsilon\phi_x(h)}\right)^{a+1} + \left(1 + \frac{n^2\epsilon^2\phi_x(h)^2}{rS_n^2}\right)^{r/2}\right) \\
&= C(R_1 + R_2)
\end{aligned}$$

Where

$$S_n^2 = \sum_{1 \; j \; n}\sum_{1 \; i \; n}|Cov(\Delta_i^k, \Delta_j^k)| = nVar(\Delta_1^k(x)) + S_n^{cov}$$

With

$$S_n^{cov} = \sum_{i \neq j}\sum_{i=1}^{n}Cov(\Delta_i^k, \Delta_j^k)$$

By taking

$$r = C(\log n)^2 \quad \text{and} \quad \epsilon = \epsilon_0\sqrt{\frac{\log n}{n\phi_x(h)}}, \quad \text{for all} \quad \epsilon_0 > 0$$

Indeed, we use the following developed by Masry ([7], 1986), such as we split the sum in two parts defined by,

$$S_1(x) = \{(i,j) \quad \text{such that} \quad 1 \leq i-j \leq m_n\}$$

And

$$S_2(x) = \{(i,j) \quad \text{such that} \quad m_n + 1 \leq i-j \leq n-1\}$$

Where $m_n \to \infty$, as $n \to \infty$, with

$$S_n^{cov} \leq \sum_{S_1} |Cov(\Delta_i^k, \Delta_j^k)| + \sum_{S_2} |Cov(\Delta_i^k, \Delta_j^k)|$$

For all $i \neq j$

$$\begin{aligned}
|Cov(\Delta_i^k, \Delta_j^k)| &= h_K^{2k} |\mathsf{E}(K(h_K^{1}\delta(x,X_i))K(h_K^{1}\delta(x,X_j)\beta_i^k\beta_j^k) \\
&- \mathsf{E}(K(h_K^{1}\delta(x,X_i)\beta_i^k)\mathsf{E}(K(h_K^{1}\delta(x,X_j)\beta_j^k)| \\
&\leq C\left|\mathsf{E}\left(K(h_K^{1}\delta(x,X_i)K(h_K^{1}\delta(x,X_j))\right)\right| \\
&+ \mathsf{E}\left|(K(h_K^{1}\delta(x,X_i)\right| \mathsf{E}\left|(K(h_K^{1}\delta(x,X_j)\right| \\
&\leq C(\mathsf{E}\left(\mathbf{1}_{B(x,h_K)}(X_i)\mathbf{1}_{B(x,h_K)}(X_j)\right)) \\
&+ C\mathsf{E}(\mathbf{1}_{B(x,h_K)}(X_i)\mathsf{E}(\mathbf{1}_{B(x,h_K)}(X_j) \\
&\leq C(\mathsf{P}(X_i,X_j) \in B(x,h_K) \times B(x,h_K) \\
&+ \mathsf{P}(X_i \in B(x,h_K))\mathsf{P}(X_j \in B(x,h_K)) \\
&\leq C(\phi_x(h)^{a/(a\ \ 1)} + \phi_x(h)^2) \\
&\leq C\phi_x(h)^{a/(a\ \ 1)}
\end{aligned}$$

The last inequality is obtained by the hypotheses (H3), (H4) and (H5), so

$$\sum_{(i,j)2S_1} Cov(\Delta_i^k, \Delta_j^k) \leq Cnm_n\phi_x(h)^{a/(a\ \ 1)} \tag{3.10}$$

On the other hand, to studied the sum over S_2, the covariance can be controlled by means of the usual Davydov's covariance inequality for mixing processes (See Rio ([20], 1999), formula 1.12a).

$$\forall i \neq j, \quad |Cov(\Delta_i^k, \Delta_j^k)| \leq C|i-j|^{\ \ a} \tag{3.11}$$

We use the assumption $(H4)$ to study the following sums.

$$\left| \sum_{(i,j)2S_2} Cov(\Delta_i^k, \Delta_j^k) \right| \leq C \frac{n{m_n}^{a+1}}{a-1} \tag{3.12}$$

By choosing $m_n = (\phi_x(h))^{\ 1/(a\ \ 1)}$ to according (3.10) and (3.12), we get

$$S_n^{cov} = O(n\phi_x(h)) \tag{3.13}$$

For the variance term, using (H3) and (H5) in order to write

$$Var(\Delta_1^k) \leq C(\phi_x(h) + \phi_x(h)^2) \leq C\phi_x(h) \tag{3.14}$$

Then by combining (3.13) with (3.14), we can get directly :

$$S_n^2 = O(n\phi_x(h)) \tag{3.15}$$

Now, we take $r = C(\log n)^2$ and $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{n\phi_x(h)}}$ to obtain

$$R_1 \leq C(\log n)^{(3a\ \ 1)/2} n^{(1\ \ a)/2} \phi_x(h)^{\ \ (a\ \ 1)/2}$$

Note that under $(H7)$, it exists some real $\nu > 0$ such that

$$R_1 \leq Cn^{\ \ 1\ \ \nu} \tag{3.16}$$

Hence, for some real $\nu > 0$, R_1 is bounded by the term of a finite series.

$$\begin{aligned} R_2 &= \exp\left(-\frac{C(\log n)^2}{2} \log\left(1 + \frac{\epsilon_0^2}{C \log n}\right)\right) \\ &= \exp\left(-(\epsilon_0^2/2) \log n \frac{\log\left(1 + \frac{\epsilon_0^2}{C\log n}\right)}{\frac{\epsilon_0^2}{C \log n}} \right) \end{aligned}$$

The fact that $\lim \log(1+x)/x = 0$ when x tends to 0, we can find under a suitable choice of all $\epsilon_0 > 0$ large enough and $\exists \nu > 0$ such that

$$R_2 \leq Cn^{\ \ 1\ \ \nu} \tag{3.17}$$

This allows us to deduce that

$$\sum_n \mathsf{P}\left\{|B_l - \mathsf{E}(B_l| > \epsilon\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} < \infty$$

Finally, the following similar arguments used to prove (3.15) we get

$$Cov(B_2, B_3) = O\left(\frac{1}{n\phi_x(h)}\right)$$

And

$$Var(B_4) = o\left(\frac{1}{n\phi_x(h)}\right)$$

Which is negligible with respect $(\sqrt{\log n/n\phi_x(h)})$. Then, the proof of our Lemma is now finished.

Proof of lemma 3.3.3. The proof of this lemma is similar to the one used for the lemma (3.1)

$$\hat{F}_N^x(y) = \underbrace{\frac{n^2h_K^2\varphi_x(h)^2}{n(n-1)\mathsf{E}(W_{12})}}_{T_1}\left[\begin{array}{c}\underbrace{\left(\frac{1}{n}\sum_{j=1}^n\frac{K_j\mathbf{1}_{\{Y_j \leq y\}}}{\phi_x(h)}\right)\left(\frac{1}{n}\sum_{i=1}^n\frac{K_i\beta_i^2}{h_K^2\phi_x(h)}\right)}_{T_2T_3}\\ -\underbrace{\left(\frac{1}{n}\sum_{j=1}^n\frac{K_j\beta_j\mathbf{1}_{\{Y_j \leq y\}}}{h_K\phi_x(h)}\right)\left(\frac{1}{n}\sum_{i=1}^n\frac{K_i\beta_i}{h_K\phi_x(h)}\right)}_{T_4T_5}\end{array}\right]$$

We use again the same kind of decomposition as in (3.5),

$$\hat{F}_N^x(y) - \mathsf{E}\hat{F}_N^x(y) = T_1\left((T_2T_3 - \mathsf{E}(T_2T_3)) - (T_4T_5 - \mathsf{E}T_4T_5)\right)$$

It suffices to follow step by step the proof of this lemma.

$$\begin{aligned}T_2T_3 - \mathsf{E}(T_2T_3) &= (T_2 - \mathsf{E}(T_2))(T_3 - \mathsf{E}(T_3)) + (T_3 - \mathsf{E}(T_3))\mathsf{E}(T_2)\\ &+ (T_2 - \mathsf{E}(T_2))E(T_3) + \mathsf{E}(T_2)\mathsf{E}(T_3) - \mathsf{E}(T_2T_3)\end{aligned}$$

And

$$\begin{aligned}T_4T_5 - \mathsf{E}(T_4T_5) &= (T_4 - \mathsf{E}(T_4))(T_5 - \mathsf{E}(T_5)) + (T_5 - \mathsf{E}(T_5))\mathsf{E}(T_4)\\ &+ (T_4 - \mathsf{E}(T_4))E(T_5) + \mathsf{E}(T_4)\mathsf{E}(T_5) - \mathsf{E}(T_4T_5)\end{aligned}$$

The result for this lemma is a direct consequence of the following assertions

$$\sum_n \mathsf{P}\left\{[T_k - \mathsf{E}(T_k)] > \epsilon\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} < \infty \quad \text{for} \quad k = 2, 3, 4, 5 \tag{3.18}$$

$$T_k = O(1) \quad \text{for} \quad \mathsf{E}(T_k) = O(1), \quad k = 2, 3, 4, 5 \tag{3.19}$$

$$Cov(T_2, T_3) = o\left(\sqrt{\frac{\log n}{n\phi_x(x)}}\right), \quad k = 2, 3, 4, 5 \tag{3.20}$$

$$Cov(T_4, T_5) = o\left(\sqrt{\frac{\log n}{n\phi_x(x)}}\right), \quad k = 2, 3, 4, 5 \tag{3.21}$$

The same arguments used when $k = 3, 5$ in (3.1). Then, we have only $k = 2, 4$

$$T_{k+2} - \mathsf{E}(T_{k+2}) = \frac{1}{n\phi_x(h)}\sum_{i=1}^{n}\Gamma_i^{l,k} \quad \text{for} \quad k = 0, 1, 2 \quad \text{and} \quad l = 0, 1, 2.$$

Where $\Gamma_i^{l,k} = \frac{1}{h_K^l}\left(K_i(x)\mathbb{1}^k_{Y_i \quad y}\beta_i^l(x) - \mathsf{E}(K_i(x)\mathbb{1}^k_{Y_i \quad y}\beta_i^l(x)\right)$.
To treat the covariance term, we are going to use an inequality of exponential type ([5], 2004) in which the following quantity appears

$$S_n^2 = \sum_{j=1}^{n}\sum_{i=1}^{n} Cov(\Gamma_i^{l,k}, \Gamma_j^{l,k}) = S_n^{2} \; + nVar(\Gamma_1^{l,k}(x))$$

Were

$$S_n^{2} \; = \sum_{i \neq j} Cov(\Gamma_i^{l,k}, \Gamma_j^{l,k})$$

Firstly, we use the technique developed by Masry ([7],1986) to calculate S_n^{2} .
Thus, we can write

$$S_n^{2} \; \leq \underbrace{\sum_{0 \quad ji \quad jj \quad m_n} \left|Cov(\Gamma_i^{l,k}, \Gamma_j^{l,k})\right|}_{P_{1,n}} + \underbrace{\sum_{m_n+1 \quad ji \quad jj \quad n \quad 1} \left|Cov(\Gamma_i^{l,k}, \Gamma_j^{l,k})\right|}_{P_{2,n}}$$

Where $m_n \longrightarrow \infty$, when $n \longrightarrow \infty$, for all $i \neq j$

$$\begin{aligned}
|Cov(\Gamma_i^{l,k}, \Gamma_j^{l,k})| &= |\mathsf{E}(\Gamma_i^{l,k}\Gamma_j^{l,k})| - \mathsf{E}(\Gamma_i^{l,k})\mathsf{E}(\Gamma_j^{l,k})| \\
&= h_K^{2l}|\mathsf{E}(K_iK_j\beta_i^l\beta_j^l\mathsf{E}(\mathbf{1}_{Y_i\ \ y}^k\mathbf{1}_{Y_j\ \ y}^k/(X_i,X_j))) \\
&- \mathsf{E}(K_i\beta_1^l\mathsf{E}(\mathbf{1}_{Y_1\ \ y}^k/X))\mathsf{E}(K_j\beta_1^l\mathsf{E}(\mathbf{1}_{Y_1\ \ y}^k/X))| \\
&\leq C\mathsf{E}\left|(K(h_K^{\ 1}(x,X_i)K(h_K^{\ 1}(x,X_j))\right| \\
&+ C\mathsf{E}\left|(K(h_K^{\ 1}(x,X_i)\right|\mathsf{E}\left|(K(h_K^{\ 1}(x,X_j)\right| \\
&\leq C(\mathsf{E}(\mathbf{1}_{B(x,h_K)}(X_i)\mathbf{1}_{B(x,h_K)}(X_j)) \\
&+ \mathsf{E}(\mathbf{1}_{B(x,h_K)}(X_i))\mathsf{E}(\mathbf{1}_{B(x,h_K)}(X_j)) \\
&\leq C(\mathsf{P}(X_i,X_j) \in B(x,h_K) \times B(x,h_K)) \\
&+ C\left(\mathsf{P}(X_i \in B(x,h_K)\mathsf{P}(X_j \in B(x,h_K))\right) \\
&\leq C(\phi_x(x)^{a/(a\ \ 1)} + \phi_x(h)^2) \\
&\leq C\phi_x(x)^{a/(a\ \ 1)}
\end{aligned}$$

Secondly, to studied the sum $P_{2,n}$, we use the Davydov-Rio's inequality ([20], 1999), then

$$\forall i \neq j, \quad |Cov(\Gamma_i^{l,k}(x), \Gamma_j^{l,k}(x))| \leq C\alpha(|i-j|)$$

So for n large enough, we get :

$$S_n^2 \ \leq C\left(nm_n\phi_x^{a/(a\ \ 1)}(h) + \frac{nm_n^{\ a+1}}{a-1}\right)$$

The choice of $m_n = (\phi_x(h))^{\ \ 1/(a\ \ 1)}$ permits to get

$$S_n^2 \ = O(n\phi_x(h)) \tag{3.22}$$

Using $(H3)$ and $(H5)$ in order to write

$$Var(\Gamma_1^{l,k}(x)) = \mathsf{E}(\Gamma_1^{l,k}(x))^2 - \mathsf{E}^2(\Gamma_1^{l,k}(x))) \leq C\phi_x(h) \tag{3.23}$$

Finally, (3.22) and (3.23) leads directly to :

$$S_n^2 = o(n\phi_x(h)) \tag{3.24}$$

Once again, similar arguments as those invoked for proving Lemma, which allow us to obtain that : for all ϵ and $r > 1$

$$\begin{aligned}\mathsf{P}\{|T_{k+2} - \mathsf{E}(T_{k+2})| > \epsilon\} &\leq C\left(\frac{n}{r}\left(\frac{r}{n\epsilon\phi_x(h)}\right)^{a+1} + \left(1 + \frac{n^2\epsilon^2\phi_x(h)^2}{rS_n^2}\right)^{\ \ r/2}\right) \\ &= C(R_1 + R_2)\end{aligned}$$

Where $\epsilon = \epsilon_0\sqrt{\dfrac{\log n}{n\phi_x(h)}}$ and $r = c(\log n)^2$. So,

$$R_1 \leq Cn^{(1\ \ a)/2}\phi_x(h)^{\ \ (a+1)/2}(\log n)^{(3a\ \ 1)/2}$$

We deduce from $(H7)$ that

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

Finally, for ϵ_0 large enough of some $\nu > 0$ such that :

$$\sum_n \mathsf{P}\left\{[T_k - \mathsf{E}(T_k)] > \epsilon_0\sqrt{\frac{\log n}{n\phi_x(h)}}\right\} < \infty \tag{3.25}$$

The both hypotheses $(H3)$ and $(H5)$ implies to write that :

$$\mathsf{E}(K_i(x)\beta_i^l(x)\mathbf{1}_{Y_i\ \ y}^k) = O(h_K^l\phi_x(h)), \quad \text{for} \quad k = 0, 1 \quad \text{and} \quad l = 0, 1, 2$$

Which implies that

$$\mathsf{E}(T_k) = O(1), \quad \text{for} \quad k = 2, 3, 4, 5. \tag{3.26}$$

The proof of this last follows exactly along the same line as the proof of (3.24).

$$Cov(T_2, T_3) = O\left(\frac{1}{n\phi_x(h)}\right) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

And

$$Cov(T_4, T_5) = O\left(\frac{1}{n\phi_x(h)}\right) = o\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

By combining (3.25), (3.26) and the both last equations to achieve the proof of the lemma.

Bibliographie

[1] Aneiros-Pérez, G., Cai, R., and Vilar-Fernández, J.M. (2008), 'Functional Methods for Time Series Prediction : A Nonparametric Approach'. Proceedings of IASC2008, Pacifico, Yokohama, Japan, pp. 91-100.

[2] Ali Laksaci and Hachemi Nawal (2012) : Note on the functional linear estimate of conditional cumulative distribution function. Journal of Probability ans Statistical Science ; 10(2), 153-160,

[3] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *of Nonparametric Statistics,* 22, No.5, Pages 617-632.

[4] Baíllo, A., and Grané, A. (2009), 'Functional Local Linear Regression with Functional Time Series Prediction', *Statistics, Probability Letters.*

[5] Boj, E., Delicado, P., and Fortiano, J. (2008), 'Local Linear Functional Based on Weighted Distance-Based Regression', in Functional and Operatorial Statistics, eds. S. Dabo-Niang and F. Ferraty, Heidelberg : Physica-Verlag, pp. 57-64.

[6] Cleveland, William S. (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots". Journal of the American Statistical Association 74 (368) : 829-836.

[7] E. Masry, Recursive probability density estimation for weakly dependant stationary processus. IEEE. Trans. Inform. Theory 32 (1986) 254-267.

[8] Fan, Jianqing and Yao, Qiwei (2003) Nonlinear time series : nonparametric and parametric methods. Springer, New York, USA. ISBN 9780387261423.

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

[10] Ferraty, F ; Vieu, P. (2006). Nonparametric functional data analysis. Springer New-York, in print.

[11] FAN, J., YAO, Q. and TONG, H. (1996) Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. Biometrika **83**, 189.206.

[12] HALL, P., WOLFF, R.C.L. AND YAO, Q. (1999). Methods for estimating a conditional distribution function. *J.* Amer. Statist. Assoc. **94**, 154-163.

[13] Hastie, T. and LOADER, C. (1993). Local regression : automatic kernel carpentry (with discussion). *Statist. Science* **8**, 120-143.

[14] Hansen, B.E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* **24**, 726-748.

[15] Li, Q. and J.S. Racine (2007), Nonparametric Econometrics : Theory and Practice. Princeton University Press. ISBN 0-691-12161-3.

[16] Lun Z. and Linton, O. (2007). Local linear fitting under near epoch dependence. *Econometric Theory* **23**, 37-70.

[17] Masry, E. (1996). Multivariate local polynomial regression for time series : uniform strong consistency and rates. *Journal of Times Series Analysis* **17**, 571-599.

[18] Nadaraya, E.A. (1964) On estimating regression. Theory of Probability and its Applications **10**,186-190.

[19] P.Doukhan, Mixing : Properties and examples. Lect. notes in statistic, 80, Springer-Verlag. Berlin, 1994.

[20] Rio, E. (1999). Théorie asymptotique des processus faiblement d´ependants. Math´ematiques, *Applications*, **31**, Springer-SMAI.

[21] Wand, M. P. and Jones, M. C. (1995). Kernel Smoothing. London : Chapman, Hall.

[22] Watson, G.S. (1964) Smooth regression analysis. Sankhya. **26**, 359-372. "

[23] YU, K. and JONES, M.C. (1998) Local linear quantile regression. J. Amer. Statist. Assoc. **93**, 228.237.

Chapitre 4

Conclusion

Le point de départ de cette thèse a développé le travail qui a été effectué par Barrientos-Marin ([2], 2010). Dans cette thèse, nous avons proposé d'étudier l'estimation semi-paramétrique de la fonction de répartition conditionnelle par la méthode linéaire locale comme outil réducteur de la dimension en estimation fonctionnelle. A partir des travaux que nous avons vu, on peut mentionner les remarques suivantes :

1. Ce modèle est insensible à l'effet de la dimension, par ce qu'il nous permet de transformer le problème de prévision à partir d'une variable fonctionnelle de dimension infinie à une prévision avec variable explicative de dimension un.
2. Les vitesses de convergence obtenues sont insensibles aux corrélations des observations. Cependant, certaines hypothèses supplémentaires ont été introduites afin de prendre en compte la dépendance des observations.
3. le terme biais confirme la supériorité par la méthode du polynôme locaux dans le cas multivarié. En outre, la variance de l'estimateur dépend de la fonction $\phi_x(h)$ qui est étroitement liée au bi-fonctionnel δ, puis à la structure topologique de l'espace fonctionnel (voir. Fan and Gijbels ([4], 1996).
4. Lorsque la densité conditionnelle est dissymétrique ou encore le cas où elle est multimodales, le quantile conditionnel prévoit mieux que la régression.

5. L'idée consiste alors à obtenir des résultats asymptotiques en fonction de la quantité $P(\delta(x, X) < h)$; selon le comportement de cette quantité en fonction de h, on en déduit différentes vitesses de convergence.
   (a) Le terme de biais $O(h^b)$ provient essentiellement de la régularité de $F^x(.)$.
   (b) Alors que le seconde $O(\sqrt{\log n/(n\phi_x(h))})$ dépend de la concentration de la mesure de probabilité sur des petites boules.

Chapitre 5

Perspectives

La richesse de recherche dans ce domaine offre de nombreuses perspectives, tant d'un point de vue théorique que pratique. Pareillement aux développements, ce domaine est d'autant plus motivant qu'aucune étude n'avait jusqu'alors abordé cette problématique. L'objectif est naturellement d'établir des propriétés asymptotiques. Nous consacrons cette dernière partie aux prolongements possibles de notre travail. D'une manière générale, les travaux que nous avons présenté dans cette thèse sont des synthèses qui relèvent pour une bonne part, comme on a pu voir dans la statistique mathématique. Pour autant, les aspects appliqués de la statistique ne sont pas absents : quand il s'agit des problèmes de choix du paramètre de lissage. Les travaux que nous avons réalisé trouvent le plus souvent des motivations et des prolongements dans la pratique. C'est dans cet esprit, c'est-à-dire la mise en perspectives théoriques à travers des applications des méthodes étudiées que nous souhaitons poursuivre notre recherche. On peut mentionner les aspects principaux ;

5.1 Erreur quadratique moyenne

Appelée aussi "risque quadratique" est trés utile pour comparer plusieurs estimateurs, notamment lorsque l'un d'eux est biaisé. Si les deux estimateurs à comparer sont sans biais, l'estimateur le plus efficace est simplement celui qui à la variance la plus petite. On peut effectivement exprimer l'erreur quadratique moyenne en fonction du biais de l'estimateur $\mathsf{E}(\hat{F}^x(y) - F^x(y))$

ainsi que sa variance. $MSE(\hat{F}^x(y)|F^x(y)) = \text{Biais}(\hat{F}^x(y))^2 + \text{Var}(\hat{F}^x(y))$. De cette façon, les procédures de choisir de façon optimale le paramètre de lissage et de construire l'intervalle de confiance peuvent être mises en œuvre. De même que pour la méthode du noyau, le choix du paramètre de lissage dans le cas linéaire local joue un rôle essentiel dans la détermination de la performance des estimateurs surtout la parti biais. Plus précisèment, la dimensionnalité du modèle dans la partie biais, alors que la dimensionnalité de l'espace fonctionnel de la variable explicative a été explicitée dans la partie dispersion.

5.2 Choix du paramètre de lissage :

La question du choix optimal du paramètre de lissage dans le cas des données fonctionnelles est reste ouverte. Ce paramètre est d'une importance capitale. L'optimisation de la partie biais est obtenue par la minimisation de ce paramètre . Par contre, il faut le maximiser pour optimiser la partie de dispersion. En effet, Fan et Gijbels ([4], 1996) ont montré que la largeur de ce paramètre contribue à déterminer la complexité du modèle, en définissant le voisinage local pour lequel l'approximation linéaire du modèle est valide. L'utilisation pratique de cette méthode nécessite des efforts supplémentaires de calcul. Donc, on considère la méthode de la validation croisée pour l'étude de l'optimalité asymptotique.

5.3 la convergence uniforme

La convergence uniforme est une forme plus exigeante que la convergence simple. Cette dernière demande seulement que, pour chaque point x, la suite ait une limite et considérée comme une étape préalable. Le choix de la semi-métrique, reste un problème ouvert pour lequel les résultats uniformes devraient utiles.

5.4 Convergence spatiale

Au futur, on généralisera notre résultats au cadre spatial lorsque les observations sont indépendants et identiquement distribuées comme Laksaci et al.([6],2013). Cependant, Les convergences "spatiales" font référence à la réalisation des variables aléatoires sur un même espace de probabilité et pas seulement à leur lois comme la convergence en loi.

5.5 Normalité asymptotique

On va suit le travail de Ezzahrioui et Ould-Saïd ([3], 2008) qui a été déterminé par la méthode du noyau par une autre approche plus générale d'ordre un, c'est la méthode linéaire local dans le cas fonctionnel. Elle nous permet de déterminer les intervalles de confiance et de faire des testes statistiques.

Chapitre 6

Annexe

6.1 Notations et définitions

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

Definition 6.1.1. *Soit $\{\Delta_i\}_{i\in\mathbb{Z}}$ une famille de variables aléatoires définie sur $(\Omega, \mathcal{A}, \mathsf{P})$ à valeurs dans un espace probabilisable $(\mathbb{E}, \xi)$. On dit que la famille $\{\Delta_i, i \in \mathcal{Z}\}$ est α-mélangent si la suite*

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

tend vers 0 quand n tend vers l'infinie. La suite $\alpha(n)$ est appelée le coefficient de mélange fort.

Definition 6.1.2. *On dit que qu'une famille $\{\Delta_i, i \in Z\}$ de variables aléatoires à valeurs dans un même espace probabilisable $(\mathbb{E}, \xi)$ est algébriquement α-mélangeant, s'il existe deux constantes $c\in\mathbb{R}^+$ et $a\in\mathbb{R}^+$ telles que le coefficient de mélange vérifié*

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

Definition 6.1.3. *On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque complètement vers X si*

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

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

Definition 6.1.4. *On dit que la suite $X_n = O(Y_n)$ en p.co. s'il existe un $\epsilon > 0$, verifiant :*

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

Definition 6.1.5. *Soient X_n, Y_n deux suites des variables aléatoires.*
La suite $\left(\dfrac{X_n}{Y_n}\right)_{n\in\mathbb{N}} \longrightarrow 0$ en p.co si

$$X_n \longrightarrow 0 \quad en \quad p.co.$$

et

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

Definition 6.1.6. *Soient X et Y deux variables aléatoires définies sur l'espace $(\Omega, \mathcal{A}, P)$, on appelle fonction de répartition conditionnelle de Y sachant $X = x$ la fonction de répartition définie par*

$$F^x(y) = P(Y \leq y/X = x).$$

6.2 Outils

Lemma 6.2.1. *([5], 2004) Soit $Z_1, \ldots, Z_n$ des variables aléatoires réelles centrées, indépendantes et identiquement distribuées,*
i) Si $\forall m \geq 2, \quad \exists C_m > 0, \quad \mathsf{E}|Z_1^m| \leq C_m a^{2(m-1)}$, on a

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

ii) Supposons que les variables dépends de n ($Z_i = Z_{i,n}$).
Si $\forall m \geq 2, \quad \exists C_m > 0, \quad \mathsf{E}|Z_1^m| \leq C_m a_n^{2(m-1)}$ et si $u_n = n^{-1} a_n^2 \log n$ vérifié $\lim_{n\to\infty} u_n = 0$, on a :

$$\frac{1}{n}\sum_{i=1}^{n} Z_i = O_{a.co}(\sqrt{u_n})$$

Cette inégalité a été donnée par Bernstein inequality (1937, [8]). Les lemmes suivants donnent les deux versions de l'inégalité de Fuk Nagaev.

Lemma 6.2.2. *([5],2004) Soit* $\{\Delta_i, i \in \mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeant, de coefficient de mélange* $\alpha(n)$ *vérifiant :*

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

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

$$\mathsf{P}\left[\left|\sum_{k=1}^{n} \Delta_k\right| > 4\varepsilon\right] \leq 4\left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{\frac{-r}{2}} + 2ncr^{\ 1}\left(\frac{2r}{\varepsilon}\right)^{a+1}. \tag{6.1}$$

Lemma 6.2.3. *([1], 2005) Soit* $\{\Delta_i, i \in \mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeant, de coefficient de mélange* $\alpha(n)$ *vérifiant :*

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

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

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

ou

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

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

Lemma 6.2.4. *([5], 2004) Soit* $\{\Delta_i, i \in \mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeant, de coefficient de mélange* $\alpha(n)$ *vérifiant :*

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

et si $\|\Delta_i\|_{1} < \infty, \forall i$*, alors, pour tout* $\varepsilon > 0$ *et* $r > 0$

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

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

Lemma 6.2.5. *([5], 2004), ([1], 2005) Soit*$\{\Delta_i, i \in \mathbb{N}\}$ *une suite des variables aléatoires réelle* α*-mélangeante, de coefficient des mélange* $\alpha(n)$*, telle que* $\|\Delta_i\|_1 < \infty, \forall i$. *On a pour tout* $i \neq j$ *:*

$$|cov(\Delta_i, \Delta_j)| \leq 4|\Delta_i|_1 \, |\Delta_j|_1 \, \alpha_{ji \quad jj}.$$

Bibliographie

[1] Ait Saidi, A.,Ferraty, F., Kassa, R., *Single functional index model for time series (2005.)* Rev. Roumaine Math.Pures Appl., in print.

[2] Barrientos-Marin-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *of Nonparametric Statistics,* 22, No.5, Pages 617-632.

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

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

[5] Ferraty, F et Vieu, Ph. Modèle de régression pour variables aléatoires uni, multi et ∞-dimensionnées (2004). Cours de DEA.

[6] Laksaci. A, Rachdi. M, Rahmani. S, (2013).s Spaciale modelization : Local linear estimation of the conditional distribution for functional data.

[7] Loève, M. 1963. Probability Theory, Third Edition, Van Nostrand, Princeton.

[8] Uspensky, J. V. (1937). *Introduction to Mathematical Probability McGraw-Hill, New York.*