

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DJILLALI LIABÈS DE SIDI BEL ABBÈS
FACULTÉ DES SCIENCES EXACTES
DÉPARTEMENT DE PROBABILITÉS ET STATISTIQUES
LABORATOIRE DE STATISTIQUE ET PROCESSUS STOCHASTIQUES (LSPS)

N° d'ordre : / 2021



THESE

Pour obtenir le Diplôme de Doctorat en Sciences délivré par

l'Université de Sidi Bel-Abbès

Spécialité : Mathématiques

Option : Probabilités et Statistique

Intitulée

CONTRIBUTION À L'ESTIMATION NON PARAMÉTRIQUE DES EXPECTILES EN STATISTIQUE FONCTIONNELLE

Présentée et soutenue publiquement le 10/02/2021, par :

M. MOHAMMEDI Mustapha

Devant le jury composé de :

Mr. **ATTOUCH Mohamed** Professeur à l'Univ. Djillali Liabès de SBA **Président**

Mr. **LAKSACI Ali** Professeur à l'Univ. Djillali Liabès de SBA **Directeur de Thèse**

Mr. **BOUZEBDA Salim** Professeur à l'Univ. de Tech. de Compiègne (France) **Co-directeur de Thèse**

Mr. **GUENDOZI Toufik** Professeur à l'Univ. Dr. Moulay Tahar de Saïda **Examineur**

Mr. **MECHAB Boubaker** Professeur à l'Univ. Djillali Liabès de SBA **Examineur**

Mme. **RAHMANI Saâdia** Professeur à l'Univ. Dr. Moulay Tahar de Saïda **Examinatrice**

ANNÉE UNIVERSITAIRE 2020-2021

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude pour mon directeur de thèse Monsieur **Laksaci Ali**, Professeur à l'Université Roi Khalid de Abha (Arabie Saoudite), qui a dirigé l'ensemble de ce travail, je le remercie pour m'avoir ouvert les portes de la recherche en Statistique fonctionnelle. Je suis particulièrement reconnaissant de la confiance qu'il m'a su m'accorder durant cette thèse ainsi pour la manière avec laquelle il a mené ce projet de thèse à son terme, à savoir sa disponibilité, ses remarques toujours pertinentes ainsi que sa gentillesse omniprésente.

J'exprime également toute ma profonde gratitude à Monsieur **Bouzebda Salim**, Professeur à l'Université de Technologie de Compiègne (France), qui me fait l'honneur de bien vouloir co-diriger cette thèse, et pour avoir suivi mon travail avec ses précieux conseils et me guider avec un enthousiasme constant et communicatif, et m'encourager lors de mon séjour au sein de Laboratoire de Mathématiques Appliquées à Compiègne (**LMAC**).

Leurs grandes qualités scientifiques et humaines ont été indispensables à l'élaboration de cette thèse. Pour tout cela, mes directeurs de thèses, je ne les remercierai jamais assez.

J'adresse mes sincères remerciements à Monsieur **Attouch Mohammed**, Professeur à l'Université Djillali Liabes de Sidi Bel Abbès, d'avoir accepté de présider le jury de cette thèse.

Je remercie chaleureusement Messieurs **Guendouzi Toufik**, Professeur à l'Université Moulay Tahar de Saïda, **Mechab Boubaker**, Professeur à l'Université Djillali Liabes de Sidi Bel Abbès, et Madame **Rahmani Saâdia**, Professeur à l'Université Moulay Tahar de Saïda, d'avoir accepté d'être les examinateurs de cette thèse. Je leur suis très reconnaissant du temps qu'ils ont consacré à l'évaluation de ce travail.

Puis je suis très reconnaissant envers mes professeurs de Saïda et Sidi Bel Abbès qui ont su me guider dans mes choix. Tout d'abord Monsieur **Kandouci Abdeldjebbar**, Professeur à l'Université Moulay Tahar de Saïda, lors des cours de la théorie de mesures pour le DES, les séminaires hebdomadaires du LMSSA. Monsieur **Tebboune Feth Allah**, Professeur à l'Université Moulay Tahar de Saïda, lors des cours d'Analyse des Données et pour l'encadrement de mon Magister. Je pense ici au professeur **Benamar Chouaf** un homme exceptionnel, nous prions DIEU Le Tout Puissant de lui ouvrir et de l'accueillir dans son vaste Paradis.

Je tiens à remercier de manière plus générale tous les membres de Laboratoire de Statistique et Processus Stochastiques (**LSPS**) de l'Université de Sidi Bel Abbès, ainsi qu'à son directeur Monsieur **Gheriballah Abdelkader**, Professeur à l'Université Djillali Liabes de Sidi Bel Abbès, pour tout le soutien moral et logistique.

Je souhaite aussi remercier Monsieur **Bouagada Djillali**, Professeur à l'Université Abdelhamid Ibn Badis de Mostaganem, pour l'intérêt qu'il a porté à mon travail, ses commentaires constructifs ainsi que sa disponibilité pour relire et améliorer ce manuscrit.

Que tous reçoivent ici l'expression de ma profonde gratitude.

Mustapha Mohammedi

Dédicace

*Cette thèse est dédiée à ma grand-mère,
toute ma famille et mes meilleurs amis...*

Table des matières

Remerciements	2
Dédicace	3
Abréviations et Symboles	8
List of Works	10
Résumé en Arabe	11
Résumé en Anglais	12
Résumé en Français	14
1 État de l’art et contribution de la thèse.	16
1.1 Introduction	16
1.2 Description et contribution de la thèse	16
Bibliographie	23
2 Introduction générale.	24
2.1 Données fonctionnelles et Estimation non paramétrique	24
2.1.1 Domaines d’applications des données fonctionnelles	29
2.1.2 Probabilités de petites boules et semi-métriques	29
2.2 Modèles conditionnels dans l’estimation non paramétrique	32

2.2.1	Estimation de la régression via la méthode du noyau	33
2.2.2	Modèle à indice fonctionnel simple (MIFS)	34
2.2.3	Quantiles et les expectiles conditionnels	36
2.2.4	Estimation de la régression via la méthode k -NN	42
Bibliographie		45
3	Modèle à indice fonctionnel simple (MIFS) pour les données de séries chronologiques quasi-associées.	55
3.1	Introduction	56
3.2	Model and estimator	58
3.3	Assumptions and notation	59
3.4	Main result	61
3.4.1	The almost consistency	61
3.4.2	The asymptotic normality	63
3.4.3	Application to the confidence intervals	64
3.5	Proofs of the intermediate results	65
References		81
4	La méthode des k plus proches voisins dans le modèle à indice fonctionnel simple pour les données de séries chronologiques quasi-associées.	84
4.1	Introduction	85
4.2	Model and estimator	87
4.3	Main result	88
4.3.1	Consistency of NW method	88
4.3.2	Consistency of k -NN method	90
4.4	Proofs of the intermediate results	90
References		102

5	Les propriétés asymptotiques de l'estimateur à noyau des expectiles conditionnels pour les données fonctionnelles.	105
5.1	Introduction	106
5.2	Model and estimators	109
5.2.1	The Conditional expectile	109
5.2.2	Nonparametric estimation of the conditional expectile	110
5.2.3	Hypotheses and notation	111
5.3	The main results	112
5.3.1	Independent case	113
5.3.2	Dependent case	114
5.3.3	Application to the confidence intervals	115
5.4	On the potential impact of the functional expectile regression in NFDA	117
5.4.1	Supervised classification and prediction : expectile regression versus classical regression	117
5.4.2	Financial risk analysis : conditional expectile versus conditional quantile	118
5.4.3	Simulation study	118
5.4.4	Real data example	121
5.5	Conclusion and perspectives	125
5.6	The demonstration of asymptotic results	125
5.7	Appendix	140
5.7.1	Demonstration of Remark 1 : the function $G(\cdot; x)$ is increasing.	140
5.7.2	Bernstein's inequality.	141
5.7.3	Lyapunov CLT.	141
	References	142
6	Conclusion et Perspectives.	146
6.1	Conclusion	146
6.2	Perspectives	147

Bibliographie

149

Abréviations et Symboles

Nous adopterons les notations ci-dessous dans les différents chapitres de cette thèse.

Abréviation

<i>NFDA</i>	Nonparametric Functional Data Analysis,
<i>ALS</i>	Asymmetric Least Squares,
<i>MCA</i>	Moindres Carrés Asymétrique,
<i>SIM</i>	Single Index Model,
<i>MIS</i>	Modèle à Indice Simple,
<i>SFIM</i>	Single Functional Index Model,
<i>MIFS</i>	Modèle à Indice Fonctionnel Simple,
<i>i.i.d.</i>	indépendantes et identiquement distribuées,
<i>ADF</i>	Analyse des Données Fonctionnelles,
<i>AF</i>	Analyse factorielle,
<i>ACP</i>	Analyse en Composantes Principales,
<i>ACPF</i>	Analyse en Composantes Principales fonctionnelles,
<i>v.a.</i>	variable aléatoire,
<i>NW</i>	Nadaraya- Watson,
<i>k – NN</i>	k plus proches voisins ou k Nearest Neighbors,
<i>VaR</i>	Valeur at Risk,
<i>ES</i>	Expected Shortfall,
<i>TVaR</i>	Tail Value at Risk,
<i>CVaR</i>	Conditional VaR,
<i>CES</i>	Conditional ES,
<i>CTE</i>	Conditional Tail Expectation,
<i>RE</i>	Régression expectile,
<i>RQ</i>	Régression quantile,
<i>p.co.</i>	Presque Complète,
<i>a.co.</i>	Almost Complete,
<i>a.s.</i>	Almost sure,
<i>cv.P.</i>	Convergence en Probabilité,
<i>CLT</i>	Central Limit Theorem.

Symboles

X	Variable aléatoire fonctionnelle,
Y	Variable aléatoire réelle,
H	Espace de Hilbert,
L^p	Les espaces de fonctions de puissance p -ième intégrable, avec $1 \leq p < \infty$,
\mathbb{N}	Ensemble des entiers naturels,
\mathbb{R}	Ensemble des nombres réels,
$\mathbb{B}_{\mathbb{R}}$	La tribu borélienne,
$P(A)$	Probabilité d'un évènement A ,
$F_Y^x(\cdot)$	Fonction de distribution conditionnelle de Y sachant $X=x$,
$N(0, 1)$	La distribution de la loi Normale de moyenne 0 et d'écart-type 1,
$\mathbf{1}_A$	Fonction indicatrice qui vaut 1 sur l'ensemble A et 0 ailleurs,
A^c	Le complémentaire d'un évènement A ,
$B(x, h)$	La boule de centre x et de rayon h ,
$\langle \cdot, \cdot \rangle$	Produit scalaire définie dans H ,
$d(\cdot, \cdot)$	Semi-métrie,
$(x)_+ = \max(x, 0)$	La partie positive de x ,
$a \vee b$ resp. $a \wedge b$	désignent le maximum resp. minimum des nombres $a, b \in \mathbb{R}$,
$E(X)$	Espérance mathématique de la variable X ,
$Var(X)$	Variance de la variable X ,
$Cov(X, Y)$	Covariance des variables X et Y ,
$u \otimes v$	Produit tensoriel,
$\min_{x:\dots}(A)$	Minimum de la quantité A par rapport à x qui satisfait les contraintes ...,
$\max_{x:\dots}(A)$	Maximum de la quantité A par rapport à x qui satisfait les contraintes ...,
$(1) \xrightarrow{M} (2)$	La quantité (1) converge en Mode M vers la quantité (2),
$:=$	La définition d'une quantité,
$a_n = o(b_n), n \rightarrow \infty$	$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 0$ avec a_n et b_n deux suites réelles,
$a_n = O(b_n), n \rightarrow \infty$	$\limsup_{n \rightarrow \infty} \left \frac{a_n}{b_n} \right < \infty$ avec a_n et b_n deux suites réelles,
$a_n \sim b_n, n \rightarrow \infty$	$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = c$ avec a_n et b_n deux suites réelles et c une constante,
$X \sim \mathbb{P}$	La v.a. X suit une probabilité \mathbb{P} .
$\lim_{n \rightarrow \infty} X_n = X$ P.	$\Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{P}(X_n - X > \epsilon) = 0$,
$\lim_{n \rightarrow \infty} X_n = X$ p.s.	$\Leftrightarrow \mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1$,
$\lim_{n \rightarrow \infty} X_n = X$ p.co.	$\Leftrightarrow \forall \epsilon > 0, \sum_{n \in \mathbb{N}} \mathbb{P}(X_n - X > \epsilon) < \infty$,
$X_n - X = O_{p.co.}(u_n)$	$\Leftrightarrow \exists \epsilon_0 > 0, \sum_{n \in \mathbb{N}} \mathbb{P}(X_n - X > \epsilon_0 u_n) < \infty$,

List of Works

Publications in international peer-reviewed journals :

- M. Mohammedi, S. Bouzebda and A. Laksaci. On the nonparametric estimation of the functional expectile regression. Published in *C. R., Math., Acad. Sci. Paris.*, Tome 358 (2020) no. 3, pp. 267-272. DOI : <https://doi.org/10.5802/crmath.27>.
- M. Mohammedi, S. Bouzebda and A. Laksaci. The Consistency and Asymptotic Normality of the Kernel type Expectile Regression Estimator for Functional Data. Published in *Journal of Multivariate Analysis (JMVA)*. Volume 181, January 2021, 104673. DOI : <https://doi.org/10.1016/j.jmva.2020.104673>.
- S. Bouzebda, A. Laksaci and M. Mohammedi. Single index regression model for functional quasi-associated times series data. [Submitted for publication : minor revision](#).
- S. Bouzebda, A. Laksaci and M. Mohammedi. The k -Nearest Neighbors method in single index regression model for functional quasi-associated times series data. [Submitted for publication](#).

نبذة مختصرة عن أطروحة الدكتوراه

عنوان الأطروحة: المساهمة في التقدير غير المعياري للتوقعات في الإحصائيات الوظيفية. الكاتب: محمدي مصطفى

في هذه الأونة الأخيرة، شعبة الإحصاء المتخصصة في دراسة المتغيرات الوظيفية شهدت ازدهارا حقيقيا ليس فقط من حيث التطور النظري وإنما عرفت كذلك تنوع و تقدم في المجالات التطبيقية. نحن مهتمون بشكل خاص في هذه الأطروحة بنماذج الانحدار التي يكون فيها المتغير العشوائي "ع" للاستجابة عددا حقيقيا، بينما المتغير التوضيحي "س" يكون وظيفي، أي لديه قيم في فضاء ذو بعد غير منتهي. إن شكل الانحدار هو الذي تم دراسته أكثر في علم الإحصاء، ونموذج الانحدار هو تلك العلاقة بين المتغيرين "س" و "ع"، و التي يمكن أن نلخصه من خلال تقدير تأثير الانحدارات على المتوسط الشرطي لمتغير الاستجابة. علاوة على ذلك، فإن السرعة المثلى للتقارب بين المقدرات غير المعيارية لوظيفة الانحدار تتناقص بسرعة كبيرة عندما يزداد بُعد المتغير التوضيحي، هذا يرجع إلى ظاهرة معروفة يطلق عليها عادة آفة البعد. لا تتعلق هذه الآفة بالمقدر الغير المعياري المستخدم بل بالأحرى بطبيعة النموذج المدروس.

في الجزء الأول من هذه الأطروحة، ومن خلال الملاحظات السابقة الذكر، نعتبر النموذج وظيفيا ذو مؤشر بسيط، ونفترض أن العينة تتكون من متغيرات عشوائية شبه مرتبطة، هذا الارتباط المستخدم كحالة خاصة من الارتباط الضعيف مما يجعل من الممكن توفير إطار ملائم للتطبيقات. ندرس في هذا الجزء من الأطروحة، التقارب الشبه الكامل للمقدر وسرعته الذي تم تشكيله اعتمادا على طريقة النواة. ويتم كذلك، الحصول على قانون التوزيع الطبيعي المقارب لهذا المقدر في ظل الظروف العامة، حيث نناقش كيفية تطبيق النتائج المتحصل عليها لإنشاء مجالات الثقة.

نقدم، في الجزء الثاني من الأطروحة، طريقة بديلة لطريقة النواة وهي طريقة أقرب الجيران، معتمدا فيها على النموذج الوظيفي الموصوف في الجزء الأول و من خلال عامل التنعيم العشوائي و الذي يعتمد على العينة المدروسة. نبني المقدر حسب طريقة أقرب الجيران و ندرس سرعة التقارب شبه الكامل لهذا المقدر. ترتبط هذه السرعة بعدد الجيران الأقرب و المعبر عنه، حسب هذه الطريقة، بعامل التنعيم العشوائي و يأخذ قيمه في فضاء منفصل و ذا بعد منتهي.

أما الجزء الثالث من هذه الأطروحة، و بالرغم من أن الاستدلال على المتوسط الشرطي مقبول بشكل عام، إلا أن الفائدة تتعلق بتقدير نيول توزيع الاستجابة بشكل مشروط للانحدارات. في هذا السياق، يكون الانحدار الكلاسيكي غير فعال ومن الضروري تجاوز تقدير المتوسط الشرطي. يقدم الإحصاء الحديث مناهج للرد على هذا النوع من المشاكل، لا سيما مع انحدار الغير متماثل للمربعات الصغرى و الموزونة أو الانحدار المتوقع. إن هذا النهج الجديد يسمح من ناحية أولى لفحص و تحليل تأثير الانحدارات على التوزيع الشرطي لمتغير الاستجابة، و بالتالي الكشف عن مجموعة متنوعة من أشكال عدم التجانس، ومن ناحية أخرى، فإن الانحدار المتوقع يعتبر بديلا جيدا للانحدار التناظري الكمي وفقا للأعمال الأخيرة المنشورة في المجالات العلمية.

نفترض، في هذا الجزء الأخير من الأطروحة، أن لدينا أزواج عشوائية مستقلة و موزعة بشكل مماثل. نبني حسب طريقة النواة مقدر للمتوقع الشرطي و نؤسس التقارب شبه الكامل و سرعته، بالإضافة إلى عرض القانون التوزيع الطبيعي المقارب للمقدر في ظل الظروف العادية للتقدير الوظيفي الغير معياري و نوضح كيفية تطبيق نتائجنا لبناء مجالات الثقة. ناقشنا التأثير المحتمل للانحدار الوظيفي للتوقعات في تحليل البيانات الوظيفية الغير معيارية مع تطبيق حقيقي و محاكاة لانحدار التوقعات مقارنة بنماذجها التنافسية (الكمية و الانحدار). في الأخير، ندرس التنفيذ السهل لهذا النموذج في الممارسة العلمية في مجال تحليل المخاطر و المقارنة بين الانحدارات الكمية و المتوقعة.

الكلمات المفتاحية:

الانحدار، التقدير غير المعياري، البيانات الوظيفية، المتغيرات العشوائية شبه المرتبطة، التقارب شبه الكامل، قانون التوزيع الطبيعي المقارب، النموذج الوظيفي ذو مؤشر بسيط، طريقة أقرب الجيران، الانحدار المتوقع، الانحدار الكمي.

ABSTRACT

Thesis Title : Contribution to the nonparametric estimation of the expectiles in functional statistics.

Author : Mustapha Mohammedi

The branch of statistics devoted to the study of functional variables has undergone remarkable development in recent decades, both in terms of theoretical evolution and of diversification of fields of application. We are more particularly interested in this thesis in one form of regression which has been the most studied in the literature, it is the regression model where the random variable response Y is real whereas the explanatory variable X is functional with values in a space of infinite dimension. It models the relationship between X and Y , and is summarized by the estimation of the influence of the regressors on the conditional mean of the response variable.

We observe, in regression models, that the optimal rate of convergence of nonparametric estimators deteriorates very quickly when the dimension of the explanatory variable increases, this is due to a well-known phenomenon which is commonly called the curse of the dimension. The latter is not related to the nonparametric estimator used but rather to the nature of the model studied.

Firstly, from the above observation, we consider a single functional index model (SFIM), assume that the sample is composed of quasi associated random variables and construct the estimator by the kernel method for the regression function. In this part of the thesis, we study the almost complete convergence of the estimator by specifying its rate. In addition, the asymptotic normality of this estimator is obtained under standard conditions, and we discuss how to apply our result to construct the confidence intervals. The concept of association used is considered as a particular case of weak dependence which makes it possible to provide an adequate framework for applications.

Secondly, we present another alternative method to the kernel method. It is the method of k nearest neighbors (or k -NN). We consider the same functional model described in the first part, and we construct the estimator according to a random smoothing parameter which data driven. We study the almost complete rate of convergence of this estimator. This rate is related to the number of neighbors' k and, in turn, this number is considered as a smoothing parameter for the estimators of k -NN which takes its values in a discrete space.

Thirdly, the inference on the conditional mean is generally acceptable, it happens that the interest relates to the estimation of the tails of the distribution of the response variable conditionally to the regressors. In this context, classical regression is ineffective and it is necessary to go beyond the estimation of the conditional mean. Modern literature offers approaches to respond to this kind of problem, in particular with asymmetric weighted least squares regression or expectile regression (ER). This new approach allows, on the one hand, to examine and analyze the influence of the regressors on the conditional distribution of the response variable, thus revealing a variety of forms of heterogeneity. On the other hand, the ER is very simple to implement compared to its analog, quantile regression (QR), and even the ER is considered a good alternative to the QR according to recent works published in the field.

We consider, in this thesis part, a random couple (X, Y) with values in $\mathcal{F} \times \mathbb{R}$ where (\mathcal{F}, d) is a semimetric space of infinite dimension and d a semi-metric. We assume that we have n random pairs $(X_i, Y_i)_{i=1 \dots n}$ independent and identically distributed as (X, Y) , construct the kernel estimator of the conditional expectile and establish the almost complete convergence and its rate of convergence. In addition, we show the

asymptotic normality of the proposed estimator under standard conditions of the functional nonparametric estimation and explain how to apply our results to construct the confidence intervals. In the latter, we also discussed the potential impact of functional expectile regression in the NFDA. Finally, through a real and simulated application of the expectile regression compared to its competitive models (quantile regression), we will examine the implementation of this model in practice. In particular, we focus on the area of risk analysis and compare between quantile and expectile regressions.

Keywords : Almost Complete Convergence, Asymptotic Normality, Expectiles Regression, Functional Data, k Nearest Neighbors, Kernel Estimate of Regression, Quantiles, Quasi-Associated Random Variables, Nonparametric Estimation, Single Functional Index Model, Weak dependence.

RÉSUMÉ

Titre de la Thèse : Contribution à l'estimation non paramétrique des expectiles en statistique fonctionnelle.
L'auteur : Mustapha Mohammedi

La branche de la statistique consacrée à l'étude de variables fonctionnelles a connu une évolution remarquable, au cours des dernières décennies, tant en terme de développements théoriques que de diversification des domaines d'application. Nous nous intéressons plus particulièrement dans cette thèse à une forme de régression qui a été la plus étudiée dans la littérature, c'est le modèle de régression où la variable aléatoire réponse Y est réelle tandis que la variable explicative X est fonctionnelle, c'est à dire, à valeurs dans un espace de dimension infinie. La régression modélise la relation entre X et Y , et se résume par l'estimation de l'influence des régresseurs sur la moyenne conditionnelle de la variable réponse.

Nous observons, dans les modèles de régression, que la vitesse optimale de convergence des estimateurs non-paramétriques se détériore très rapidement lorsque la dimension de la variable explicative augmente, cela est dû à un phénomène bien connu que l'on appelle couramment fléau de la dimension. Ce dernier n'est pas lié à l'estimateur non-paramétrique utilisé mais plutôt à la nature du modèle étudié.

Dans un premier temps, et à partir de ce constat, nous considérons un *modèle à indice fonctionnel simple* (MIFS) et on suppose que l'échantillon est composé de *variables aléatoires quasi-associées* et nous construisons l'estimateur par la méthode du noyau pour la fonction de régression. Nous étudions, dans cette partie de thèse, la convergence presque complète de l'estimateur en précisant sa vitesse. De plus, la normalité asymptotique de cet estimateur est obtenue sous des conditions standards, ainsi que nous discutons comment appliquer notre résultat pour construire les intervalles de confiance. Le concept d'association utilisé est considéré comme un cas particulier de la condition de dépendance faible qui permet de donner un cadre adéquat pour les applications.

Dans un second temps, on présente une autre méthode alternative à la méthode du noyau c'est la méthode des *k plus proches voisins* (ou k -NN). Nous considérons le même modèle fonctionnel décrit dans la première partie, et nous construisons l'estimateur en fonction d'un paramètre de lissage aléatoire et qui dépend de l'échantillon. Nous étudions la vitesse de convergence presque complète de cet estimateur. Cette vitesse est liée au nombre de voisins k et, à son tour, ce nombre est considéré comme un paramètre de lissage pour les estimateurs des k -NN qui prend ses valeurs dans un espace discret.

Dans un troisième temps, l'inférence sur la moyenne conditionnelle est généralement acceptable, il arrive que l'intérêt porte sur l'estimation des queues de la distribution de la variable réponse conditionnellement aux régresseurs. Dans ce contexte, la *régression classique* est inefficace et il faut aller au-delà de l'estimation de la moyenne conditionnelle. La littérature moderne offre des approches pour répondre à ce genre de problématique, notamment avec la *régression asymétrique des moindres carrés pondérés ou régression expectile* (RE). Cette nouvelle approche permet d'une part d'examiner et d'analyser l'influence des régresseurs sur la distribution conditionnelle de la variable réponse, révélant ainsi une variété de formes d'hétérogénéité, et d'autre part la RE est très simple à mettre en œuvre comparativement à son analogue, la *régression quantile* (RQ), et même l'expectile est considéré comme un bon alternatif au quantile selon des travaux récents publiés dans le domaine.

On considère, dans cette partie de thèse, un couple aléatoire (X, Y) à valeurs dans $\mathcal{F} \times \mathbb{R}$ où (\mathcal{F}, d) un espace semi-métrique de dimension infinie et d une semi-métrique. Nous supposons qu'on a n paires aléatoire $(X_i, Y_i)_{i=1, \dots, n}$ indépendantes et identiquement distribuées comme (X, Y) , et nous construisons l'estimateur à noyau de l'expectile conditionnel en établissant la convergence presque complète de l'estimateur et sa

vitesse de convergence. De plus, nous montrons la normalité asymptotique de l'estimateur proposé sous des conditions standards de l'estimation non-paramétrique fonctionnelle et nous expliquons comment appliquer nos résultats pour construire les intervalles de confiance. Dans ce dernier, nous avons discuté l'impact potentiel de la régression fonctionnelle des expectiles dans la NFDA. Finalement, Une application réelle et simulée de la régression expectile par rapport à ses modèles compétitifs (quantile et régression). Nous examinerons ainsi, dans cette dernière partie, la mise en œuvre de ce modèle dans la pratique. En particulier, nous nous concentrons sur le domaine de l'analyse des risques et nous comparons entre les régressions quantiles et expectiles.

Mots clés : Convergence Presque Complète, Donnée Fonctionnelle, Estimation Non-paramétrique, Estimation à noyau de la régression, Quantile, Méthode des k plus proches voisins, Modèle à Indice Fonctionnel Simple, Normalité Asymptotique, Régression Expectile, Variables Aléatoires Quasi-Associées.

Chapitre 1

État de l'art et contribution de la thèse.

1.1 Introduction

On suppose pour commencer que l'on s'intéresse au cas de la régression d'une variable aléatoire réelle (v. a. réponse Y) sur une variable aléatoire fonctionnelle (v. a. explicative X), c'est la forme de la régression qui a été la plus étudiée dans la littérature. Cette régression modélise la relation entre les régresseurs et la variable réponse. La modélisation se résume par l'estimation de l'influence de la variable explicative sur la moyenne conditionnelle de la variable réponse. Dans cette thèse, Nous cherchons d'une part à généraliser la méthode de régression classique en développant des alternatives de cette dernière lorsque les données sont générées à partir d'un modèle de régression à indice fonctionnel simple, et d'autre part de proposer la régression expectile (RE) comme une alternative parfaitement raisonnable aux mesures de risques telles que la Valeur at Risk (VaR), Expected Shortfall (ES) et au régression quantile (RQ). Cette RE, non seulement est une bonne mesure de risque mais permet, en plus, d'examiner et d'analyser l'influence des régresseurs sur la distribution conditionnelle de la variable réponse.

1.2 Description et contribution de la thèse

Nous présentons dans cette thèse, brièvement les principaux résultats établis dans chaque chapitre tout en exposant la problématique de recherche traitée. Nous présentons également les questionnements à l'origine du sujet de recherche.

Le chapitre 2, est un chapitre introductif, qui présente une étude bibliographique des problèmes liés à l'analyse statistique des données fonctionnelles ainsi qu'à l'estimation non paramétrique, que ce soit dans le cadre de la dimension finie ou infinie, des paramètres conditionnels. Ensuite, dans le troisième chapitre, nous abordons le modèle à indice fonctionnel simple (MIFS) pour des données fonctionnelles dépendantes. Plus précisément, nous considérons un couple de variables aléatoires (X, Y) quasi-associées à valeurs dans un espace mesurable $(H \times \mathbb{R}, \mathcal{A}_d \otimes \mathbb{B}_{\mathbb{R}})$ où H est un espace de Hilbert réel séparable, (\mathcal{A}, d) est un espace semi-métrique et \mathcal{A}_d est la σ -algèbre engendrée par la topologie de H qui est définie par la semi-métrique $d(\cdot, \cdot)$ et $\mathbb{B}_{\mathbb{R}}$ est la tribu borélienne dans \mathbb{R} . Ce chapitre est présenté comme suit :

Chapitre 3 : Modèle à indice fonctionnel simple (MIFS) pour les données de séries chronologiques quasi-associées.

On dispose de n couples $(X_i, Y_i)_{i=1, \dots, n}$ quasi-associés et identiquement distribués selon la loi de (X, Y) .

Nous considérons la semi-métrique $d_\theta(\cdot, \cdot)$ associée à l'indice simple $\theta \in H$ définie, pour tout $u, v \in H$, par

$$d_\theta(u, v) := |\langle \theta, u - v \rangle|.$$

Soit θ fixé dans $\Theta \subset H$, où les observations $(X_i, Y_i)_{i=1, \dots, n}$ sont liées par la relation suivante :

$$Y_i = r(\langle \theta, X_i \rangle) + \varepsilon_i, \quad \forall i = 1, \dots, n, \quad (1.1)$$

où $r(\cdot)$ est la fonction réelle et pour $i = 1, \dots, n$, ε_i est la variable aléatoire réelle telle que $\mathbb{E}(\varepsilon_i | X_i) = 0$.

L'estimateur à noyau $\hat{r}_\theta(\cdot, h_n)$ de l'opérateur de la régression

$$r_\theta(\cdot) = r(\langle \theta, \cdot \rangle) = \mathbb{E}(Y | \langle \theta, X_i = \cdot \rangle)$$

est défini par

$$\hat{r}_\theta(x, h_n) = \frac{\sum_{i=1}^n Y_i K_i(x)}{\sum_{i=1}^n K_i(x)}, \quad \text{pour tout } x \in H, \quad (1.2)$$

où $K_i(x) := K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right)$ est un noyau et h_n est un paramètre (de lissage) de nombres positifs qui décroît vers 0 quand $n \rightarrow \infty$.

Dans ce chapitre, notre but est d'étudier l'estimation non paramétrique de la fonction de régression dans un modèle à indice fonctionnel simple sous la condition de dépendance de quasi-association. Le principal résultat de ce chapitre est l'obtention des propriétés asymptotiques de l'estimateur et la vitesse de convergence presque complète (voir le Théorème 1.1). De plus, la normalité asymptotique de cet estimateur est obtenue sous des conditions standards de l'estimation nonparamétrique fonctionnelle (voir le Théorème 1.2 et le Corollaire 1.1) et nous discutons comment appliquer notre résultat pour construire les intervalles de confiance. La démonstration et les détails de conditions imposées font l'objet d'un article récemment soumis et qui est présenté dans le chapitre 3 de cette thèse.

Théorème 1.1 *On démontre sous certaines hypothèses, quand $n \rightarrow \infty$, que*

$$\hat{r}_\theta(x, h_n) - r_\theta(x) = O_{a.co.} \left(h_n^\beta + \sqrt{\frac{\log n}{n^{1-\gamma} G_\theta(x, h_n)}} \right), \quad (1.3)$$

où la probabilité des petites boules est définie par $G_\theta(x, h_n) := \mathbb{P}(|\langle \theta, x - X_i \rangle| \leq h_n)$.

Le théorème suivant montre la normalité asymptotique de $\hat{r}_\theta(x)$.

Théorème 1.2 *Sous certaines hypothèses, nous avons, quand $n \rightarrow \infty$,*

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) \xrightarrow{\mathcal{D}} {}^1N(0, \sigma_{\theta,1}^2(x)), \quad (1.4)$$

et

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta,1}(x) - r_{\theta}(x)\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - r_{\theta}(x)\widehat{r}_{\theta,0}(x))) \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,2}^2(x)), \quad (1.5)$$

où $\sigma_{\theta,1}^2(x) = \frac{C_2}{C_1^2} \frac{\varphi(x)}{f_1(x)}$ et $\sigma_{\theta,2}^2(x) = \frac{C_2}{C_1^2} \frac{(\varphi(x) - r_{\theta}^2(x))}{f_1(x)}$ avec $\varphi(z) := \mathbb{E}(Y_1^2 | X_1 = z)$ et $f_1(\cdot)$ est la densité de probabilité de X_1 telle que $f_1(x) > 0$.

Corollaire 1.1 *Sous les hypothèses du Théorème 1.2 nous avons, quand n tends vers l'infini,*

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta}(x, h_n) - r_{\theta}(x)) \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,2}^2(x)). \quad (1.6)$$

Dans le chapitre qui suit, nous étudions les propriétés asymptotiques de l'estimateur à noyau de la fonction de régression avec la méthode des k -NN sous le même modèle MIFS et quand la variable explicative est à valeurs dans un espace semi-métrique.

Chapitre 4 : La méthode des k plus proches voisins dans le modèle à indice fonctionnel simple pour les données de séries chronologiques quasi-associées.

On s'intéresse au modèle de régression fonctionnelle non-paramétrique (1.1). Le principal objectif est la caractérisation de la consistance de la fonction de régression.

La méthode des k plus proches voisins notée k -NN consiste donc à considérer les k variables X_i les plus proches de x au sens de la semi-métrique d . Soit $x \in H$ fixé, nous nous plaçons dans le cadre de l'estimation de l'opérateur $r_{\theta}(\cdot) = \mathbb{E}(Y | \langle \theta, X_i = \cdot \rangle)$.

L'estimateur des k plus proches voisins de r en x est défini de la façon suivante

$$\widehat{r}_{k,\theta}(x, H_{n,k}) = \frac{\sum_{i=1}^n Y_i K_i(x)}{\sum_{i=1}^n K_i(x)} \quad \text{pour tout } x \in H, \quad (1.7)$$

où $K_i(x) := K\left(\frac{\langle \theta, x - X_i \rangle}{H_{n,k}}\right)$ est le noyau. Pour k fixé, le paramètre $H_{n,k}$ est une variable aléatoire de \mathcal{A} dans \mathbb{R}^+ qui dépend de (X_1, \dots, X_n) , et est définie comme suit :

$$H_{n,k}(x) = \min \left\{ h_n \in \mathbb{R}^+ : \sum_{i=1}^n \mathbb{1}_{B(x, h_n)}(x_i) = k \right\}, \quad (1.8)$$

1. $\xrightarrow{\mathcal{D}}$ notée la convergence en distribution

où $B(x, h_n)$ est la boule de centre x et de rayon $h_n > 0$, définie par l'ensemble $\{x' \in H, d_\theta(x, x') \leq h_n\}$ et $\mathbb{1}_A$ est la fonction indicatrice de l'ensemble A .

Nous utilisons, pour montrer la convergence presque complète de l'estimateur à noyau des k plus proches voisins, un résultat établi par Collomb (1980), et qui a été ensuite généralisé par Burba *et al.* (2009) au cadre fonctionnel.

Lemme 1.1 (voir Burba *et al.* (2009)) *Soit $(D_n(x))_{n \geq 1}$ une suite de variables aléatoires réelles. Si, $\forall \beta \in]0, 1[$ fixé, il existe deux suites de variables aléatoires réelles $(D_n^-)_{n \geq 1}$ et $(D_n^+)_{n \geq 1}$ telles que :*

$$\text{C1. } \forall n \geq 1, \forall x \in H, D_n^-(\beta_n, x) \leq D_n^+(\beta_n, x) \text{ et } \mathbb{1}_{\{D_n^-(\beta_n, x) \leq D_n(x) \leq D_n^+(\beta_n, x)\}} \xrightarrow{p.co.} 1,$$

$$\text{C2. } \frac{\sum_{i=1}^n L(D_n^-(\beta_n, x), (x, A_i))}{\sum_{i=1}^n L(D_n^+(\beta_n, x), (x, A_i))} - \beta_n = O_{p.co.}(\mathcal{U}_n),$$

$$\text{C3. } R_{n,x}(D_n^-(\beta_n, x)) - R(x) = O_{p.co.}(\mathcal{U}_n) \text{ et } R_{n,x}(D_n^+(\beta_n, x)) - R(x) = O_{p.co.}(\mathcal{U}_n).$$

Alors,

$$R_{n,x}(D_n(x)) - R(x) = O_{p.co.}(\mathcal{U}_n). \quad (1.9)$$

Sous la méthode des k -NN, nous établissons la convergence presque complète de $\hat{r}_{k,\theta}(H_{n,k}, x)$ vers $r_\theta(x)$.

Théorème 1.3 *Sous certaines hypothèses, en outre, nous supposons que la suite de nombres réels positifs $k = k_n$ satisfait $\frac{k}{n} \rightarrow 0$ et qu'il existe $\gamma \in (0, 1)$ tel que $\frac{n^\gamma \log(n)}{k} \rightarrow 0$ quand n tends vers l'infini et pour tout $\epsilon_1, \epsilon_2 > 0$ tel que $\frac{n \log(n)}{n^{1-\gamma-\epsilon_2}} \leq k \leq \frac{n}{(\log(n))^{1+\epsilon_1}}$. Alors*

$$\hat{r}_{k,\theta}(H_{n,k}, x) - r_\theta(x) = O_{p.co.}(\mathcal{U}_n), \quad (1.10)$$

Avec $\phi(\epsilon)$ une fonction non négative, tends vers zéro quand $\epsilon \rightarrow 0$, et $\mathcal{U}_n = (\phi^{-1}(\frac{k}{n}))^\alpha + \sqrt{\frac{n^\gamma \log(n)}{k}}$.

Dans le présent chapitre, la méthode k -NN, a été mise en oeuvre dans le modèle de régression fonctionnel à indice simple pour une variable réponse réelle et une variable explicative à valeurs dans un espace semi-métrique (de dimension infinie), et sous la condition de dépendance de quasi-association. Notre apport principal dans ce chapitre est la caractérisation des propriétés asymptotiques de l'estimateur, telles que la vitesse de convergence presque complète. Les résultats obtenus sont basés d'une part sur les résultats de l'estimation fonctionnelle classique, et d'autre part sur les lemmes de Kallabis and Neumann (2006), Douge (2007) et Burba *et al.* (2009).

Une extension des résultats des deux chapitres précédents permet de chercher à définir une autre mesure de risque appelée l'expectile qui répond le mieux aux besoins de la modélisation des risques. Dans le chapitre 5, présenté ci-dessous, nous établissons la consistance et la normalité asymptotique de l'estimateur à noyau des expectiles conditionnels.

Chapitre 5 : Les propriétés asymptotiques de l'estimateur à noyau des expectiles conditionnels pour les données fonctionnelles.

Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathcal{A} \times \mathbb{R}$, où (\mathcal{A}, d) un espace semi-métrique de dimension infinie équipé d'une semi-métrique $d(\cdot, \cdot)$. Pour $0 < p < 1$, nous définissons l'expectile conditionnel $\theta(p; x)$ d'ordre p de la fonction croissante $G(\cdot; x)$, défini ci-dessous, comme une solution unique par rapport à t du problème d'optimisation suivant :

$$\min_{t \in \mathbb{R}} \left\{ \mathbb{E} [p(Y - t)^2 \mathbf{1}_{(Y-t) > 0} | X = x] + \mathbb{E} [(1-p)(Y - t)^2 \mathbf{1}_{(Y-t) \leq 0} | X = x] \right\}. \quad (1.11)$$

Par une simple manipulation, nous montrons que $\theta(p; x)$ est la solution unique pour t de

$$\zeta(p; x) = G(t; x) := \frac{G^1(t; x)}{G^2(t; x)}, \quad (1.12)$$

où

$$\begin{cases} \zeta(p; x) &= \frac{p}{1-p}, \\ G^1(t; x) &= -\mathbb{E} [(Y - t) \mathbf{1}_{(Y-t) \leq 0} | X = x], \\ G^2(t; x) &= \mathbb{E} [(Y - t) \mathbf{1}_{(Y-t) > 0} | X = x]. \end{cases}$$

Notons que l'expectile conditionnel peut également s'interpréter comme l'espérance conditionnelle de la queue de la distribution de Y (voir l'équation (5.4) du chapitre 5). De plus, l'équation (1.12) est plus pertinente dans le contexte de la régression, et qui révèle que les expectiles conditionnel, tout comme la moyenne conditionnelle, sont des moyennes conditionnelles pondérées. La seule particularité est que le poids est une variable aléatoire, fonction de la distribution de Y .

Ainsi, l'expectile conditionnel $\theta(p; x)$ est défini comme le quantile conditionnel de la manière suivante

$$\theta(p; x) = \inf \{ t \in \mathbb{R} : G(t; x) \geq \zeta(p; x) \}. \quad (1.13)$$

Nous considérons n couples $(X_i, Y_i)_{i=1, \dots, n}$ de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) comme (X, Y) . Soit $x \in \mathcal{A}$ fixé. L'estimateur de $G(\cdot; x)$ en x est défini de la façon suivante

$$\widehat{G}_{n, h_n}(t; x) = \frac{-\sum_{i=1}^n W_{ni}(x)(Y_i - t) \mathbf{1}_{(Y_i - t) \leq 0}}{\sum_{i=1}^n W_{ni}(x)(Y_i - t) \mathbf{1}_{(Y_i - t) > 0}}, \quad \text{for } t \in \mathbb{R}, \quad (1.14)$$

où

$$W_{ni}(x) = \frac{K_i(x)}{\sum_{i=1}^n K_i(x)} \text{ et } K_i(x) = K(h^{-1}d(x, X_i)).$$

En utilisant les variables pondérées locales du noyau fonctionnel, nous pouvons écrire

$$\widehat{G}_{n, h_n}(t; x) = \frac{\widehat{G}_{n, h_n; -}(t; x)}{\widehat{G}_{n, h_n; +}(t; x)}, \quad (1.15)$$

où

$$\begin{cases} \widehat{G}_{n,h_n;-}(t; x) &= -\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)(Y_i - t)\mathbb{1}_{(Y_i - t) \leq 0}, \\ \widehat{G}_{n,h_n;+}(t) &= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)(Y_i - t)\mathbb{1}_{(Y_i - t) > 0}. \end{cases}$$

Par conséquent, nous estimons l'expectile conditionnel $\theta(p; x)$ par

$$\widehat{\theta}_{n,h_n}(p; x) = \arg \min_{t \in \mathbb{R}} \widehat{\Psi}_p(x, t), \quad (1.16)$$

où $\widehat{\Psi}_p(x, t)$ est l'estimateur de $\Psi_p(x, t)$ défini par

$$\widehat{\Psi}_p(x, t) = \sum_{i=1}^n W_{ni}(x) \rho_p(Y, t), \quad \text{pour } t \in \mathbb{R}.$$

Il s'ensuit alors que l'estimateur du noyau de l'expectile conditionnel $\theta(p; x)$, noté $\widehat{\theta}_{n,h_n}(p; x)$ est explicitement défini par

$$\widehat{\theta}_{n,h_n}(p; x) = \inf\{t \in \mathbb{R} : \widehat{G}_{n,h_n}(t; x) \geq \zeta(p; x)\}. \quad (1.17)$$

Dans le cas i.i.d., nous établissons la convergence presque complète de $\widehat{\theta}_{n,h_n}(p; x)$ vers $\theta(p; x)$

$$\sum_{n \geq 1} \mathbb{P}\left(\left|\widehat{\theta}_{n,h_n}(p; x) - \theta(p; x)\right| > \epsilon\right) < \infty. \quad (1.18)$$

Proposition 1.1 *Nous démontrons sous certaines hypothèses que l'on a, pour tout $t \in \mathbb{R}$, quand $n \rightarrow \infty$,*

$$\left|\widehat{G}_{n,h_n}(t; x) - G(t; x)\right| = O\left(h_n^{k_1}\right) + O\left(h_n^{k_2}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right), \quad (1.19)$$

où h_n et $\phi_x(h_n)$ sont respectivement le paramètre de lissage et la probabilité des petites boules.

Théorème 1.4 *Soit $\widehat{\theta}_{n,h_n}(p; x)$ l'estimateur de l'expectile conditionnel $\theta(p; x)$. Sous les hypothèses de la Proposition 1.1, alors, pour tout $\epsilon > 0$, on obtient*

$$\sum_{n \geq 1} \mathbb{P}\left(\left|\widehat{\theta}_{n,h_n}(p; x) - \theta(p; x)\right| > \epsilon\right) < \infty. \quad (1.20)$$

Notons que le résultat du Théorème 1.4 peut être considéré comme une généralisation du Théorème 6.1 dans Ferraty and Vieu (2006) sur l'opérateur de régression, par le fait que la régression expectile d'ordre $p = 0.5$ se réduit à la régression classique $\mathbb{E}[Y | X = x]$.

Théorème 1.5 *Sous certaines hypothèses nous avons, quand $n \rightarrow \infty$,*

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\widehat{\theta}_{n,h_n}(p; x) - \theta(p; x)\right) \xrightarrow{\mathcal{D}} N(0, 1), \quad (1.21)$$

où

$$\sigma_p^2(x) = \frac{\alpha_2(x)\gamma_p(\theta(p;x);x)}{\alpha_1^2(x)\Lambda_p^2(\theta(p;x);x)}, \quad (1.22)$$

$$\alpha_j(x) = K^j(1) - \int_0^1 (K^j)'(s)\beta_x(s)ds \text{ for } j = 1, 2, \quad (1.23)$$

et

$$\gamma_p(\theta(p;x);x) = \left(\frac{p}{1-p}\right)^2 R_+^x(\theta(p;x)) + R_-^x(\theta(p;x)),$$

où

$$R_+(\theta(p;x);x) = \mathbb{E}\left((Y_1 - \theta(p;x))^2 \mathbb{1}_{\{Y_1 > \theta(p;x)\}} \mid X = x\right),$$

$$R_-(\theta(p;x);x) = \mathbb{E}\left((Y_1 - \theta(p;x))^2 \mathbb{1}_{\{Y_1 \leq \theta(p;x)\}} \mid X = x\right),$$

et

$$\Lambda_p(\theta(p;x);x) = G^{1'}(\theta(p;x);x) - \left(\frac{p}{1-p}\right) G^{2'}(\theta(p;x);x).$$

De plus, la normalité asymptotique dans le Théorème 1.5 généralise la normalité asymptotique de l'opérateur de la régression obtenue dans le Théorème 2 de Ferraty *et al.* (2007). Notons, ici, que la normalité asymptotique présente un intérêt particulier pour certaines études asymptotiques diverses.

Cas dépendant : Le théorème suivant énonce la propriété asymptotique de l'estimateur $\widehat{\theta}_{n,h_n}(p;x)$ dans le cas des séries chronologiques fonctionnelles en utilisant la condition α -mélangeante.

Théorème 1.6 *Soit $\widehat{\theta}_{n,h_n}(p;x)$ l'estimateur de l'expectile conditionnel $\theta(p;x)$. Sous des hypothèses décrites dans le chapitre 5 nous avons, pour tout $\epsilon > 0$, quand $n \rightarrow \infty$,*

$$\mathbb{P}\left(\left|\widehat{\theta}_{n,h_n}(p;x) - \theta(p;x)\right| > \epsilon\right) \rightarrow 0. \quad (1.24)$$

Ce dernier résultat sera établi dans le chapitre 5, de cette thèse, d'une manière assez courte.

La principale contribution de cette thèse sur ce modèle est l'obtention des propriétés asymptotiques de l'estimateur de l'expectile de régression non paramétrique. Plus précisément, nous construisons un estimateur de type noyau de l'expectile de régression et nous établissons la convergence presque complète avec la vitesse de convergence. De plus, nous montrons la normalité asymptotique de l'estimateur proposé sous des conditions standards de l'estimation nonparamétrique fonctionnelle. Par ailleurs, nous expliquons comment appliquer nos résultats pour construire les intervalles de confiance et discutons l'impact potentiel de la régression fonctionnelle des expectiles dans la NFDA avec un accent particulier sur la classification supervisée, la prédiction et le risque financier. Les résultats obtenus sont détaillés dans le chapitre 5 de cette thèse, et nous pensons que ce sont les premiers résultats disponibles dans la littérature portant sur l'estimation de l'expectile conditionnel dans le cas fonctionnel via la méthode du noyau.

Enfin, dans le chapitre 6 nous exposons une conclusion générale et des perspectives de recherches permettant d'étendre et de généraliser les résultats de cette thèse.

Bibliographie

- Burba, F., Ferraty, F., and Vieu, P. (2009). k-nearest neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics*, **21**(4), 453–469.
- Collomb, G. (1980). Estimation de la régression par la méthode des k points les plus proches avec noyau : quelques propriétés de convergence ponctuelle. In *Nonparametric asymptotic statistics (Proc. Conf., Rouen, 1979) (French)*, volume 821 of *Lecture Notes in Math.*, pages 159–175. Springer, Berlin.
- Douge, L. (2007). Vitesses de convergence dans la loi forte des grands nombres et dans l'estimation de la densité pour des variables aléatoires associées. *Comptes Rendus Mathématique*, **344**(8), 515 – 518.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data : inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**(3), 267–286.
- Kallabis, R. S. and Neumann, M. H. (2006). An exponential inequality under weak dependence. *Bernoulli*, **12**(2), 333–350.

Chapitre 2

Introduction générale.

2.1 Données fonctionnelles et Estimation non paramétrique

Au cours des deux dernières décennies, une immense innovation sur le développement informatique a vu le jour, et ceci grâce à l'augmentation des capacités de stockage qui permet d'enregistrer des données volumineuses. Ce développement technologique a contribué à l'évolution de la branche statistique, consacrée à l'analyse fonctionnelle, tant en termes des développements théoriques et méthodologiques que de la diversification des domaines d'application. Ainsi, a été développée pour traiter les observations comme des éléments aléatoires fonctionnels. Dans ce cadre, nous donnons quelques définitions permettant de fixer le vocabulaire utilisé. Tout d'abord, rappelons qu'une variable aléatoire fonctionnelle est tout simplement une variable aléatoire à valeurs dans un espace de dimension éventuellement infinie que nous noterons \mathcal{A} (cf. [Ferraty and Vieu \(2006\)](#)). Par exemple, cet espace \mathcal{A} peut être un espace de fonctions, d'opérateur linéaires..., et le fait que ces observations sont supposées appartenir à un tel espace, c'est la principale source de difficulté dans ce domaine.

Une observation d'une variable fonctionnelle $X = \{X_t, t \in T\}$ est appelée donnée fonctionnelle, i.e. pour chaque $t \in T$, $X_t : \Omega \rightarrow \mathbb{R}$ est une fonction (cf. [Ferraty and Vieu \(2006\)](#)). D'où l'appellation de données fonctionnelles. La variable est associée à un ensemble unidimensionnel $T \subset \mathbb{R}$ dans le cas d'une courbe. Il est essentielle de signaler que la notion de variable fonctionnelle occupe une place plus importante que l'analyse des courbes.

Suite à notre contexte de données fonctionnelles, on définira le modèle fonctionnel non-paramétrique :

Définition 2.1 (cf. [Ferraty and Vieu \(2006\)](#)) *On appelle modèle fonctionnel, tout modèle prenant en compte au moins une variable aléatoire fonctionnelle.*

Soit X une variable aléatoire à valeurs dans un espace de dimension infini \mathcal{A} et soit une application φ définie dans \mathcal{A} et dépendant de la distribution de X . Un modèle d'estimation de φ consiste à introduire une certaine contrainte de la forme

$$\varphi \in \mathcal{C},$$

d'où la définition suivante :

Définition 2.2 (cf. [Ferraty and Vieu \(2006\)](#)) *Un modèle fonctionnel est dit paramétrique réel (ou vectoriel) si la classe de fonctions \mathcal{C} est indexable par un nombre fini de paramètres réels (ou vectoriels). Dans le cas contraire, il est dit non-paramétrique, i.e. que la classe \mathcal{C} est indexable par un nombre infini de paramètres d'un espace de dimension infinie (espace fonctionnel).*

Les données fonctionnelles sont largement employées dans les problèmes non paramétriques. Historiquement, les premiers travaux dans ce domaine sont relativement anciens, l'analyse en composantes principales (ACP) et l'analyse factorielle (AF) fonctionnelles qui ont été étudiées par [Rao \(1958\)](#) et [Tucker \(1958\)](#). [Ramsay \(1982\)](#) a dégagé la notion de données fonctionnelles et a soulevé la question de l'adaptation des méthodes utilisées en analyse statistique de données multivariées (en dimension finie) au cadre fonctionnel. [Deville \(1974\)](#) quant à lui, a généralisé l'ACP à des courbes. La généralisation de l'ACP dans un espace de Hilbert a été introduite par [Dauxois and Pousse \(1976\)](#). Quant à eux, [Besse and Ramsay \(1986\)](#) ont montré ensuite qu'une ACP fonctionnelle correspond à une norme, tandis que [Besse et al. \(1997\)](#) ont réalisé une ACP fonctionnelle avec contraintes de lissage sur la norme. Nous renvoyons le lecteur au livre de [Ramsay and Silverman \(2005\)](#) qui a donné un aperçu général et de l'état de l'art de l'analyse en composantes principales fonctionnelles. Ainsi, de nombreux auteurs ont étudiés les variables fonctionnelles comme étant des variables aléatoires de carrés intégrables c'est-à-dire à valeurs dans $L^2(0, 1)$ (cf. [Crambes et al. \(2009\)](#)) ou plus généralement dans un espace de Hilbert (cf. [Preda \(2007\)](#)), ou de Banach (cf. [Cuevas and Fraiman \(2004\)](#)) ou un espace métrique (cf. [Dabo-Niang and Rhomari \(2003\)](#)).

Rappelons que, l'approche multivariée classique ne permet pas de prendre en compte ni de tirer profit de la nature fonctionnelle des données observées, et se heurte à un certain nombre d'écueils théoriques et pratiques. Tout d'abord, le nombre de points de discrétisation ou la dimension de la variable à étudier est souvent importante par rapport à la taille de l'échantillon. De plus, les points de discrétisation ne sont pas toujours espacés régulièrement et il arrive fréquemment que la nature des points de discrétisation, ainsi que leur nombre, soit très différents d'une courbe à une autre. D'autre part, les différentes composantes de la variable multivariée étudiée sont clairement dépendantes car elles résultent de la discrétisation d'une même fonction. Si on souhaite, par exemple, estimer de manière non-paramétrique, l'effet de n courbes discrétisées en p points sur n variables réelles, d'après les résultats de type minimax obtenus par ([Stone \(1980\)](#), [Stone \(1982\)](#), [Stone \(1983\)](#)), la vitesse optimale de convergence des estimateurs non-paramétriques (de la fonction de régression, de la densité et les quantiles) se détériore très rapidement lorsque la dimension de la variable explicative augmente. Cela est dû à un phénomène bien connu que l'on appelle couramment fléau de la dimension. Ce dernier n'est pas lié à l'estimateur non-paramétrique utilisé mais plutôt à la nature du modèle étudié. C'est à partir de ce constat qu'ont été faites des propositions de modèles pour réduire la dimension. On peut notamment citer le modèle additif ([Stone \(1985\)](#)), le modèle additif généralisé ([Hastie and Tibshirani \(1986\)](#), [Hastie and Tibshirani \(1990\)](#)), le modèle à transformations optimales ([Breiman and Friedman \(1985\)](#)), le modèle à projections révélatrices ([Diaconis and Shahshahani \(1984\)](#)), le modèle à indice simple ([Ichimura \(1993\)](#), [Hristache et al. \(2001\)](#) et [Delecroix et al. \(2003\)](#)) et le modèle à indices multiples ([Li \(1991\)](#)).

Enfin, si ces modèles sont satisfaisants en statistique multivariée en terme de vitesse de convergence en ayant contourné le fléau de la dimension, leur application à des données fonctionnelles ne prend pas en compte la nature fonctionnelle des données et n'en tire pas profit, et lorsque le nombre p de points de discrétisation augmente, c'est-à-dire lorsque nous avons une meilleure connaissance de nos courbes, les estimateurs non-paramétriques construits à partir des valeurs discrétisées ont des vitesses de convergence moins bonnes, c'est pour quoi il est important de construire un modèle qui puisse s'adapter facilement à n'importe quelle finesse de discrétisation sans perdre de la qualité.

Avec l'arrivée des méthodes d'acquisition et la précision des appareils de mesures modernes permettent souvent de récolter des données discrétisées sur des grilles de plus en plus fines, ce qui les rend fondamentalement fonctionnelles et de considérer, par exemple, les dérivées, les primitives, l'intégrale, les valeurs en points non observés, la périodicité, la régularité, le taux de variation, le déphasage de deux courbes, comme des caractères typiquement fonctionnels. De plus, on n'est plus confronté au paradoxe qu'une meilleure connaissance des données (discrétisation sur une grille plus fine), entraîne une détérioration des vitesses de convergence de nos estimateurs. D'autre part, l'ensemble de p variables dépendantes peut être remplacé par une seule variable fonctionnelle pour laquelle on construit des modèles et des méthodes adéquates. Ce type de variables se retrouve dans de nombreux domaines, comme la météorologie (cf. Besse *et al.* (2000)), les données d'expression temporelle des gènes (cf. Leng and Müller (2006)), la reconnaissance de la parole (cf. Rossi and Villa (2008)), la spectrométrie de masse (cf. Martin-Barragan *et al.* (2014)), les processus chimiques (cf. Blanquero *et al.* (2018)). Ainsi, pour d'autres aspects appliqués, on peut citer le livre de Ramsay and Silverman (2002). Parmi les ouvrages de référence en la matière, on peut citer aussi Ramsay and Silverman (2005) et Bosq (2012) pour les aspects théoriques, Ferraty and Vieu (2006) pour une étude non paramétrique et Ferraty and Romain (2011) pour des développements plus récents. Dans le même contexte, nous renvoyons à Manteiga and Vieu (2007) ainsi que Ferraty (2010). De plus, même si les données dont dispose le statisticien ne sont pas de nature fonctionnelle, celui-ci peut être amené à étudier des variables fonctionnelles construites à partir de son échantillon initial. Un exemple classique est celui où l'on observe plusieurs échantillons de données réelles indépendantes et où l'on est ensuite amené à comparer les densités de ces différents échantillons ou bien à considérer des modèles où elles interviennent (cf. Ramsay and Silverman (2002)). Dans le cas particulier de l'étude des séries temporelles, l'approche introduite par Bosq (1991) fait apparaître une suite de données fonctionnelles dépendantes qui modélisent la série chronologique observée. Cette approche consiste tout d'abord à considérer le processus non pas à travers sa forme discrétisée mais comme étant un processus à temps continu puis à le découper en un échantillon de courbes successives. Citons aussi, les travaux de Cuevas *et al.* (2002), qui se sont intéressés au problème de la régression linéaire d'une variable fonctionnelle sur un ensemble de données fonctionnelles déterministes, et Benhenni *et al.* (2010) ont considéré le problème d'estimation de l'opérateur de régression quand les données fonctionnelles sont déterministes et les erreurs sont corrélées. D'autre part, Cardot *et al.* (2005) quant à eux, ils ont proposé un estimateur non paramétrique de l'opérateur de régression quand le facteur prédictif est réel et la variable réponse est une courbe.

Par ailleurs, l'étude du modèle de régression non linéaire est beaucoup plus récente que celle du cas linéaire. En effet, les premiers résultats ont été obtenus par Gasser *et al.* (1998), les auteurs se sont intéressés à l'estimation non paramétrique du mode de distribution d'une variable fonctionnelle vérifiant une condition fractale. Sous la même condition, Ferraty and Vieu (2000) ont établi la convergence presque complète d'un estimateur de noyau de la fonction de régression, lorsque les observations sont indépendantes et identiquement distribuées. Ensuite, ces résultats ont été étendus par Ferraty and Vieu (2002) en traitant le cas de données dépendantes et en établissant des convergences fortes de l'estimateur à noyau de la régression.

Une extension du modèle à indice simple au cadre fonctionnel dont les variables aléatoires explicatives est à valeurs dans un espace de dimension infinie est proposé par Ferraty *et al.* (2003). Ils ont introduit des hypothèses sur la loi de la v.a. explicative moins restrictives que celles utilisées habituellement dans le cadre vectoriel. Dabo-Niang and Rhomari (2003) ont étudié la convergence en norme L^p de l'estimateur à noyau de l'opérateur de régression et ils ont expérimenté leur résultats à la discrimination de variables d'un espace semi-métrique. Ferraty and Vieu (2004) ont étudié la convergence presque complète pour le cas fortement mélangeant, Masry (2005) a montré la normalité asymptotique de l'estimateur de Ferraty and Vieu (2004) pour la fonction de régression. Quant à Delsol (2007), a étudié la normalité asymptotique

de l'estimateur à noyau pour des données α -mélangeantes fonctionnelles. Par ailleurs, la convergence en moyenne quadratique a été étudiée par [Ferraty et al. \(2007\)](#). Plus précisément, ils ont explicité d'une manière exacte le terme de l'erreur quadratique asymptotique. Ce résultat a été utilisé par [Rachdi and Vieu \(2007\)](#) pour donner une contribution basée sur la construction d'un critère sur le choix automatique et optimal du paramètre de lissage pour l'estimateur de la régression quand le régresseur est de type fonctionnel. La version locale de ce critère a été étudiée par [Benhenni et al. \(2007\)](#). Ils ont présenté une étude comparative entre l'approche locale et globale. La normalité asymptotique des estimateurs à noyau du mode conditionnel et des quantiles conditionnels a été étudiée par ([Ezzahrioui and Ould-Saïd \(2008\)](#), [Ezzahrioui and Ould Saïd \(2008\)](#), [Ezzahrioui and Ould-Saïd \(2008a\)](#) et [Dabo-Niang and Laksaci \(2012\)](#)) en traitant deux cas (i.i.d. et le cas mélange). Une autre méthode d'estimation des quantiles conditionnels a été proposée par [Laksaci et al. \(2009\)](#). Ils ont illustré des résultats asymptotiques tels que la convergence presque complète et la normalité asymptotique dans le cas i.i.d. Quant à [Laksaci et al. \(2011\)](#), ont étudié, en norme L^1 , la convergence presque complète et la normalité asymptotique de l'estimateur à noyau des quantiles conditionnels.

Le concept de dépendance de quasi-association pour des variables aléatoires qui prennent ses valeurs dans un espace de Hilbert a été introduit par [Douge \(2010\)](#), et a obtenu quelques théorèmes limites pour ce type de variables. Ce concept d'association peut être considéré comme un cas particulier de la condition de dépendance faible qui permet de traiter la corrélation de mélange et l'association dans une approche unifiée. Dans le travail de [Mechab and Laksaci \(2016\)](#) ont établi la régression relative non paramétrique pour les variables aléatoires associées, cette méthode est basée sur la minimisation de l'erreur relative quadratique moyenne. Plus récemment, [Douge \(2018\)](#) a étudié la normalité asymptotique de la fonction de régression lorsque la variable explicative prend ses valeurs dans un espace de Hilbert et sous des données quasi-associées.

Notons que les quantiles conditionnels ne sont pas toujours satisfaisants et peuvent être critiqués du fait qu'elles sont difficiles à calculer car la fonction de perte correspondante n'est pas continûment différentiable. En particulier, sa grande sensibilité à la valeur extrême. Une autre mesure de risque, appelée expectile, est construite en combinant les idées de régression des moindres carrés à celles du quantile conditionnel permettant d'accumuler les avantages des deux modèles. L'expectile conditionnel a été largement étudié dans des domaines appliqués tels que l'économétrie, la finance et la science actuarielle, voir par exemple [Kuan et al. \(2009\)](#) et [Ehm et al. \(2016\)](#) en tant que travaux pionniers.

Cependant, d'un point de vue théorique, un tel modèle de risque a été introduit par les auteurs [Aigner et al. \(1976\)](#) et [Newey and Powell \(1987\)](#) dans le contexte des modèles de régression linéaire. La régression expectile est une généralisation de la régression classique et qui constitue un bon modèle de risque alternatif à la régression quantile. Malgré cette importance, la régression expectile est inexplorée par rapport aux deux régressions compétitives (la moyenne conditionnelle et quantile). A titre d'exemple, on cite [Abdous and Rémillard \(1995\)](#), sous l'hypothèse de la symétrie pondérée, ont établi qu'une caractéristique importante est que les quantiles et les expectiles possèdent la même distribution, et coïncident avec le centre de la symétrie pondérée (cf. Théorème 2.2 dans [Abdous and Rémillard \(1995\)](#)). Ainsi, ont souligné, que l'inférence sur les expectiles est beaucoup plus facile que l'inférence sur les quantiles. Par ailleurs, la méthode du noyau basée sur l'ajustement localement linéaire est adaptée à l'estimation non paramétrique de la régression d'expectiles et des centiles en utilisant l'approche des moindres carrés asymétriques (ALS), et sous les hypothèses de stationnarité et de mélange, la normalité asymptotique pour les estimateurs des expectiles conditionnels est entièrement discuté dans le papier de [Yao and Tong \(1996\)](#), les tests de la symétrie pondée non paramétriques ont été étudiés par [Abdous et al. \(2003\)](#).

En se basant sur les distributions marginales, [Maume-Deschamps et al. \(2017\)](#) ont étudié le comportement asymptotique des expectiles multivariés dans le cadre des variations régulières multivariées. [Bellini](#)

et al. (2018) quant à eux, ils ont généralisé l'espérance conditionnelle aux expectiles conditionnels par la minimisation d'une fonction de perte quadratique asymétrique et ils ont révélé ses principales caractéristiques. Une étude rigoureuse des techniques d'estimations des expectiles extrêmes dans le contexte stationnaire et faiblement dépendant a été amenée par Padoan and Stupfler (2020), ils ont montré la convergence dans le cas de mélange. Par ailleurs, une discussion plus approfondie sur les avantages de ce modèle par rapport aux mesures de risque classiques telles que les quantiles et le "expected shortfall" (ES). Bellini and Bernardino (2017) ont indiqué dans leur papier, que les expectiles sont des alternatives parfaitement raisonnables aux mesures de risques telles que la "value at risk" (VaR) et le ES. Daouia *et al.* (2018) ont proposé une estimation de VaR, ES et "conditional value at risk" (CVaR). De plus, selon les travaux de Bellini *et al.* (2014) et Bellini and Bignozzi (2015), ont montrés que l'expectile est la seule mesure de risque cohérente et élicitable. Pour plus de détails sur la signification statistique de la propriété d'élicitabilité et de cohérence, nous renvoyons le lecteur intéressé aux articles de Gneiting (2011) et Ziegel (2016).

D'autres travaux se sont intéressés à l'estimation de la fonction de régression en utilisant différentes approches : la méthode des k plus proches voisins ou " k -Nearest Neighbors" (k -NN) par Burba *et al.* (2009), les techniques robustes par Azzedine *et al.* (2008), Crambes *et al.* (2008) et Attouch *et al.* (2009), ou aussi l'estimation par la méthode simplifiée de polynômes locaux par Barrientos-Marin *et al.* (2010).

La méthode k -NN pour estimer la régression est largement utilisée et étudiée depuis de nombreuses années dans le cas multivarié. Les premiers travaux ont été présentés par Royall (1966), Collomb (1980) et Stone (1982). Mack (1981) a étudié la convergence L^2 et la distribution asymptotique. Sous diverses conditions de bruit, Devroye (1978) a démontré que les estimations sont fortement uniformément cohérentes. Pour une étude complète voir, par exemple, Bosq and Lecoutre (1987) et Györfi *et al.* (2002).

Le véritable intérêt de la méthode k -NN vient de la nature du paramètre de lissage. En effet, dans la méthode traditionnelle du noyau, le paramètre de lissage h_n est un nombre réel positif, tandis que dans la méthode k -NN, le nombre de voisins k est le paramètre de lissage et il prend ses valeurs dans un ensemble discret. Cependant, la méthode k -NN présente une double difficulté technique : la première est la sélection aléatoire des voisins les plus proches. La deuxième difficulté est liée à la nature fonctionnelle des données, et du fait que nous ne supposons pas l'existence d'une densité (car, dans les espaces de dimension infinie, nous n'avons pas de mesure standard comme la mesure de Lebesgue dans le cas multivarié).

De nombreux travaux ont été consacrés à l'étude des modèles non paramétriques via la méthode k -NN, on retrouve dans Collomb (1980) un lemme fondamental qui a été généralisé par Burba *et al.* (2009) aux espaces fonctionnels semi métriques. Ils ont prouvé la convergence presque complète de l'estimateur à noyau des k plus proches voisins avec une vitesse de convergence presque complète pour des données indépendantes. L'article de Müller and Dippon (2011) examine la méthode des k plus proches voisins dans la régression fonctionnelle non paramétrique pour les données α -mélangeantes. Lian (2011) quant à lui, il a présenté la vitesse de convergence presque sûre de l'estimation des k les plus proches voisins lorsque la réponse est un élément d'un espace de Hilbert. Nous renvoyons le lecteur aux résultats les plus récents, de l'analyse des données fonctionnelles non paramétriques des k -NN, qui se trouvent dans Kudraszow and Vieu (2013) et Kara *et al.* (2017) et ses références. Récemment, Ling *et al.* (2019) ont considéré le modèle de régression non paramétrique et en établissant la vitesse de convergence uniforme presque complète de l'estimateur k -NN pour des données chronologiques fonctionnelles fortement mélangeantes. Enfin, pour un état de l'art ainsi que des pistes futures importantes sur ce sujet de l'analyse des données fonctionnelles non paramétriques, nous renvoyons le lecteur au travail de Ling and Vieu (2018).

2.1.1 Domaines d'applications des données fonctionnelles

Nous présentons dans cette section quelques exemples de données fonctionnelles provenant de domaines assez variés (Économétrie, Chimiométrie et Climatologie). Pour illustrer les avantages et les défis associés à leur étude à travers une approche fonctionnelle, et de montrer aussi que l'étude de ces données fonctionnelles est en plein essor à la fois en ce qui concerne le nombre d'articles qui leurs sont consacrés et la diversité des phénomènes étudiés.

1. On trouve de nombreux phénomènes qui sont modélisés par des variables fonctionnelles. Parmi ces phénomènes, on peut citer la volatilité des marchés financiers (cf. Müller *et al.* (2011)), le rendement d'une entreprise (cf. Kawasaki and Ando (2004)), et le commerce électronique (cf. Jank and Shmueli (2006)). Par ailleurs, nous citons un exemple, qui consiste à observer les fluctuations d'un indice boursier en fonction du temps : il s'agit typiquement d'une série temporelle qu'on découpe selon des sous-intervalles de l'espace temps (cf. Bosq (2002)).
2. Un autre domaine est celui de l'industrie alimentaire. Les données spectrométriques, représentées par le taux de graisse, sont obtenues en mesurant pour chaque morceau de viande l'absorbance de lumières de différentes longueurs d'onde. Ce sont des données de nature intrinsèquement fonctionnelle comme le souligne Leurgans *et al.* (1993). On peut donc les résumer par des courbes (appelées spectrométriques) représentant l'absorbance en fonction de la longueur d'onde. Ces courbes spectrométriques sont très régulières et de formes semblables en dehors d'une translation verticale. Comme les courbes sont très lisses et de formes semblables, il est donc intéressant d'utiliser les courbes dérivées plutôt que les courbes elles-mêmes.
3. Dans le domaine climatique. L'observation d'un phénomène climatologique assez important comme El Niño avait provoqué des perturbations dans la température. Ce phénomène est un grand courant marin dont l'origine est assez mal connue, et qui survient de manière exceptionnelle le long des côtes péruviennes à la fin de l'hiver. Le courant provoque des dérèglements climatiques à l'échelle de la planète. Par conséquent, ce jeu de données est constitué de relevés de température mensuels de la surface océanique effectués depuis 1950 dans une zone située au large du nord du Pérou (aux coordonnées $0 - 10^\circ$ Sud, $80 - 90^\circ$ Ouest) où peut apparaître le courant marin El Niño. Pour plus de détails sur les données et leur description, nous renvoyons au site suivant : [http : //www.cpc.ncep.noaa.gov/data/indices/](http://www.cpc.ncep.noaa.gov/data/indices/) (voir Delsol (2008)).

Il existe plusieurs domaines dans lesquels les statistiques fonctionnelles sont utilisées comme par exemple la prédiction de la consommation électrique (Voir Ferraty *et al.* (2002)), le traitement de signaux sonores (voir Lucero (1999)) ou l'enregistrement par un radar (voir Hall *et al.* (2001)), la météorologie (voir Ramsay and Silverman (2005), Hall and Vial (2006)). Pour plus de domaines d'application de ces données fonctionnelles, en se référant au livre de Ramsay and Silverman (2002) où on trouve des applications dans des domaines assez variés comme la criminologie. Dans ce livre, ils ont modélisé et comparé l'évolution de la criminalité d'un individu au cours du temps.

Ensuite, dans la section suivante, nous présentons brièvement les semi-métriques et les probabilités de petites boules.

2.1.2 Probabilités de petites boules et semi-métriques

Comme mentionné dans la section introduction aux données fonctionnelles que, dans le cas du modèle de régression nonparamétrique multivarié, le problème du fléau de la dimension provoque une décroissance

exponentielle des vitesses de convergences des estimateurs nonparamétriques en fonction de la dimension (voir Stone (1982)). Par conséquent, il est légitime de penser que les méthodes non-paramétriques dans des modèles de régression fonctionnelle risquent d'avoir une vitesse de convergence très lente. Dans le cas où la variable explicative est multivariée (i.e. à valeurs dans un espace de dimension finie k), les vitesses de convergence de l'estimateur à noyau sont exprimées en fonction d'un terme de la forme h_n^k , provenant de la probabilité que la variable explicative appartienne à la boule de centre x et de rayon h_n . Dans le cas d'une variable explicative fonctionnelle (i.e. à valeur dans un espace semi-métrique de dimension infinie (\mathcal{A}, d) où d est une semi-métrique), les résultats asymptotiques sont exprimés à partir de quantités plus générales appelées probabilités de petites boules et définies par :

$$\phi_x(h_n) := \mathbb{P}(d(x, X) \leq h_n). \quad (2.1)$$

On peut observer, dans la littérature concernant la convergence non-paramétrique de la régression, que la vitesse de convergence est liée à la fois à la loi de la variable explicative, à la topologie que l'on considère et par conséquent à la manière de ces probabilités de petites boules décroissent vers zéro. On peut trouver dans de nombreux papiers de probabilités d'équivalences asymptotiques pour ces probabilités de petites boules lorsque d est une norme (voir par exemple Lifshits *et al.* (2006), Gao and Li (2007)) ou une semi-norme spécifique (voir Lifshits and Simon (2005), Aurzada and Simon (2007)). Dans le cas d'un processus non lisse tels que le mouvement brownien, processus Ornstein-Ullénbeck, ...), ces probabilités de petites boules ont une forme exponentielle (par rapport à h_n) et donc la vitesse de convergence est une puissance de $\log(n)$ (voir Section 5 dans Ferraty *et al.* (2006) et paragraphe 13.3.2. dans Ferraty and Vieu (2006), pour une discussion plus approfondie sur ce sujet).

Pour étudier et donner un sens aux données fonctionnelles on a souvent besoin d'avoir une notion de distance entre celles-ci. Définir une telle distance représente l'une des difficultés majeure dans l'analyse des données fonctionnelles, on est donc amené à donner une écriture probabiliste qui nous conduit à des hypothèses agissant directement sur la distribution de la fonctionnelle explicative plutôt que sur la densité. En effet, dans ce contexte, nous ne pouvons pas manipuler la fonction de densité de probabilité aussi facilement que dans le cas réel ou vectoriel. De plus, en dimension finie toutes les normes sont équivalentes. Mais, n'est évidemment pas le cas en dimension infinie. Ce qui justifie que le choix de la métrique et de la topologie associée est encore plus crucial pour l'étude des variables aléatoires fonctionnelles que pour les statistiques multivariées. Alors, pour de tenter de remédier ce problème est de chercher à remplacer les topologies usuelles induites par des normes par une topologie qui restitue de façon pertinente les proximités entre les données fonctionnelles (voir Ferraty and Vieu (2006), Chapitre 3).

Ce choix de la semi-métrique peut constituer une alternative aux problèmes liés aux données de grande dimension, et se fait en fonction de la nature des données, du problème à étudier et en fonction aussi de nos objectifs pour atteindre les résultats espérés. Il y a deux grandes familles de semi-métriques qui sont les plus utilisables, mais, bien sûr, beaucoup d'autres peuvent être construites. La première semi-métrique est bien adaptée aux courbes irrégulières, tandis que la deuxième concerne des données fonctionnelles qui sont assez lisses.

Analyse en composantes principales fonctionnelle (ACPF). L'Analyse en composantes principales classique (PCA) peut être vue comme un bon outil pour visualiser des données dans un espace de dimension plus réduite. Cette technique a été prolongée aux données fonctionnelles et plus récemment employée pour différents buts statistiques. Nous verrons que l'ACPF ou FPCA (Functional Principal Components Analysis) est devenue un bon outil pour calculer des proximités entre les courbes. Si la variable fonctionnelle est carrée

intégrable, on peut obtenir la décomposition suivante :

$$X = \sum_{k=1}^{\infty} \left(\int X(t) \vartheta_k(t) dt \right) \vartheta_k,$$

où $\vartheta_1, \vartheta_2, \dots$ sont les fonctions propres orthonormales de l'opérateur de covariance

$$\Gamma_X(s, t) = \mathbb{E}(X(s)X(t))$$

associées aux valeurs propres qui sont rangées de manière décroissante $\lambda_1 \geq \lambda_2 \geq \dots$

Notons que, l'entier q un paramètre de réglage indiquant le niveau de résolution auquel le problème est considéré. On en déduit une famille de semi-métriques comme suit :

$$d_q^{ACP}(X_i, x) = \sqrt{\sum_{k=1}^q \left(\int (X_i(t) - x(t)) \vartheta_k(t) dt \right)^2}. \quad (2.2)$$

Si, par exemple, nous discrétisons deux courbes x_{l_i} , pour $i = 1, 2$ alors, la quantité $d_q^{ACP}(x_{l_1}, x_{l_2})$ sera approximée par sa version empirique :

$$d_q^{ACP}(x_{l_1}, x_{l_2}) = \sqrt{\sum_{k=1}^q \left(\sum_{j=1}^J \omega_j (x_{l_1}(t_j) - x_{l_2}(t_j)) \vartheta_k(t_j) dt \right)^2}, \quad (2.3)$$

avec $\omega_1, \omega_2, \dots, \omega_J$ se sont les poids qui définissent l'approximation de l'intégrale, alors qu'un choix standard de poids est donné par $\omega_j = t_j - t_{j-1}$, pour $j \in J$.

Il est important de préciser que cette semi-métrique ne peut être utilisée que si les données sont équilibrées (i.e. les courbes sont observées aux mêmes points) et la grille des mesures suffisamment fine. Cela pourrait apparaître comme un inconvénient pour l'utilisation de ce type de semi-métriques, mais d'un autre côté, elles ont l'avantage d'être utilisables même si les courbes sont irrégulières (cf. [Ferraty and Vieu \(2006\)](#), section 3.4. page 28).

Semi-métrique basée sur des dérivées. Une autre façon de construire une famille de semi-métriques entre les courbes consiste à considérer une distance entre leurs dérivées. Plus précisément, étant donné deux courbes observées X_i et $X_{i'}$. Elle est définie de la manière suivante :

$$d_q^{deriv}(X_i, X_{i'})^2 = \int \left(X_i^{(q)}(t) - X_{i'}^{(q)}(t) \right)^2 dt, \quad (2.4)$$

où $X^{(q)}$ désigne la q -ième dérivée de X .

On remarque, dans le cas où $q = 0$ et $X_{i'} = 0$, que la semi-métrique basée sur des dérivées $d_q^{deriv}(X_i, X_{i'})$ correspond à la norme classique L^2 de X_i .

Notons aussi, dans ce cas, que le calcul des dérivées successives est très sensible numériquement. Afin de contourner ce problème de stabilité numérique, nous pouvons utiliser une approximation B-splines pour les courbes. Une fois que nous avons obtenu un développement analytique sur une base de B-splines pour chaque courbe, les dérivées successives sont directement calculées en différenciant plusieurs fois leur forme analytique. Dans la pratique, cette classe de semi-métriques sera bien adaptée et employée pour les courbes lisses.

2.2 Modèles conditionnels dans l'estimation non paramétrique

Les modèles non paramétriques conditionnels représentent un outil principal aux problèmes de prévision et à l'étude des liens entre les variables aléatoires. Ce type de problème peut être modélisé de la manière suivante :

$$\min_{t \in \mathbb{R}} \psi(x, t), \quad (2.5)$$

où $\psi(x, t) := \mathbb{E}(\rho(Y - t) \mid X = x)$ est la fonction de risque (ou d'erreur) et ρ est une fonction de perte. Si cette fonction de perte est convexe et admet un minimum unique, alors on peut prendre la quantité

$$\hat{r}(x) = \arg \min_{t \in \mathbb{R}} \mathbb{E}(\rho(Y - t) \mid X = x), \quad (2.6)$$

comme approximation de Y sachant $X = x$. En conséquence, on peut dire qu'en statistique non paramétrique, les outils naturels pour faire la prévision sont les modèles liés à la distribution conditionnelle de ces variables aléatoires. Parmi ces modèles les plus sollicités citons la régression classique, la régression modale (mode conditionnel), les quantiles de régression et l'expectile de régression et d'autres. Dans la suite de ce paragraphe, nous rappelons quelques définitions de ces outils.

La Régression classique. L'espérance conditionnelle de Y sachant qu'on a observé X est l'approximation la plus utilisée comme prédicteur de Y . La fonction de perte associée à ce prédicteur est

$$\rho(z) = z^2,$$

donc, la régression classique est l'unique solution de

$$\hat{r}(x) = \mathbb{E}(Y \mid X = x).$$

C'est une estimation équivalente à l'estimation par la méthode de maximum vraisemblance classique lorsque X et Y sont liées par la relation suivante :

$$Y = r(X) + \varepsilon, \quad (2.7)$$

où ε est la variable aléatoire de loi normale, de variance finie σ^2 et indépendante de X .

L'équation (2.7) est le modèle de régression, et c'est l'un des modèles le plus fréquemment rencontré en statistique paramétrique ou non paramétrique.

Les quantiles de régression. L'inconvénient de la régression classique est que l'estimation de la fonction de régression est sensible aux valeurs aberrantes et peut se montrer inappropriée dans certains cas, comme lorsque la distribution est multimodale ou fortement asymétrique. Ce manque de robustesse peut être résolu par la prévision en utilisant les quantiles conditionnels. Cet outil de prévision est obtenu en introduisant la fonction de perte pour une valeur fixe $p \in (0, 1)$:

$$\rho_p(t) = (2p - 1)t + |t|, \quad (2.8)$$

donc, le prédicteur est la solution du problème d'optimisation suivant :

$$\hat{r}(x) = \arg \min_{t \in \mathbb{R}} \mathbb{E}(\rho_p(Y - t) \mid X = x). \quad (2.9)$$

La robustesse de cette méthode provient du fait que la fonction de perte est convexe et bornée. Ces deux propriétés sont indispensables pour la robustesse d'un modèle, la première sert à assurer l'unicité de la

solution du problème d'optimisation (2.9), tandis que la deuxième est utilisée pour réduire l'influence des observations aberrantes sur la prévision.

L'expectiles de régression. L'expectile est apparu en 1987 pour la première fois dans le papier de [Newey and Powell \(1987\)](#) tandis que la régression expectile (RE) a été introduite dans la littérature, par [Aigner et al. \(1976\)](#), pratiquement aux mêmes moments que la régression quantile (RQ) (ou quantile conditionnel) qui a été introduite par [Koenker and Bassett \(1978\)](#). L'outil de prévision est aussi obtenu en introduisant la fonction de perte avec un paramètre d'asymétrie p :

$$\rho_p(z) = |p - \mathbf{1}_{z \leq 0}|z^2 \quad (2.10)$$

La RE généralise la régression classique de la moyenne conditionnelle, et constitue un bon modèle de risque alternatif à la régression quantile par le fait que la RE donne une solution unique, ce qui n'est pas le cas pour la RQ que seulement lorsque la fonction de répartition de la variable Y est absolument continue. De plus, la RE est plus facile à estimer alors que la RQ fait appel à la programmation linéaire. Malgré ces avantages, l'expectile ou l'expectile conditionnel reste inexploré par rapport aux deux régressions compétitives (la régression classique pour la moyenne conditionnelle et le quantile).

2.2.1 Estimation de la régression via la méthode du noyau

Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathcal{A} \times \mathbb{R}$, où (\mathcal{A}, d) un espace semi-métrique équipé d'une semi-métrique $d(\cdot, \cdot)$. On considère une suite $\{(X_i, Y_i)\}_{i=1, \dots, n}$ de variables aléatoires identiquement distribuées comme la variable aléatoire (X, Y) .

L'estimateur à noyau de Nadaraya-Watson est le premier estimateur rencontré dans la littérature (cf. [Nadaraya \(1964\)](#) et [Watson \(1964\)](#)), noté estimateur (NW), introduit pour la régression non paramétrique de dimension finie. Cet estimateur est composé d'une fonction de noyau $K(\cdot)$ et d'une fenêtre h_n , de manière analogue à l'estimateur à noyau de la fonction de densité introduit par Parzen-Rosenblatt (cf. [Parzen \(1962\)](#) et [Rosenblatt \(1956\)](#)), ensuite il est largement étudié depuis ce temps, on cite par exemple [Youndjé \(1993\)](#).

En se plaçant dans le modèle (2.7) avec le régresseur X est à valeurs dans un espace de dimension infinie et Y un réel. De leur côté, [Ferraty and Vieu \(2006\)](#) ont proposé une extension fonctionnelle de l'estimateur NW pour l'opérateur non linéaire $r(\cdot)$ de la régression défini comme suit :

$$\hat{r}_n(x) = \sum_{i=1}^n W_{ni}(x) Y_i, \quad (2.11)$$

où

$$W_{ni}(x) = \frac{K_i(x)}{\sum_{i=1}^n K_i(x)} \text{ et } K_i(x) = K(h^{-1}d(x, X_i)).$$

Les deux théorèmes suivants, présentés par [Ferraty and Vieu \(2006\)](#), et donnent respectivement la consistance et la vitesse presque complète de l'estimateur à noyau de la fonction de régression.

Théorème 2.1 (voir [Ferraty and Vieu \(2006\)](#)) *Sous certaines hypothèses décrites dans [Ferraty and Vieu \(2006\)](#) on a*

$$\lim_{n \rightarrow \infty} \hat{r}(x) = r(x) \text{ p.co.} \quad (2.12)$$

Théorème 2.2 (voir [Ferraty and Vieu \(2006\)](#)) Sous certaines hypothèses décrites dans [Ferraty and Vieu \(2006\)](#) on a

$$\widehat{r}(x) - r(x) = O\left(h_n^b\right) + O_{p.co.}\left(\sqrt{\frac{\log(n)}{n\phi(h_n)}}\right), \quad (2.13)$$

où $\phi(h_n)$ est la concentration de la mesure de probabilité de la variable fonctionnelle X dans la boule de centre x et de rayon h_n .

Dans le cas des observations α -mélangeantes, [Ferraty and Vieu \(2004\)](#) ont donné une nouvelle preuve des vitesses de convergence de certains types d'estimateurs à noyau (estimateur de NW) et l'ont exploité l'importance de la modélisation non-paramétrique des données fonctionnelles en appliquant leur étude à la discrimination des courbes et à la prévision.

Théorème 2.3 (voir [Ferraty and Vieu \(2004\)](#))

$$\sup_{x \in C} |\widehat{r}(x) - r(x)| = O\left(h_n^\beta\right) + O_{p.s.}\left(\sqrt{\frac{S_n \log(n)}{n\phi(h_n)}}\right). \quad (2.14)$$

La normalité asymptotique de l'estimateur de [Ferraty and Vieu \(2004\)](#) pour la fonction de régression a été montrée par [Masry \(2005\)](#).

Théorème 2.4 (voir [Masry \(2005\)](#))

$$\sqrt{(n\phi(h_n))}(\widehat{r}(x) - r(x) - B_n(x)) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x)), \quad (2.15)$$

où $\sigma^2(x) := \frac{C_2 g_2(x)}{C_1^2 f_1(x)}$. Avec $f_1(x)$ est la même fonction que nous avons supposé dans l'hypothèse B2 du chapitre 3 et $g_2(x) = \text{Var}(Y_j | X_j = x)$ est indépendante de l'indice j et est continue au voisinage de x . Pour ce qui concerne les constantes C_1 et C_2 sont spécifiées dans la condition 1' de [Masry \(2005\)](#).

Si l'on impose une hypothèse plus forte sur le paramètre de lissage h_n , alors on peut supprimer le terme de biais $B_n(x)$ du Théorème 2.4.

2.2.2 Modèle à indice fonctionnel simple (MIFS)

Le modèle à indice simple (MIS) est un modèle de régression semi-paramétrique, ce modèle a reçu une attention croissante dans la régression non paramétrique. En raison de ces avantages, il cumule les avantages des modèles de régression paramétrique et non paramétrique. Les références clés sur ce sujet dans les statistiques multivariées sont [Ichimura \(1993\)](#), [Hristache et al. \(2001\)](#) et [Delecroix et al. \(2003\)](#) pour les résultats précédents et [Huang \(2010\)](#), [Li et al. \(2011\)](#) et [Strzalkowska-Kominiak and Cao \(2013\)](#) pour les avancées et références récentes.

Cependant, dans la littérature de la statistique fonctionnelle, le modèle à indice fonctionnel simple est strictement limité dans le cas où les données sont fonctionnelles (une courbe). Le premier résultat dans ce contexte a été donné par [Ferraty et al. \(2003\)](#). Ils ont obtenu la convergence presque complète de la fonction

de régression $r(\cdot)$ dans le cas indépendant et identiquement distribué (i.i.d.). La généralisation de ce résultat au cas dépendant a été étudié par [Ait-Saïdi et al. \(2005\)](#) et [Masry \(2005\)](#). L'estimation des deux paramètres (non paramétrique $r(\cdot)$ et le coefficient linéaire θ) ont été traités par [Ait-Saïdi et al. \(2008\)](#). Ils ont proposé une estimation basée sur la procédure de la validation croisée. L'optimalité asymptotique de cette procédure est également étudiée.

De leur côté, [Mechab and Laksaci \(2016\)](#) ont établi la régression relative non-paramétrique pour les variables aléatoires associées, cette méthode est basée sur la minimisation de l'erreur relative quadratique moyenne. Plus récemment, [Douge \(2018\)](#) a étudié la normalité asymptotique de la fonction de régression sous des données quasi-associées lorsque la variable explicative prend ses valeurs dans un espace de Hilbert.

Soit $(X_i, Y_i)_{i=1, \dots, n}$ une suite de variables aléatoires distribuées selon la loi de (X, Y) , où la v.a. Y considérée comme une réponse scalaire et X une v.a.f. explicative, supposée à valeurs dans un espace hilbertien séparable H muni du produit scalaire $\langle \cdot, \cdot \rangle$. Le modèle à indice fonctionnel simple est défini par :

$$Y_i = r(\langle \theta, X_i \rangle) + \varepsilon_i, \forall i = 1, \dots, n, \quad (2.16)$$

où $r(\cdot)$ est la fonction réel, et pour $i = 1, \dots, n$, ε_i est la variable aléatoire réelle telle que $\mathbb{E}(\varepsilon_i | X_i) = 0$.

Sous la condition de l'indépendance, [Ferraty et al. \(2003\)](#) ont montré la consistance et la vitesse de convergence presque complète :

Théorème 2.5 (voir [Ferraty et al. \(2003\)](#)) *Sous des conditions de régularité, hypothèses sur les moments conditionnels de la variable aléatoire réponse Y , et sur le noyau K , le paramètre de lissage h_n et la probabilité des petites boules $G_\theta(x, h)$, on a*

$$\widehat{r}_\theta(x) - r_\theta(x) = O\left(h_n^\beta\right) + O_{p.co.}\left(\sqrt{\frac{\log(n)}{nG_\theta(x, h)}}\right). \quad (2.17)$$

[Douge \(2018\)](#) a étudié la normalité asymptotique de la fonction de régression sous des données quasi-associées lorsque la variable explicative prend ses valeurs dans un espace de Hilbert.

Théorème 2.6 (voir [Douge \(2018\)](#)) *Sous certaines hypothèses et pour $b > 0$, $\phi(h_n) = O(h_n^b)$ nous avons*

$$\sqrt{(n\phi(h_n))}(\widehat{r}_n(x) - r(x)) \xrightarrow{\mathcal{D}} N(0, \sigma_2^2(x)) \text{ quand } n \rightarrow \infty, \quad (2.18)$$

où $\sigma_2^2(x) := \frac{C_2}{C_1^2} \left(\frac{g_2(x) - r^2(x)}{f_1(x)} \right)$ avec la fonction g_2 est uniformément continue au voisinage de x et définie pour $x \in H$ par $g_2(x) := \mathbb{E}(\varphi^2(Y) | X_1 = x)$. La fonction $f_1(x)$ et les constantes (C_1 et C_2) ont été définies par [Douge \(2018\)](#) dans les hypothèses A1(i) et A2(iii) respectivement.

Les propriétés asymptotiques de cet estimateur ($\widehat{r}_n(\cdot)$) ont été étudiées par [Ferraty and Vieu \(2004\)](#) et [Masry \(2005\)](#) dans le cas de processus fortement mélangeant.

2.2.3 Quantiles et les expectiles conditionnels

Quantile conditionnel

Les quantiles conditionnels sont largement étudiés lorsque la variable explicative est à valeurs dans un espace de dimension finie et qu'il existe de nombreuses références sur ce sujet (voir, par exemple, [Samanta \(1989\)](#) pour les résultats précédents et [Gannoun *et al.* \(2003\)](#) pour les avancées et références récentes). Citons également [Zhou and Liang \(2000\)](#) et [Lin and Li \(2007\)](#), ils ont montré la normalité asymptotique dans le cas de la médiane conditionnelle (où $p = 0.5$) sous des conditions α -mélangeantes et de l'association, respectivement.

Cependant, dans le cas où la variable X est fonctionnelle, on peut citer par exemple, [Cardot *et al.* \(2004\)](#) qui ont utilisé une approche B-spline pour étudier le modèle linéaire de régression sur les quantiles avec une variable explicative à valeurs dans un espace de Hilbert. Les auteurs ont obtenu la convergence en L^2 de l'estimateur.

Dans le contexte non paramétrique, la convergence presque complète (p.co.) de l'estimateur à noyau du quantile conditionnel est établie dans [Ferraty *et al.* \(2006\)](#) lorsque les observations sont indépendantes et identiquement distribuées (i.i.d.) alors que le cas dépendant est traité dans [Ferraty *et al.* \(2005\)](#) avec application aux données climatologiques. La normalité asymptotique de cet estimateur a été étudiée dans les deux cas (i.i.d. et du mélange fort) par [Ezzahrioui and Ould Saïd \(2008\)](#). Quant à eux [Dabo-Niang and Laksaci \(2007\)](#), ont établi la convergence en norme L^p sous des conditions moins restrictives. Ils ont considéré la propriété des probabilités de petites boules de la loi de la variable explicative fonctionnelle et aussi la régularité de la densité conditionnelle, dans le cas i.i.d.

De leur côté, [He *et al.* \(2018\)](#) ont proposé un nouveau cadre pour estimer les quantiles conditionnels extrêmes avec une covariable fonctionnelle qui combine les techniques de modélisation non paramétrique et la théorie des valeurs extrêmes. Tous ces développements théoriques ont favorisé l'accessibilité et l'utilisation de la RQ. Pour une bonne revue de littérature sur le sujet, on peut consulter l'excellent livre de [Koenker \(2005\)](#) ou le plus récent manuel de [Koenker *et al.* \(2018\)](#).

Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathcal{A} \times \mathbb{R}$, où (\mathcal{A}, d) est un espace semi-métrique équipé d'une semi-métrique $d(\cdot, \cdot)$. On suppose qu'on a n couples $(X_i, Y_i)_{i=1, \dots, n}$ de variables aléatoires à valeurs dans $\mathcal{A} \times \mathbb{R}$ et qui sont identiquement distribuées comme (X, Y) .

Nous considérons, pour $p \in (0, 1)$, le quantile conditionnel d'ordre p de la distribution conditionnelle de Y sachant $X = x$ notée par $F_Y^x(\cdot)$

$$t_p(x) = \inf\{t \in \mathbb{R} : F_Y^x(t) \geq p\}.$$

Le quantile conditionnel $t_p(x)$ est défini comme une solution unique par rapport à t du problème d'optimisation (2.5) avec fonction de perte $\rho_p(t) = (2p - 1)t + |t|$.

Un estimateur naturel de $t_p(x)$ est alors

$$\hat{t}_p(x) = \arg \min_{t \in \mathbb{R}} \widehat{\psi}_p(x, t), \quad (2.19)$$

où $\psi_p(x, t) := \mathbb{E}(\rho_p(Y - t) | X = x)$ et $\widehat{\psi}_p(x, \cdot)$ est l'estimateur à noyau de Nadaraya-Watson de $\psi_p(x, \cdot)$

défini par

$$\widehat{\psi}_p(x, t) = \sum_{i=1}^n W_{ni}(x) \rho_p(Y_i - t), \quad \forall t \in \mathbb{R}$$

avec

$$W_{ni}(x) = \frac{K_i(x)}{\sum_{i=1}^n K_i(x)} \quad \text{et} \quad K_i(x) = K(h^{-1}d(x, X_i)),$$

l'estimation de la fonction de répartition conditionnelle a été traité comme une étude préliminaire de l'estimation des quantiles conditionnels par plusieurs auteurs, la normalité asymptotique de cet estimateur dans les deux cas (i.i.d. et α -mélangeant) a été étudié par [Ezzahrioui and Ould Saïd \(2005, 2006\)](#). Les quantiles conditionnels ont été proposés via une autre méthode d'estimation par [Laksaci et al. \(2009\)](#). Les résultats asymptotiques de cette étude est la consistance et la normalité asymptotique dans le cas i.i.d.

Proposition 2.1 (voir [Laksaci et al. \(2009\)](#)) *Sous certaines Hypothèses, nous avons*

$$\sup_{t \in [t_p(x) - \delta, t_p(x) + \delta]} |\widehat{F}_Y^x(t) - F_Y^x(t)| = O(h_n^b) + O_{p.co.} \left(\sqrt{\frac{\log(n)}{n\phi_x(h)}} \right). \quad (2.20)$$

Théorème 2.7 (voir [Laksaci et al. \(2009\)](#)) *Sous les hypothèses de la Proposition (2.1), et si la densité conditionnelle $f^x(t_p(x))$ est strictement positif, nous avons*

$$\widehat{t}_p(x) - t_p(x) = O(h_n^b) + O_{p.co.} \left(\sqrt{\frac{\log(n)}{n\phi_x(h)}} \right). \quad (2.21)$$

Théorème 2.8 (voir [Laksaci et al. \(2009\)](#)) *Sous des hypothèses sur la concentration $\phi_x(h)$, nous avons*

$$\sqrt{\frac{n\phi_x(h)}{\sigma^2(x)}} (\widehat{t}_p(x) - t_p(x)) \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{quand } n \rightarrow \infty, \quad (2.22)$$

où $\sigma^2(x) := \frac{p(1-p)a_2(x)}{(f_Y^x(t_p(x)))^2 a_1^2(x)}$ et $a_j(x) = K^j(1) - \int_0^1 (K^j)'(s) \beta_x(s) ds$ pour $j=1, 2$.

La fonction $\beta_x(s)$ est définie dans l'hypothèse $(H1')$ de [Laksaci et al. \(2009\)](#).

Dans le cas dépendant (α -mélangeantes), [Ferraty et al. \(2005\)](#) ont montré la consistance et la vitesse de convergence presque complète du quantile conditionnel.

Théorème 2.9 (voir [Ferraty et al. \(2005\)](#)) *Sous certaines hypothèses et si de plus*

$$\exists \varsigma > 0, \quad C n^{\frac{3-a}{a+1} + \varsigma} \leq \phi_x(h_K) \leq C' n^{\frac{1}{1-a}}$$

avec $a > \frac{5 + \sqrt{17}}{2}$, nous avons

$$\widehat{t}_p(x) - t_p(x) \xrightarrow{p.co.} 0 \quad \text{quand } n \rightarrow \infty.$$

Théorème 2.10 (voir [Ferraty et al. \(2005\)](#)) Sous certaines conditions, nous avons

$$\widehat{t}_p(x) - t_p(x) = O\left(h_K^{\frac{b_1}{j}} + h_H^{\frac{b_2}{j}}\right) + O_{p.co.}\left(\left(\frac{\log(n)}{n\phi_x(h)}\right)^{\frac{1}{2j}}\right).$$

La normalité asymptotique de cet estimateur a été étudiée dans les deux cas (i.i.d. et α -mélangeant) par [Ezzahrioui and Ould Saïd \(2008\)](#).

Théorème 2.11 (voir [Ezzahrioui and Ould Saïd \(2008\)](#)) Soit l'ordre unique p du quantile conditionnel tel que $p = F_Y^x(t_p(x)) = p = F_n^x(\widehat{t}_p(x))$. Soit $\Xi = \{x : x \text{ appartenant à un compact } S \text{ de } \mathbb{R} : \Sigma^2(x, t_p(x)) \neq 0\}$. Sous certaines hypothèses, nous avons, pour tout $x \in \Xi$,

$$\sqrt{\frac{n\phi_x(h)}{\Sigma^2(x, t_p(x))}} (\widehat{t}_p(x) - t_p(x)) \xrightarrow{\mathcal{D}} N(0, 1) \text{ quand } n \rightarrow \infty, \quad (2.23)$$

où $\Sigma(x, t_p(x)) := \sigma^2(x)$. Même expression de variance que celle définie dans [Laksaci et al. \(2009\)](#).

Expectile conditionnel

Malgré que la *régression quantile* (RQ) et la *régression expectile* (RE) ont été introduites dans la littérature pratiquement au même moment, la mesure de risque comme *quantile* est plus connu et le plus développé d'une manière assez rapide que la mesure de risque *expectile*. Par ailleurs, les quantiles ne sont pas toujours satisfaisants et peuvent être critiqués pour leur difficulté à calculer car la fonction de perte correspondante n'est pas partout différentiable, les méthodes d'estimation traditionnelles, par exemple l'algorithme de Gauss-Newton, ne s'appliquent plus pour générer les estimateurs de la RQ. Pour contourner ce problème, Koenker et Bassett ont reformulé le problème d'optimisation dans le cadre de la programmation linéaire, (Voir [Koenker and Bassett \(1978\)](#)). Le principal avantage de l'expectile sur le quantile est son efficacité, sa facilité de calcul et d'estimation.

Notons qu'il est facile de voir que les *expectiles conditionnels* sont caractérisés par les *espérances conditionnelles* de queue (tail conditionals expectations) de la même manière que les *quantiles conditionnels* sont caractérisés par la fonction de distribution conditionnelle, pour plus de détails, nous renvoyons le lecteur intéressé aux travaux de [Newey and Powell \(1987\)](#) et [Abdous et al. \(2003\)](#). De plus, l'*expectile conditionnel* considère la sensibilité aux valeurs aberrantes comme un avantage important, car si nous mesurons les pertes potentielles, nous voulons que notre mesure soit sensible aux valeurs extrêmes et aberrantes. Par conséquent, cet avantage permet une gestion des risques plus prudente et réactive.

Nous présentons ici la définition d'une mesure de risque dans un cadre général ainsi que certaines critères possibles pour le choix d'une mesure de risque.

Définition 2.3 (voir [Artzner et al. \(1999\)](#)) Une mesure de risque θ est une fonction définie sur l'espace des variables aléatoires, et prenant ses valeurs dans \mathbb{R} .

Les mesures de risque sont caractérisées par des propriétés qui reflètent des comportements économiques souhaitables. [Artzner et al. \(1999\)](#) ont énoncé quatre axiomes que toute mesure de risque utilisée devrait satisfaire pour une régulation et une gestion efficace du risque. Ces mesures de risque sont alors dites *cohérentes*.

Cette notion de cohérence des mesures de risque représente un point de repère dans le développement de la théorie du risque, elle est considérée aujourd'hui comme la référence en matière d'analyse des propriétés des mesures. La cohérence est regroupe donc certaines propriétés mathématiques et qui sont représentés comme des critères possibles pour le choix d'une mesure de risque.

- Une mesure de risque θ est *invariante par translation* si $\theta(Y + h) = \theta(Y) + h$ pour toute constante h .
- Homogénéité : Une mesure de risque θ est *homogène* si $\theta(hY) = h\theta(Y)$ pour toute constante positive h .
- Une mesure de risque θ est *sous-additive* si $\theta(Y_1 + Y_2) \leq \theta(Y_1) + \theta(Y_2)$ quels que soient les risques Y_1 et Y_2 .
- Une mesure de risque θ est *monotone* si $\mathbb{P}[Y_1 \leq Y_2] = 1 \Rightarrow \theta(Y_1) \leq \theta(Y_2)$ quels que soient les risques Y_1 et Y_2 .

Selon Artzner *et al.* (1999), une mesure de risque qui satisfait ces quatre propriétés est qualifiée de *cohérente*. D'autres propriétés non considérées comme des *axiomes de cohérences* sont aussi souhaitables et parfois naturelles. Nous donnons à titre d'exemples :

- ▷ *L'invariance par loi* : Pour un risque Y , la mesure $\theta(Y)$ ne dépend que de sa loi de probabilités. Pour tous risques X et Y de même loi, $\theta(X) = \theta(Y)$.
- ▷ *Additivité comonotone* : Pour tout risque X et toutes fonctions croissantes f et g , $\theta(f(X) + g(X)) = \theta(f(X)) + \theta(g(X))$.
L'additivité comonotonique est une autre propriété des mesures de risque qui présente principalement un intérêt complémentaire à la sous-additivité.
- ▷ *Convexité* : Pour tous risques X et Y définis sur le même espace de probabilité, pour toute constante $p \in [0, 1]$, $\theta(pX + (1 - p)Y) \leq p\theta(X) + (1 - p)\theta(Y)$.

On cite quelques mesures usuelles les plus utilisées dans les applications et qui sont présentées dans la littérature de la théorie du risque :

- ▷ *La Value at Risk* (VaR) de seuil p est le quantile d'ordre p de la distribution de X

$$VaR_p(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq p\}.$$

Si la variable aléatoire X est de distribution continue alors

$$VaR_p(X) = \inf\{x \in \mathbb{R} : F(x) \geq p\} = F^{-1}(p),$$

où F est la fonction de répartition de X . La VaR n'est pas une mesure cohérente car elle ne satisfait pas la propriété de la sous-additivité. Elle est comonotoniquement additive et vérifiée, pour toute fonction croissante et continue φ ,

$$VaR_p(\varphi(X)) = \varphi(VaR_p(X)).$$

En pratique, le principal inconvénient de la VaR est son incapacité à fournir l'information sur le comportement du risque au-delà de sa valeur.

- ▷ *La Tail Value at Risk* (TVaR) avec un niveau de confiance p est définie comme une moyenne des VaR dépassant la $VaR_p(X)$:

$$TVaR_p(X) = \frac{1}{1 - p} \int_p^1 VaR_\mu(X) d\mu.$$

La TVaR est une mesure cohérente pour tous les risques de distribution continue. Elle représente la plus petite mesure cohérente majorant la VaR, et elle fournit l'information sur le comportement du risque et l'épaisseur de la queue de sa distribution.

▷ La *Conditional Tail Expectation* (CTE) est définie par

$$CTE_p(X) = \mathbb{E}[X \mid X > VaR_p(X)],$$

la CTE et la TVaR coïncident pour les risques de distribution continue.

▷ La *Conditional VaR* (CVaR) est définie par

$$CVaR_p(X) = \mathbb{E}[X - VaR_p(X) \mid X > VaR_p(X)] = CTE_p(X) - VaR_p(X),$$

▷ L'*Expected Shortfall* (ES) est définie par

$$ES_p(X) = \mathbb{E}[(X - VaR_p(X))_+] = CVaR_p(X)P(X \mid X > VaR_p(X)).$$

La *variance* est aussi considérée comme une mesure de risque importante, utilisée en particulier dans l'allocation des actifs financiers. Ainsi, en assurance, *l'espérance et la variance* sont des mesures de risques limitées dans leur utilisation mais restent néanmoins utiles.

Nous nous appuyons, dans cette section, sur une littérature récente qui se focalise sur une nouvelle propriété désirable statistiquement pour toute mesure de risque. Une mesure de risque θ est dite *élicitable* si elle peut être écrite, pour toute variable aléatoire X , sous la forme d'une solution du problème d'optimisation d'une fonction de perte

$$\theta(X) = \arg \min_x \mathbb{E}(\rho(x, X)), \quad X \sim \mathbb{P}.$$

Cette propriété mathématique *d'élicitabilité*, satisfaite par certaines mesures de risque, introduite dans la théorie du risque par [Gneiting \(2011\)](#), est importante pour comparer les différentes méthodes d'estimations des mesures de risque. Elle donne un sens aux procédures de *backtesting*, notamment pour des tests comparatifs sur des modèles de risques et aussi aux performances de ces modèles, et qui sont nécessaires dans la gestion pratique des risques financiers et actuariels.

L'élicitabilité a fait l'objet de plusieurs articles, on peut citer par exemple [Ziegel \(2016\)](#), [Steinwart et al. \(2014\)](#) ont étudié cette propriété dans un cadre purement mathématique et analysée comme une mesure de risque par [Bellini and Bigozzi \(2015\)](#), [Wang and Ziegel \(2015\)](#), [Emmer et al. \(2015\)](#) et [Bellini et al. \(2014\)](#).

La VaR est une mesure élicitable pour tout seuil $p \in [0, 1]$, elle s'écrit sous la forme

$$VaR_p(X) = \arg \min_{x \in \mathbb{R}} \{\mathbb{E}[p(X - x)_+ + (1 - p)(x - X)_+]\},$$

tandis que la TVaR n'est pas une mesure élicitable. Ce constat a remis en question la notion de cohérence des mesures de risque, et a lancé une nouvelle discussion académique pour définir le choix de mesure qui répond le mieux aux besoins de la modélisation des risques. Une des réponses proposées est fondée sur la caractérisation des mesures de risque qui sont à la fois cohérentes et élicitables. Dans ce sens, les mesures *expectiles* sont apparues récemment comme une alternative souhaitable et défendable de la VaR et la TVaR.

Les *expectiles* ont été introduits par [Newey and Powell \(1987\)](#). Ils ont défini, cette nouvelle mesure de risque, pour une variable aléatoire X comme solution du problème d'optimisation suivant :

$$\theta_p(X) = \arg \min_{x \in \mathbb{R}} \{\mathbb{E}[p(X - x)_+^2 + (1 - p)(x - X)_+^2]\}, \quad (2.24)$$

où $(x)_+ = \max(x, 0)$ et p le seuil de confiance.

Plusieurs études ont été menées récemment sur les propriétés de l'expectile. Certains analysent les propriétés de cohérence et d'élicitabilité dans le contexte des mesures de risques, d'autres se consacrent à trouver une caractérisation des mesures à la fois *cohérente et élicitable*. [Bellini et al. \(2014\)](#) ont démontré que ces mesures sont *cohérentes et élicitables* pour un niveau de confiance $p \geq 1/2$, elles sont, selon [Bellini and Bignozzi \(2015\)](#), les seules mesures de risques, *invariantes par loi*, qui vérifient à la fois ces deux propriétés.

L'expectile conditionnel d'ordre p , pour $p \in (0, 1)$, ou p -expectile conditionnel, peut être définie de façon unique par

$$\theta(p; x) = \arg \min_{t \in \mathbb{R}} \mathbb{E}(\rho_p(Y, t) \mid X = x), \quad (2.25)$$

où $\rho_p(Y, t) = |p - \mathbb{1}_{(Y-t) \leq 0}|(Y-t)^2 = p(Y-t)^2 \mathbb{1}_{(Y-t) > 0} + (1-p)(Y-t)^2 \mathbb{1}_{(Y-t) \leq 0}$, qui représente la fonction de perte avec un paramètre d'asymétrie p .

La RE est construite en combinant les idées des modèles de régression à celles du quantile conditionnel permettant d'accumuler les avantages des deux modèles. Dans l'article [Newey and Powell \(1987\)](#), on retrouve une étude détaillée des propriétés de la fonction expectile. Notamment, le fait que la fonction expectile est invariante par translation affine. Ils ont présenté également la méthode de la RE pour le modèle linéaire simple et l'ont montré les propriétés asymptotiques de cette nouvelle classe d'estimateurs.

Après quelques années, tout en vantant ses propriétés analogues à celles des quantiles, [Efron \(1991\)](#) a illustré les bienfaits de l'utilisation de cette nouvelle classe d'estimateurs. Par la suite, peu ou pas d'études sur le modèle de la RE ont été publiées dans la littérature, probablement comme il a mentionné le papier de [Eilers \(2013\)](#), parce que les efforts étaient concentrés sur le développement du modèle de la RQ. Plusieurs caractéristiques et expressions existantes pour les quantiles ont leur analogue avec les expectiles. Par exemples, la fonction expectile caractérise la fonction de répartition de la variable aléatoire, de la même manière que la fonction quantile. En particulier, à chaque valeur de la fonction de répartition, on peut associer un seul expectile. De plus, à chaque quantile t_p , de niveau p , on peut trouver un expectile θ_{τ_p} de niveau τ_p tel que $t_p = \theta_{\tau_p}$. Quant à lui [Koenker \(1993\)](#), a montré que l'expectiles et les quantiles coïncident lorsque la fonction de répartition de de la variable aléatoire Y est définie par

$$F(y) = 1/2\{1 + \text{sign}(y)\sqrt{1 + 4/(4 + y^2)}\}, y \in \mathbb{R}.$$

Ce résultat est très utile dans des situations où l'estimation des quantiles est très longue et que son unicité n'est pas garantie, et est également vrai pour la classe des modèles de régression, comme le montre le papier de [Yao and Tong \(1996\)](#). D'ailleurs, [Yao and Tong \(1996\)](#) ont utilisé cette correspondance pour estimer, dans un cadre non-paramétrique, les quantiles conditionnels à partir de l'expectiles conditionnels. En plus, d'autres conditions ont été étudiées où les expectiles et les quantiles coïncident, c'est le cas de [Abdous and Rémillard \(1995\)](#), ont montré que l'expectile d'une distribution F est un quantile d'une distribution G .

Après les travaux de [Efron \(1991\)](#), [Koenker \(1993\)](#), [Abdous and Rémillard \(1995\)](#) et [Yao and Tong \(1996\)](#), le modèle de la RE est resté dans l'ombre pendant des décennies. Aujourd'hui, le sujet refait surface et intéresse les chercheurs comme le démontre le grand nombre de publications sur le sujet.

L'application de la RE a été considéré pour le modèle de spline et de lissage (voir [Schnabel and Eilers \(2009\)](#), [De Rossi and Harvey \(2009\)](#), [Sobotka and Kneib \(2012\)](#) et [Sobotka et al. \(2013\)](#)). Ainsi, sur le problème de la croissance humaine (croissance des courbes) et sur la RE (voir [Kneib \(2013\)](#), [Schnabel and Eilers \(2013\)](#) et [Waltrup et al. \(2015\)](#)). Récemment d'autres travaux ont été publiés, notamment sur les caractéristiques de l'expectile multivarié (voir par exemple [Maume-Deschamps et al. \(2017\)](#) et [Maume-Deschamps et al. \(2018\)](#)).

Finalement, la RE est combiné à de nombreux types de modèles, on cite par exemple : bayésiens (voir [Majumdar and Paul \(2016\)](#), [Waldmann et al. \(2017\)](#) et [Xing and Qian \(2017\)](#)), non-paramétriques (voir [Righi et al. \(2014\)](#) et [Yang and Zou \(2015\)](#)), non-linéaires (voir [Kim and Lee \(2016\)](#)).

2.2.4 Estimation de la régression via la méthode k -NN

La méthode des k plus proches voisins (k -NN) consiste à considérer les k variables X_i les plus proches de x au sens de la semi-métrique d . Cette méthode présente de nombreux avantages, principalement celui de respecter la structure locale des données, ce qui est primordial en dimension infinie.

Dans le cadre fonctionnel, on trouve [Laloë \(2008\)](#) qui a examiné cette méthode des k -NN lorsque la variable fonctionnelle est un élément d'un espace de Hilbert séparable H . La convergence faible a été établie par le même auteur en appliquant des techniques multivariées sur des données projetées pour effectuer la régression k -NN.

De leur côté, [Attouch et al. \(2017\)](#), [Chikr-Elmezouar et al. \(2019\)](#) et [Ling et al. \(2020\)](#) ont combinés entre la procédure k -NN et l'estimation linéaire locale fonctionnelle pour montrer la consistance et la vitesse de convergence. Quant à eux [Burba et al. \(2009\)](#), ont basé sur une approche purement fonctionnelle, pour des résultats récents en dimensions infinies (voir par exemple [Ferraty and Vieu \(2006\)](#), [Cérou, Frédéric and Guyader, Arnaud \(2006\)](#) et [Chen and Shah \(2018\)](#)). Dans ce sens [Burba et al. \(2009\)](#), ont étendu un lemme de [Collomb \(1980\)](#) en examinant le problème sur un espace fonctionnel semi-métrique. Ils ont prouvé la convergence presque complète de l'estimateur k -NN ainsi que sa vitesse pour des données indépendantes.

[Lian \(2011\)](#) a présenté les vitesses de convergence presque sûre de l'estimation des k plus proches voisins lorsque la réponse est un élément d'un espace de Hilbert. Les résultats les plus récents de l'analyse des données fonctionnelles non paramétriques des k -NN se trouvent dans [Kudraszow and Vieu \(2013\)](#) et [Kara et al. \(2017\)](#) et ses références.

Dans l'analyse des données fonctionnelles dépendantes, [Müller and Dippon \(2011\)](#) ont examiné la méthode des k plus proches voisins dans la régression fonctionnelle non paramétrique pour des données α -mélangeantes. Ils ont prouvé la convergence presque complète et la vitesse de l'estimateur à noyau des k -NN. Plus récemment, [Ling et al. \(2019\)](#) ont considéré le modèle de régression non paramétrique et l'ont établi la vitesse de convergence uniforme presque complète de l'estimateur k -NN pour les données chronologiques fonctionnelles et fortement mélangeantes.

L'estimateur à noyau des k plus proches voisins de r en x est défini de la façon suivante

$$\hat{r}_{k,\theta}(x, H_{n,k}) = \frac{\sum_{i=1}^n Y_i K_i(x)}{\sum_{i=1}^n K_i(x)} \quad \text{pour tout } x \in H, \quad (2.26)$$

où $K_i(x) := K\left(\frac{\langle \theta, x - X_i \rangle}{H_{n,k}}\right)$ est le noyau. Pour k fixé, le paramètre $H_{n,k}$ est une variable aléatoire de \mathcal{A} dans \mathbb{R}^+ qui dépend de (X_1, \dots, X_n) , et est définie de la manière suivante :

$$H_{n,k}(x) = \min \left\{ h_n \in \mathbb{R}^+ : \sum_{i=1}^n \mathbb{1}_{B(x, h_n)}(x_i) = k \right\}, \quad (2.27)$$

où $B(x, h_n)$ est la boule de centre x et de rayon $h_n > 0$ définie comme suit

$B(x, h_n) = \{x' \in H, d_\theta(x, x') \leq h_n\}$ et $\mathbb{1}_A$ est la fonction indicatrice de l'ensemble A .

Cette méthode de k -NN est largement étudiée, [Burba et al. \(2009\)](#) ont généralisé un lemme de [Collomb \(1980\)](#) et l'ont examiné le problème sur un espace fonctionnel semi-métrique. Ils ont prouvé la convergence presque complète et la vitesse de convergence presque complète pour des données indépendantes.

Théorème 2.12 (voir [Burba et al. \(2009\)](#)) *Sous le modèle de continuité, et des hypothèses sur la suite, $k = k_n$, de nombres réels positives : $\frac{k}{n} \rightarrow 0$ et $\frac{\log(n)}{k} \rightarrow 0$, nous avons*

$$\lim_{n \rightarrow \infty} \widehat{r}_k(x, H_{n,k}) = r(x) \text{ p.co.} \quad (2.28)$$

Théorème 2.13 (voir [Burba et al. \(2009\)](#)) *Sous le modèle de Lipschitz et les hypothèses sur $k : \frac{k}{n} \rightarrow 0$ et $\frac{\log(n)}{k} \rightarrow 0$, nous avons*

$$\widehat{r}_k(x, H_{n,k}) - r(x) = O\left(\phi_x^{-1}\left(\frac{k}{n}\right)^\beta\right) + O_{p.co.}\left(\sqrt{\frac{\log(n)}{k}}\right) \quad (2.29)$$

[Müller and Dippon \(2011\)](#) ont généralisé les résultats de [Burba et al. \(2009\)](#) au cas des données α -mélangeantes ; où ils ont supposé que la probabilité d'observer la variable aléatoire fonctionnelle explicative X autour de x est strictement positive (i.e. $\phi_x(h) > 0$), et ils ont supposé une condition plus générale sur la fonction de distribution conjointe. Soit alors, la quantité suivante

$$\psi(h) := \max_{i,j \in \{1, \dots, n\}, i \neq j} \mathbb{P}(X_i \in B(x, h), X_j \in B(x, h)),$$

tel que

$$\frac{\log(n)\chi(x, h)^{1-s}n^{1+s}}{k^2} \rightarrow 0 \text{ quand } n \rightarrow \infty, \quad (2.30)$$

où $\chi(x, h) := \max\{1, \frac{\psi(h)}{(\phi_x(h))^2}\}$ et $s := \frac{1}{b+1}$ avec b la vitesse du coefficient de mélange.

Théorème 2.14 (voir [Müller and Dippon \(2011\)](#)) *Sous le modèle de continuité, et les hypothèses sur la concentration $\phi_x(h)$, le paramètre k , le noyau K , nous avons*

$$\lim_{n \rightarrow \infty} \widehat{r}_k(x, H_{n,k}) = r(x) \text{ p.co.} \quad (2.31)$$

Théorème 2.15 (voir [Müller and Dippon \(2011\)](#)) *Sous le modèle de Lipschitz et sous certaines hypothèses, nous avons*

$$\widehat{r}_k(x, H_{n,k}) - r(x) = O\left(\phi_x^{-1}\left(\frac{k}{n}\right)^\beta\right) + O_{p.co.}\left(\sqrt{\frac{\log(n)}{k}}\right). \quad (2.32)$$

Sous le modèle de Lipschitz et de plus la condition (2.30) est satisfaite, nous avons

$$\begin{aligned} \widehat{r}_k(x, H_{n,k}) - r(x) &= O\left(\phi_x^{-1}\left(\frac{k}{n}\right)^\beta\right) + O_{p.co.}\left(\sqrt{\frac{\log(n)}{k}}\right) \\ &+ O_{p.co.}\left(\sqrt{\frac{n^{1+s} \log(n)}{k^2} \chi\left(x, \phi_x^{-1}\left(\frac{k}{n}\right)\right)^{1-s}}\right). \end{aligned} \quad (2.33)$$

Notons que, dans le Théorème 2.15, le terme de la covariance S_n^2 est disparu dans l'équation (2.32). Sous la condition (2.30), ils ont obtenu, comme le montre l'équation (2.33), trois termes pour la vitesse de convergence presque complète. Le premier terme de l'équation (2.33) trouve son origine dans la régularité de la fonction de régression, le second et le troisième terme représentent la dépendance des variables aléatoires.

L'idée principale dans la consistance uniforme, est de couvrir un sous-ensemble S de \mathcal{F} par un nombre fini de boules. Ferraty *et al.* (2010) ont introduit un concept topologique défini comme suit :

Définition 2.4 Soit S un sous-ensemble d'un espace semi-métrique \mathcal{F} , et soit $\varepsilon > 0$ donné. Un ensemble fini de points x_1, \dots, x_N dans \mathcal{F} est appelé ε -net pour S si $S \subset \bigcup_{k=1}^N B(x_k, \varepsilon)$. La quantité $\Psi_S(\varepsilon) = \log(N_\varepsilon(S))$ où $N_\varepsilon(S)$ est le nombre minimal de boules ouvertes dans \mathcal{F} de rayon ε qui est nécessaire pour couvrir S , est appelée ε -entropie de Kolmogorov de l'ensemble S .

La convergence uniforme d'un estimateur à noyau pour la fonction de régression dans le cas i.i.d. a été étudié par Kudraszow and Vieu (2013) :

Théorème 2.16 (voir Kudraszow and Vieu (2013)) Sous certaines conditions, nous avons

$$\sup_{x \in S} |\hat{r}_k(x, H_{n,k}) - r(x)| = O\left(\phi^{-1}\left(\frac{k}{n}\right)^b\right) + O_{p.co.}\left(\sqrt{\frac{\Psi_S\left(\frac{\log(n)}{n}\right)}{k}}\right) \quad (2.34)$$

Les résultats de Kudraszow and Vieu (2013) sont généralisés aux données α -mélangeantes par Ling *et al.* (2019) :

Théorème 2.17 (voir Ling *et al.* (2019)) Sous certaines conditions, nous avons

$$\sup_{x \in S} |\hat{r}_k(x, H_{n,k}) - r(x)| = O\left(\phi^{-1}\left(\frac{k}{n}\right)^\beta\right) + O_{p.co.}\left(\sqrt{\frac{S_n^2 \Psi_S\left(\frac{\log(n)}{n}\right)}{n^2}}\right) \quad (2.35)$$

Il est nécessaire de noter que le Théorème 2.17 de Ling *et al.* (2019) étend le Théorème 2.16 de Kudraszow and Vieu (2013) au cas de données α -mélangeantes. Il étend également les résultats de Müller and Dippon (2011) en établissant la convergence uniforme avec vitesse presque complète, et l'ont montré aussi le comportement asymptotique de l'estimateur k -NN de la régression $\hat{r}_k(x, H_{n,k})$ pour les données fonctionnelles α -mélangeantes. Deuxièmement, la vitesse de convergence presque complète de l'estimateur est divisée en deux parties. L'une ne dépend que de la régularité de l'opérateur de régression et du paramètre de lissage k , et l'autre montre l'effet de mélange fort et de l'entropie de Kolmogorov $\Psi_S(\cdot)$.

Bibliographie

- Abdous, B. and Rémillard, B. (1995). Relating quantiles and expectiles under weighted-symmetry. *Ann. Inst. Statist. Math.*, **47**(2), 371–384.
- Abdous, B., Ghoudi, K., and Rémillard, B. (2003). Nonparametric weighted symmetry tests. *Canad. J. Statist.*, **31**(4), 357–381.
- Aigner, D. J., Amemiya, T., and Poirier, D. J. (1976). On the estimation of production frontiers : maximum likelihood estimation of the parameters of a discontinuous density function. *Internat. Econom. Rev.*, **17**(2), 377–396.
- Ait-Saïdi, A., Ferraty, F., Kassa, R., and Vieu, P. (2008). Cross-validated estimations in the single-functional index model. *Statistics*, **42**(6), 475–494.
- Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, **9**(3), 203–228.
- Attouch, M., Laksaci, A., and Ould-Saïd, E. (2009). Asymptotic distribution of robust estimator for functional nonparametric models. *Communications in Statistics - Theory and Methods*, **38**(8), 1317–1335.
- Attouch, M., Laksaci, A., and Rafea, F. (2017). Estimation locale linéaire de la régression non paramétrique fonctionnelle par la méthode des k plus proches voisins. *Comptes Rendus Mathématique*, **355**(7), 824 – 829.
- Aurzada, F. and Simon, T. (2007). Small ball probabilities for stable convolutions. <http://dx.doi.org/10.1051/ps :2007022>, **11**.
- Azzedine, N., Laksaci, A., and Ould-Saïd, E. (2008). On robust nonparametric regression estimation for a functional regressor. *Statistics & Probability Letters*, **78**(18), 3216 – 3221.
- Barrientos-Marin, J., Ferraty, F., and Vieu, P. (2010). Locally modelled regression and functional data. *Journal of Nonparametric Statistics*, **22**(5), 617–632.
- Bellini, F. and Bernardino, E. D. (2017). Risk management with expectiles. *The European Journal of Finance*, **23**(6), 487–506.
- Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quant. Finance*, **15**(5), 725–733.
- Bellini, F., Klar, B., Müller, A., and Rosazza Gianin, E. (2014). Generalized quantiles as risk measures. *Insur. Math. Econ.*, **54**, 41–48.

- Bellini, F., Bignozzi, V., and Puccetti, G. (2018). Conditional expectiles, time consistency and mixture convexity properties. *Insurance Math. Econom.*, **82**, 117–123.
- Benhenni, K., Ferraty, F., Rachdi, M., and Vieu, P. (2007). Local smoothing regression with functional data. *Computational Statistics*, **22**, 353–369.
- Benhenni, K., Hedli-Griche, S., and Rachdi, M. (2010). Estimation of the regression operator from functional fixed-design with correlated errors. *Journal of Multivariate Analysis*, **101**, 476–490.
- Besse, P. and Ramsay, J. (1986). Principal components analysis of sampled functions. *Psychometrika*, **51**, 285–311.
- Besse, P., Cardot, H., and Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations. *Scandinavian Journal of Statistics*, **27**(4), 673–687.
- Besse, P. C., Cardot, H., and Ferraty, F. (1997). Simultaneous non-parametric regressions of unbalanced longitudinal data. *Computational Statistics & Data Analysis*, **24**(3), 255 – 270.
- Blanquero, R., Carrizosa, E., Jiménez-Cordero, A., and Martín-Barragan, B. (2018). Variable selection in classification for multivariate functional data. *Information Sciences*.
- Bosq, D. (1991). *Modelization, Nonparametric Estimation and Prediction for Continuous Time Processes*, pages 509–529. Springer Netherlands, Dordrecht.
- Bosq, D. (2002). Estimation of mean and covariance operator of autoregressive processes in banach spaces. *Statistical Inference for Stochastic Processes*, **5**, 287–306.
- Bosq, D. (2012). *Linear Processes in Function Spaces : Theory and Applications*. Lecture Notes in Statistics. Springer New York.
- Bosq, D. and Lecoutre, J. (1987). *Théorie de l'estimation fonctionnelle*. Collection Economie et statistiques avancées. Série Ecole nationale de la statistique et de l'administration et du Centre d'études des programmes économiques. Économica.
- Breiman, L. and Friedman, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, **80**(391), 580–598.
- Burba, F., Ferraty, F., and Vieu, P. (2009). k-nearest neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics*, **21**(4), 453–469.
- Cardot, H., Crambes, C., and Sarda, P. (2004). Estimation spline de quantiles conditionnels pour variables explicatives fonctionnelles. *C. R. Math. Acad. Sci. Paris*, **339**(2), 141–144.
- Cardot, H., Crambes, C., and Sarda, P. (2005). Quantile regression when the covariates are functions. *Journal of Nonparametric Statistics*, **17**(7), 841–856.
- Cérou, Frédéric and Guyader, Arnaud (2006). Nearest neighbor classification in infinite dimension. *ESAIM : PS*, **10**, 340–355.
- Chen, G. H. and Shah, D. (2018). Explaining the success of nearest neighbor methods in prediction. *Foundations and Trends in Machine Learning*, **10**(5-6), 337–588.

- Chikr-Elmezouar, Z., Almanjahie, I. M., Laksaci, A., and Rachdi, M. (2019). Fda : strong consistency of the knn local linear estimation of the functional conditional density and mode. *Journal of Nonparametric Statistics*, **31**(1), 175–195.
- Collomb, G. (1980). Estimation de la régression par la méthode des k points les plus proches avec noyau : quelques propriétés de convergence ponctuelle. In *Nonparametric asymptotic statistics (Proc. Conf., Rouen, 1979) (French)*, volume 821 of *Lecture Notes in Math.*, pages 159–175. Springer, Berlin.
- Crambes, C., Delsol, L., and Laksaci, A. (2008). Robust nonparametric estimation for functional data. *J. Nonparametr. Stat.*, **20**(7), 573–598.
- Crambes, C., Kneip, A., and Sarda, P. (2009). Smoothing splines estimators for functional linear regression. *Ann. Statist.*, **37**(1), 35–72.
- Cuevas, A. and Fraiman, R. (2004). On the bootstrap methodology for functional data.
- Cuevas, A., Febrero-Bande, M., and Fraiman, R. (2002). Linear functional regression : The case of fixed design and functional response. *Canadian Journal of Statistics*, **30**, 285 – 300.
- Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique de mode conditionnel pour variable explicative fonctionnelle. *C. R. Math. Acad. Sci. Paris*, **344**(1), 49–52.
- Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Communications in Statistics - Theory and Methods*, **41**(7), 1254–1268.
- Dabo-Niang, S. and Rhomari, N. (2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes Rendus Mathématique*, **336**(1), 75 – 80.
- Daouia, A., Girard, S., and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc., Ser. B, Stat. Methodol.*, **80**(2), 263–292.
- Dauxois, J. and Pousse, A. (1976). *Les analyses factorielles en calcul des probabilités et en statistique : Essai d'étude synthétique*.
- De Rossi, G. and Harvey, A. (2009). Quantiles, expectiles and splines. *J. Econometrics*, **152**(2), 179–185.
- Delecroix, M., Härdle, W., and Hristache, M. (2003). Efficient estimation in conditional single-index regression. *Journal of Multivariate Analysis*, **86**(2), 213–226.
- Delsol, L. (2007). Clt and lq errors in nonparametric functional regression. *Comptes Rendus Mathématique*, **345**(7), 411 – 414.
- Delsol, L. (2008). *Régression sur variable fonctionnelle : Estimation, Test de Structure et Applications*. Thèse de doctorat, Université de Toulouse 3 Paul Sabatier, Toulouse (France).
- Deville, J.-C. (1974). Méthodes statistiques et numériques de l'analyse harmonique. *Annales de l'INSEE*, (15), 3–101.
- Devroye, L. (1978). The uniform convergence of nearest neighbor regression function estimators and their application in optimization. *IEEE Trans. Inf. Theory*, **24**, 142–151.

- Diaconis, P. and Shahshahani, M. (1984). On nonlinear functions of linear combinations. *SIAM Journal on Scientific and Statistical Computing*, **5**(1), 175–191.
- Douge, L. (2010). Théorèmes limites pour des variables quasi-associées hilbertiennes. *Ann. I.S.U.P.*, **54**(1-2), 51–60.
- Douge, L. (2018). Nonparametric regression estimation for quasi-associated hilbertian processes.
- Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statist. Sinica*, **1**(1), 93–125.
- Ehm, W., Gneiting, T., Jordan, A., and Krüger, F. (2016). Of quantiles and expectiles : consistent scoring functions, choquet representations and forecast rankings. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, **78**(3), 505–562.
- Eilers, P. H. C. (2013). Discussion : the beauty of expectiles [mr3179527]. *Stat. Model.*, **13**(4), 317–322.
- Emmer, S., Kratz, M., and Tasche, D. (2015). What is the best risk measure in practice ? a comparison of standard measures. *Journal of Risk*, **18** :2, 31–60.
- Ezzahrioui, M. and Ould Saïd, E. (2005). Asymptotic normality of nonparametric estimators of the conditional mode function for functional data. Technical Report 249, Technical report, LMPA.
- Ezzahrioui, M. and Ould Saïd, E. (2006). Asymptotic results of the kernel estimator of the conditional quantile in the normed space under α -mixing hypothesis. Technical Report 277, Technical report, LMPA.
- Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *Journal of Nonparametric Statistics*, **20**(1), 3–18.
- Ezzahrioui, M. and Ould Saïd, E. (2008). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East Journal of Theoretical Statistics*, **25**.
- Ezzahrioui, M. and Ould-Saïd, E. (2008a). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Communications in Statistics - Theory and Methods*, **37**(17), 2735–2759.
- Ezzahrioui, M. and Ould-Saïd, E. (2008b). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Comm. Statist. Theory Methods*, **37**(16-17), 2735–2759.
- Ferraty, F. (2010). Special issue : Statistical methods and problems in infinite-dimensional spaces. selected papers based on the presentations at the meeting, toulouse, france, june 19–21, 2008. *Journal of Multivariate Analysis*, **101**(2), 305–490.
- Ferraty, F. and Romain, Y. (2011). *The Oxford Handbook of Functional Data Analysis*.
- Ferraty, F. and Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, **330**(2), 139 – 142.
- Ferraty, F. and Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist.*, **17**(4), 545–564.
- Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. *Journal of Nonparametric Statistics*, **16**(1-2), 111–125.

- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Goia, A., and Vieu, P. (2002). Functional nonparametric model for time series : A fractal approach for dimension reduction. *TEST : An Official Journal of the Spanish Society of Statistics and Operations Research*, **11**, 317–344.
- Ferraty, F., Peuch, A., and Vieu, P. (2003). Modèle à indice fonctionnel simple. *C. R. Math. Acad. Sci. Paris*, **336**(12), 1025–1028.
- Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic " el niño " phenomenon. *Sankhya : The Indian Journal of Statistics*, **67**.
- Ferraty, F., Laksaci, A., and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, **9**, 47–76.
- Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data : inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**(3), 267–286.
- Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference*, **140**(2), 335–352.
- Gannoun, A., Saracco, J., and Yu, K. (2003). Nonparametric prediction by conditional median and quantiles. *J. Statist. Plann. Inference*, **117**(2), 207–223.
- Gao, F. and Li, W. V. (2007). Small ball probabilities for the slepian gaussian fields. *Transactions of the American Mathematical Society*, **359**(3), 1339–1350.
- Gasser, T., Hall, P., and 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.
- Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.*, **106**(494), 746–762.
- Györfi, L., Kohler, M., Krzyzak, A., and Walk, H. (2002). A distribution-free theory of non-parametric regression.
- Hall, P. and Vial, C. (2006). Assessing the finite dimensionality of functional data. *Journal of the Royal Statistical Society : Series B*, **68**(4), pp. 689–705.
- Hall, P., Poskitt, D. S., and Presnell, B. (2001). A functional data—analytic approach to signal discrimination. *Technometrics*, **43**(1), 1–9.
- Hastie, T. and Tibshirani, R. (1986). Generalized additive models. *Statist. Sci.*, **1**(3), 297–310.
- Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.
- He, F. Y., Cheng, Y. B., and Tong, T. J. (2018). Nonparametric estimation of extreme conditional quantiles with functional covariate. *Acta Math. Sin. (Engl. Ser.)*, **34**(10), 1589–1610.
- Hristache, M., Juditsky, A., and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Ann. Statist.*, **29**(3), 595–623.

- Huang, Z. (2010). Empirical likelihood for single-index varying-coefficient models with right-censored data. *J. Korean Statist. Soc.*, **39**(4), 533–544.
- Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J. Econometrics*, **58**(1-2), 71–120.
- Jank, W. and Shmueli, G. (2006). Functional data analysis in electronic commerce research. *Statist. Sci.*, **21**(2), 155–166.
- Kara, L.-Z., Laksaci, A., Rachdi, M., and Vieu, P. (2017). Data-driven knn estimation in nonparametric functional data analysis. *Journal of Multivariate Analysis*, **153**, 176 – 188.
- Kawasaki, Y. and Ando, T. (2004). Functional data analysis of the dynamics of yield curves. Physica-Verlag/Springer.
- Kim, M. and Lee, S. (2016). Nonlinear expectile regression with application to value-at-risk and expected shortfall estimation. *Computational Statistics & Data Analysis*, **94**, 1 – 19.
- Kneib, T. (2013). Beyond mean regression. *Stat. Model.*, **13**(4), 275–303.
- Koenker, R. (1993). When are expectiles percentiles? *Econometric Theory*, **9**(3), 526–527.
- Koenker, R. (2005). *Quantile regression*, volume 38 of *Econometric Society Monographs*. Cambridge University Press, Cambridge.
- Koenker, R. and Bassett, Jr., G. (1978). Regression quantiles. *Econometrica*, **46**(1), 33–50.
- Koenker, R., Chernozhukov, V., He, X., and Peng, L., editors (2018). *Handbook of quantile regression*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, Boca Raton, FL.
- Kuan, C.-M., Yeh, J.-H., and Hsu, Y.-C. (2009). Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*, **150**(2), 261 – 270. Recent Development in Financial Econometrics.
- Kudraszow, N. L. and Vieu, P. (2013). Uniform consistency of k NN regressors for functional variables. *Statist. Probab. Lett.*, **83**(8), 1863–1870.
- Laksaci, A., Lemdani, M., and Ould Saïd, E. (2009). A generalized l1-approach for a kernel estimator of conditional quantile with functional regressors : Consistency and asymptotic normality. *Statistics & Probability Letters*, **79**, 1065–1073.
- Laksaci, A., Lemdani, M., and Ould-Saïd, E. (2011). Asymptotic results for an L^1 -norm kernel estimator of the conditional quantile for functional dependent data with application to climatology. *Sankhya A*, **73**(1), 125–141.
- Laloë, T. (2008). A k-nearest neighbor approach for functional regression. *Statistics & Probability Letters*, **78**(10), 1189 – 1193.
- Leng, X. and Müller, H.-G. (2006). Time ordering of gene coexpression. *Biostatistics*, **7**(4), 569–584.
- Leurgans, S. E., Moyeed, R. A., and Silverman, B. W. (1993). Canonical correlation analysis when the data are curves. *Journal of the Royal Statistical Society. Series B (Methodological)*, **55**(3), 725–740.

- Li, K.-C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, **86**(414), 316–327.
- Li, T.-T., Yang, H., Wang, J.-L., Xue, L.-G., and Zhu, L.-X. (2011). Estimation for a partial-linear single-index model [correction on mr2589322]. *Ann. Statist.*, **39**(6), 3441–3443.
- Lian, H. (2011). Convergence of functional k -nearest neighbor regression estimate with functional responses. *Electron. J. Stat.*, **5**, 31–40.
- Lifshits, M. and Simon, T. (2005). Small deviations for fractional stable processes. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, **41**(4), 725 – 752.
- Lifshits, M. A., Linde, W., and Shi, Z. (2006). Small deviations of riemann-liouville processes in lq spaces with respect to fractal measures. *Proceedings of the London Mathematical Society*, **92**(1), 224–250.
- Lin, Z. and Li, D. (2007). Asymptotic normality for L_1 -norm kernel estimator of conditional median under association dependence. *J. Multivariate Anal.*, **98**(6), 1214–1230.
- Ling, N. and Vieu, P. (2018). Nonparametric modelling for functional data : selected survey and tracks for future. *Statistics*, **52**(4), 934–949.
- Ling, N., Meng, S., and Vieu, P. (2019). Uniform consistency rate of knn regression estimation for functional time series data. *Journal of Nonparametric Statistics*, **31**(2), 451–468.
- Ling, N., Aneiros, G., and Vieu, P. (2020). knn estimation in functional partial linear modeling. *Statistical Papers*, **61**(1), 423–444.
- Lucero, J. (1999). Computation of the harmonics-to-noise ratio of a voice signal using a functional data analysis algorithm. *Journal of Sound and Vibration*, **222**(3), 512 – 520.
- Mack, Y. P. (1981). Local properties of k-nn regression estimates. *SIAM Journal on Algebraic Discrete Methods*, **2**(3), 311–323.
- Majumdar, A. and Paul, D. (2016). Zero expectile processes and Bayesian spatial regression. *J. Comput. Graph. Statist.*, **25**(3), 727–747.
- Manteiga, W. G. and Vieu, P. (2007). Statistics for functional data. *Computational Statistics & Data Analysis*, **51**(10), 4788 – 4792.
- Martin-Barragan, B., Lillo, R., and Romo, J. (2014). Interpretable support vector machines for functional data. *European Journal of Operational Research*, **232**, 146–155.
- Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stochastic Process. Appl.*, **115**(1), 155–177.
- Maume-Deschamps, V., Rullière, D., and Said, K. (2017). Multivariate extensions of expectiles risk measures. *Depend. Model.*, **5**(1), 20–44.
- Maume-Deschamps, V., Rullière, D., and Said, K. (2018). Asymptotics multivariate expectiles. working paper or preprint.

- Mechab, W. and Laksaci, A. (2016). Nonparametric relative regression for associated random variables. *Metron*, **74**(1), 75–97.
- Müller, H.-G., Sen, R., and Stadtmueller, U. (2011). Functional data analysis for volatility. *Journal of Econometrics*, **165**, 233–245.
- Müller, S. and Dippon, J. (2011). k-nn kernel estimate for nonparametric functional regression in time series analysis.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, **9**(1), 141–142.
- Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, **55**(4), 819–847.
- Padoan, S. A. and Stupfler, G. (2020). Extreme expectile estimation for heavy-tailed time series. working paper or preprint.
- Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**(3), 1065–1076.
- Preda, C. (2007). Regression models for functional data by reproducing kernel hilbert spaces methods. *Journal of Statistical Planning and Inference*, **137**, 829–840.
- Rachdi, M. and Vieu, P. (2007). Nonparametric regression for functional data : Automatic smoothing parameter selection. *Journal of Statistical Planning and Inference*, **137**(9), 2784 – 2801.
- Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, **47**, 379–396.
- Ramsay, J. O. and Silverman, B. W. (1997). *Functional data analysis*. Springer Series in Statistics. Springer, New York, NY, first edition.
- Ramsay, J. O. and Silverman, B. W. (2002). *Applied functional data analysis : Methods and case studies*. Springer Series in Statistics. Springer, New York, NY, second edition.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer Series in Statistics. Springer, New York, second edition.
- Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, **14**(1), 1–17.
- Righi, M., Yang, Y., and Ceretta, P. (2014). *Nonparametric Expectile Regression for Conditional Autoregressive Expected Shortfall Estimation*, volume 96, pages 83–95.
- Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, **27**(3), 832–837.
- Rossi, F. and Villa, N. (2008). Recent advances in the use of svm for functional data classification. In *Functional and Operatorial Statistics*, pages 273–280, Heidelberg. Physica-Verlag HD.
- Royall, R. M. (1966). A class of nonparametric estimators of a smooth regression function.
- Saidi, A. A., Ferraty, F., and Kassa, R. (2005). Single functional index model for a time series. *Rev. Roumaine Math. Pures Appl.*, **50**(4), 321–330.

- Samanta, M. (1989). Nonparametric estimation of conditional quantiles. *Statist. Probab. Lett.*, **7**(5), 407–412.
- Schnabel, S. K. and Eilers, P. H. (2013). A location-scale model for non-crossing expectile curves. *Stat*, **2**(1), 171–183.
- Schnabel, S. K. and Eilers, P. H. C. (2009). Optimal expectile smoothing. *Comput. Statist. Data Anal.*, **53**(12), 4168–4177.
- Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression. *Comput. Statist. Data Anal.*, **56**(4), 755–767.
- Sobotka, F., Kauermann, G., Schulze Waltrup, L., and Kneib, T. (2013). On confidence intervals for semi-parametric expectile regression. *Stat. Comput.*, **23**(2), 135–148.
- Steinwart, I., Pasin, C., Williamson, R., and Zhang, S. (2014). Elicitation and identification of properties. *Journal of Machine Learning Research*, **35**, 482–526.
- Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.*, **8**(6), 1348–1360.
- Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.*, **10**(4), 1040–1053.
- Stone, C. J. (1983). Optimal uniform rate of convergence for nonparametric estimators of a density function or its derivatives. In *Recent Advances in Statistics*, pages 393 – 406. Academic Press.
- Stone, C. J. (1985). Additive regression and other nonparametric models. *Ann. Statist.*, **13**(2), 689–705.
- Strzalkowska-Kominiak, E. and Cao, R. (2013). Maximum likelihood estimation for conditional distribution single-index models under censoring. *J. Multivariate Anal.*, **114**, 74–98.
- Tucker, L. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19–23.
- Waldmann, E., Sobotka, F., and Kneib, T. (2017). Bayesian regularisation in geoadditive expectile regression. *Stat. Comput.*, **27**(6), 1539–1553.
- Waltrup, L., Otto-Sobotka, F., Kneib, T., and Kauermann, G. (2015). Expectile and quantile regression—david and goliath? *Statistical Modelling*, **15**, 433–456.
- Wang, R. and Ziegel, J. F. (2015). Elicitable distortion risk measures : a concise proof. *Statist. Probab. Lett.*, **100**, 172–175.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhyā : The Indian Journal of Statistics, Series A (1961-2002)*, **26**(4), 359–372.
- Xing, J.-J. and Qian, X.-Y. (2017). Bayesian expectile regression with asymmetric normal distribution. *Communications in Statistics - Theory and Methods*, **46**(9), 4545–4555.
- Yang, Y. and Zou, H. (2015). Nonparametric multiple expectile regression via er-boost. *Journal of Statistical Computation and Simulation*, **85**(7), 1442–1458.

-
- Yao, Q. and Tong, H. (1996). Asymmetric least squares regression estimation : a nonparametric approach. *J. Nonparametr. Statist.*, **6**(2-3), 273–292.
- Youndjé, E. (1993). *Estimation non paramétrique de la densité conditionnelle par la méthode du noyau*. Thèse de doctorat en sciences et techniques communes, Université de Rouen, Rouen (France).
- Zhou, Y. and Liang, H. (2000). Asymptotic normality for L_1 norm kernel estimator of conditional median under α -mixing dependence. *J. Multivariate Anal.*, **73**(1), 136–154.
- Ziegel, J. F. (2016). Coherence and elicibility. *Math. Finance*, **26**(4), 901–918.

Chapitre 3

Modèle à indice fonctionnel simple (MIFS) pour les données de séries chronologiques quasi-associées.

Dans ce chapitre, nous considérons le problème de l'estimation non paramétrique de la fonction de régression sous la condition de dépendance de quasi-association et dans un modèle à indice fonctionnel simple (MIFS). L'apport principal de ce chapitre est la caractérisation des propriétés asymptotiques de l'estimateur, telles que la convergence presque complète avec précision de la vitesse de convergence, la normalité asymptotique est obtenue sous certaines conditions standards et une discussion sur l'applicabilité de nos résultats pour construire les intervalles de confiance.

Le chapitre fait l'objet d'un article soumis dans une revue internationale.

Single index regression model for functional quasi-associated times series data.

Salim Bouzebda¹, Ali Laksaci² and Mustapha Mohammedi³

¹Alliance Sorbonne Université, Université de Technologie de Compiègne,
L.M.A.C., Compiègne, France

²Department of Mathematics, College of Science,
King Khalid University, Abha, 61413, Saudi Arabia

³Université Djillali Liabès, BP 89, 22000, Sidi Bel Abbès, Algérie,
L.S.P.S., Sidi Bel Abbès, Algérie

Abstract

The mixing condition is often considered to modeling the functional time series data. Alternatively, in this work we consider the problem of nonparametric estimation of the regression function in single functional model under the quasi-association dependence condition. The main result of this work is the establishment of the asymptotic properties of the estimator, such as the almost complete convergence rates. Furthermore, the asymptotic normality of the constructed are obtained under some mild conditions and we discuss how to apply our result to construct the confidence intervals.

Key words : Almost Complete Convergence, Asymptotic Normality, Functional Hilbert space, Kernel regression estimation, Mixing, Quasi-associated variables, Single functional index model, Weak dependence.

AMS subject classification : 62G05 ; 62G08 ; 62L12 ; 62G20.

3.1 Introduction

The statistical study of single index models have been studied and developped by several authors from a pratical and theorical point of view. The case of a vector explanatory variable was studied by [Härdle *et al.* \(1993\)](#) and [Hristache *et al.* \(2001\)](#). In the work of [Ferraty *et al.* \(2003\)](#) propose an extension of the single index model to the functional framework whose explanatory random variables have values in space of infinite dimension (Hilbertian or generally functional) called single functional index model. These models of single index are very popular in the econometric communaity because it respond two importants preoccupations. The first concerns dimension reduction since this type of model makes it possible to provide a solution to the problem of the curse of dimensionality. The second is related to the interpretability of the index θ introduced in these models. The model in which we consider, in this paper, is between the simple index model and the functional regression model. Moreover, this model can also be considered as a generalization of the functional linear model (see [Cardot *et al.* \(2003\)](#)). For this topics, [Chen *et al.* \(2011\)](#) introduced a new technique to estimate the link function in nonparametric framework. They proposed a multi-index modeling using an adaptive linear projections approach for functional data, and show that their method makes it possible to

predict with polynomial convergence rates. More recently, [Ding et al. \(2017\)](#) generalizes the functional linear regression model to a class of functional partially linear single index models. Thus, proposes a least squares approach with the same convergence rates (polynomial) than the previous technique cited in [Chen et al. \(2011\)](#) to estimate the slope and the link function.

Therefore, the single functional index model (SFIM) accumulate the advantages of single index model, and inherits the potential of the functional linear model in terms of applications (see [Ramsay and Silverman \(2005\)](#)) and that of the non parametric regression functional model (see [Ferraty and Vieu \(2002\)](#), [Ferraty and Vieu \(2006\)](#) and [Ferraty et al. \(2011\)](#)).

The modelization of functional data, has been developed intensively. The motivation of such statistical analysis is justified by the recent technological development of the measuring instruments that offers the opportunity to observe phenomena in an increasingly accurate way, but this accuracy obviously generates a large amount of data observed over a finer grid, which can be considered as observations varying over a continuum. The most theoretical results are obtained under independence condition. However, in practice, it is rarely that we have an independent identically distributed observations and of functional nature. The functional time series presents the more realistic situation. Thus it is really crucial to study the functional statistical models when the usual independence condition on the statistical sample is relaxed. In this paper, we consider the problem of the nonparametric estimation of the regression function in single functional index model when the data are weakly dependant.

Usually the dependence structure is molded with the strong mixing hypothesis, in this paper we focus in some more general correlation, that is the quasi-associated condition. The latter has been introduced for real valued random fields by [Bulinski and Suquet \(2001\)](#), which generalizes the positively associated variables introduced in [Esary et al. \(1967\)](#).

From practical point of view, this kind of data has great importance in practice, in particular, in reliability theory, mathematical physics and in percolation theory (see , for instance [Richard E. Barlow \(1975\)](#)) for more discussion in the practical interest of these random variables. Moreover, in theoretical point of view, the concept of quasi-association correlation can be viewed as a particular case of the weak dependence condition for real-valued stochastic processes introduced by [Doukhan and Louhichi \(1999\)](#) which allows treating the mixing condition and association correlation in a unified approach.

Noting that, the single index model is a semi-parametric regression model, thus, it couples the advantages of both parametric and nonparametric regression models. Because of these advantages, it has received an increasing amount of attention in the nonparametric regression literature. Key references on this topic in multivariate statistic are [Ichimura \(1993\)](#), [Hristache et al. \(2001\)](#) and [Delecroix et al. \(2003\)](#) for previous results and [Huang \(2010\)](#), [Li et al. \(2011\)](#), and [Strzalkowska-Kominiak and Cao \(2013\)](#) for recent advances and references.

However, in the literature of functional statistic, the single functional index model is strictly limited in the case where the data is functional (a curve). The first result in this context, was given by [Ferraty et al. \(2003\)](#). They obtained the almost complete convergence, of the regression function $r(\cdot)$ in the independent and identically distributed (i.i.d.) case. The generalization of this result to the dependent case has been studied by [Ait-Saïdi et al. \(2005\)](#) and [Masry \(2005\)](#). The estimation of both parameters (non parametric $r(\cdot)$ and linear coefficient θ) has treated by [Ait-Saïdi et al. \(2008\)](#). They proposed an estimation based on the cross-validation procedure. The asymptotic optimality of this procedure is also studied in this article. [Shang \(2018\)](#) uses a Bayesian method to estimate the bandwidths in the kernel form error density and regression function, under an autoregressive error structure, and according to empirical studies, the author considered

that the single functional index model gives improved estimation and prediction accuracies compared to any nonparametric functional regression considered.

The concept of quasi-association for random variables taking its values in a Hilbert space has been introduced by Douge (2010), and obtained some limit theorems for this type of variables. Attaoui *et al.* (2011) considered the conditional density estimation in functional index model. They obtained the almost complete convergence of their estimation in this context of functional single index model. The asymptotic normality of this last model when the observations are dependent was obtained by Ling and Xu (2012), and the asymptotic results for an M -estimator of the regression function for quasi-associated processes were obtained by Attaoui *et al.* (2015). Recently, Attaoui and Ling (2016) studied the asymptotic normality and uniform almost complete convergence rates of the conditional cumulative distribution estimator for dependent variables and by single index structure. The estimate of second derivative of the link function in a single functional index model are proposed by the local quadratic approximation in Ye and Hooker (2018). In Mechab and Laksaci (2016) established the non parametric relative regression for associated random variables, this method is based on minimization of the mean squared relative error. More recently, Douge (2018) studied the asymptotic normality of regression function under quasi-associated data when the explanatory variable takes its values in a Hilbert space.

The main purpose of the present paper is to establish the asymptotic properties of the estimator $\hat{r}_\theta(\cdot)$, when the variables are functional quasi-associated and in single index structure, such as the almost complete convergence rates. Furthermore, the asymptotic distribution is obtained under some mild conditions.

We point out that the mixing and the association concern two distinct classes of processes but not disjoint and offer two complementary approaches to study the dependence. Moreover, the functional quasi-associated data analysis has great importance in various domains such as the reliability theory or the statistical mechanics. Furthermore, it should be noted that the dependence condition considered here allow to avoid the widely used strong mixing condition which is very easy to verified in practice.

The rest of this work is organized as follows. In Section 3.2, we describe the single index regression model for functional data and in the quasi-associated framework, the next section is devoted to the introduction of the notation and hypotheses needed to state our main results. In Section 3.4, we will establish our main results of the almost complete convergence of the kernel estimators and the asymptotic normality under non restrictive conditions. In Section 3.4.3, we discuss an application of our results for the construction of the confidence interval. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the Section 3.5.

3.2 Model and estimator

We start by giving a definition of quasi-association adapted to the functional framework. In the real valued random fields, Bulinski and Suquet (2001) define the quasi-association dependence in the Definition 3.1 and it adapted to functional random variables in the Definition 3.2 by Douge (2010) as follows.

Definition 3.1 *A sequence $(X_n)_{n \in \mathbb{N}}$ of r.v.'s. is said to be quasi-associated, if for any disjoint subsets I and J of \mathbb{N} and all bounded Lipschitz functions $f_1 : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ and $f_2 : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ satisfying :*

$$|\text{Cov}(f_1(X_i, i \in I), f_2(X_j, j \in J))| \leq \text{Lip}(f_1) \text{Lip}(f_2) \sum_{i \in I} \sum_{j \in J} |\text{Cov}(X_i, X_j)|, \quad (3.1)$$

where $|I|$ denotes cardinality of a finite set I , and the Lipschitz of a function $f(\cdot)$ defined by

$$\text{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1}, \quad \text{with } \|(x_1, \dots, x_k)\|_1 = \sum_{k=1}^n |x_k|.$$

Definition 3.2 A sequence $(X_i)_{i \in \mathbb{N}}$ of r.v.'s. taking values in a Hilbert space H is called quasi-associated relative to an orthonormal basis $\{e_p : p \geq 1\}$ of H , if for any $p \geq 1$, $(\langle X_i, e_1 \rangle, \dots, \langle X_i, e_p \rangle)_{i \in \mathbb{N}}$ is a sequence of random vectors quasi-associates.

Now, we consider a sequence of quasi-associated random variables $\{(X_i, Y_i)\}_{i \in \mathbb{N}}$ identically distributed as (X, Y) , which are valued in $H \times \mathbb{R}$ where H is a separable real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and a orthonormal basis $\{e_p : p \geq 1\}$. The semi-metric $d_\theta(\cdot, \cdot)$ associated to the single-index $\theta \in H$ defined by $\forall u, v \in H$:

$$d_\theta(u, v) := |\langle \theta, u - v \rangle|.$$

The purpose of this paper is to study the estimation of the nonparametric regression of Y given $\langle \theta, X \rangle$ structure, denoted by

$$r(\langle \theta, X_i = x \rangle) = \mathbb{E}(Y \mid \langle \theta, X_i = x \rangle). \quad (3.2)$$

Such structure suppose that the explanation of Y from X is done through an fixed functional index θ in Θ . Now, we suppose that exists a $\theta \in \Theta \subset H$ where the observations $(X_i, Y_i)_{i=1, \dots, n}$ are related by the following relation

$$Y_i = r(\langle \theta, X_i \rangle) + \varepsilon_i, \quad \forall i = 1, \dots, n \quad (3.3)$$

where $r(\cdot)$ is a real function, and for $i = 1, \dots, n$, ε_i is a real random variable such that $\mathbb{E}(\varepsilon_i \mid X_i) = 0$. We consider that the single functional index model is identifiable, i.e., if the regression function is differentiable and if $\langle \theta, e_1 \rangle = 1$, where e_1 is the first element of an orthonormal basis of H . Then, if $r_1(\langle \theta_1, x \rangle) = r_2(\langle \theta_2, x \rangle)$ implies that $r_1 \equiv r_2$ and $\theta_1 \equiv \theta_2$. This hypothesis that we consider is demonstrated by [Ferraty et al. \(2003\)](#) once we have the differentiability of the regression operator r . For more details on the problem of identifiability of the single functional index model, one can refer to [Ferraty et al. \(2003\)](#) and [Ait-Saïdi et al. \(2008\)](#). The kernel estimator $\hat{r}_\theta(\cdot, h_n)$ of regression operator $r_\theta(\cdot) = r(\langle \theta, \cdot \rangle)$ is defined by

$$\hat{r}_\theta(x, h_n) = \frac{\sum_{i=1}^n Y_i K_i(x)}{\sum_{i=1}^n K_i(x)}, \quad \text{for all } x \in H, \quad (3.4)$$

where $K_i(x) := K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right)$ kernel function and h_n is the bandwidth parameter decreases to zero as n goes to infinity.

3.3 Assumptions and notation

In the sequel, we will denote by C_i with $i = 1, 2, \dots$ some strictly positive generic constants.

(H₁) $K(\cdot)$ is a bounded continuous Lipschitz function such that

$$C_1 \mathbb{1}_{[0,1]}(t) < K(t) < C_2 \mathbb{1}_{[0,1]}(t);$$

(H₂) Let $E_i(x) := \langle \theta, x - X_i \rangle$ so that $E_i(x)$ is a real-valued random variable,

$$G_\theta(x, h_n) = \mathbb{P}(|E_i(x)| \leq h_n) > 0;$$

(H₃) Let $B(x, h_n)$ be a ball centered at $x \in H$ with radius h_n . The joint distribution functions satisfy

$$\begin{aligned} 0 &< \sup_{i \neq j} \mathbb{P}\left((X_i, X_j) \in B(x, h_n) \times B(x, h_n)\right) \\ &= \max_{i \neq j} \left\{ \mathbb{P}\left(|E_i| \leq h_n\right), \mathbb{P}\left(|E_j| \leq h_n\right) \right\} = O(G_\theta^2(x, h_n)); \end{aligned}$$

(H₄) For all $u, v \in H$ for some $\beta > 0$,

$$|r_\theta(u) - r_\theta(v)| \leq C_3 |\langle \theta, u - v \rangle|^\beta;$$

(H₅) We have

- (i) $\mathbb{E}(\exp(|Y|)) \leq C_4 < \infty$,
- (ii) $\mathbb{E}(|Y|^\nu | X = x) \leq C_5 < \infty$, a.s. for all $\nu \geq 2$,
- (iii) $\forall i \neq j \mathbb{E}(|Y_i Y_j| | X_i, X_j) \leq C_6 < \infty$, a.s. ;

(H₆) Denote λ_r the covariance coefficient defined as :

$$\lambda_r := \sup_{s \geq r} \sum_{|i-j| \geq s} \lambda_{i,j},$$

where

$$\begin{aligned} \lambda_{i,j} &= \sum_{k \geq 1} \sum_{l \geq 1} \left| \text{Cov}(X_i^k, X_j^l) \right| \\ &\quad + \sum_{k \geq 1} \left| \text{Cov}(X_i^k, Y_j) \right| + \sum_{l \geq 1} \left| \text{Cov}(Y_i, X_j^l) \right| + |\text{Cov}(Y_i, Y_j)|, \end{aligned}$$

and $X_i^p := \langle X_i, e_p \rangle$. We assume that the covariance coefficient λ_r , $r \in \mathbb{N}$ satisfying :

$$\exists a > 0, \exists C_7 > 0, \text{ such that } \lambda_r \leq C_7 e^{-ar};$$

(H₇) There exist $\gamma \in (0, 1)$ and $\xi_1, \xi_2 > 0$ such that

$$\frac{\log n}{n^{1-\gamma-\xi_2}} \leq G_\theta(x, h_n) \leq \frac{C}{(\log n)^{1+\xi_1}};$$

(B₁) The function $\varphi(z) := \mathbb{E}(Y_1^2 | X_1 = z)$, $z \in H$ exists and is uniformly continuous in some neighborhood of z .

(B₂) There exists a differentiable function $\psi(\cdot)$ such that for $x \in H$

$$0 < C_8\psi(z)f_1(x) \leq G_\theta(x, z) \leq C_9\psi(z)f_1(x) \text{ where } \psi(z) \longrightarrow 0 \text{ as } z \longrightarrow 0$$

and

$$\lim_{z \rightarrow 0} \frac{G_\theta(x, z)}{\psi(z)} =: f_1(x), \text{ with } f_1 \text{ is a probability density of } X_1.$$

We assume, as in Gasser *et al.* (1998), that the asymptotic value of the integral

$$I_j(h_n) = \frac{1}{\psi(h_n)} \int_0^{h_n} K^j\left(\frac{z}{h_n}\right) \psi'(z) dz \longrightarrow C_j, \text{ as } h_n \longrightarrow 0. \quad (3.5)$$

for some positive constants $C_j, j = 1, 2$.

Some comments on the assumptions

Assumptions (H₁), (B₁) are standard and technical conditions on the kernel $K(\cdot)$ and the function $\varphi(\cdot)$. The concentration property of the explanatory variable in small balls under single index topological structure is defined in the assumption (H₂). Assumption (H₃) gives the behavior of the joint distribution of $(E_i(x), E_j(x))$ and to establish the asymptotic normality of our model in the quasi-associated variables setting. The hypothesis (H₄) is used to control the regularity condition which characterizes the functional space of our model and is needed to evaluate the bias in the asymptotic results of this paper. Assumptions (H₅ (i),(ii),(iii)) are a mild smoothness assumptions on the regression functional. The hypothesis (H₆) used for the quasi-associated data, (H₇) is also classical in the functional estimation. We need the condition (B₂) which is similar to the assumption considered in Gasser *et al.* (1998) and the equation (3.5) in (B₂) is used to obtain an explicite expression of the asymptotic variance.

3.4 Main result

3.4.1 The almost consistency

Our aim is to establish the almost complete convergence (a.co.)¹ of $\widehat{r}_\theta(x)$ to $r_\theta(x)$, the main result is given by the following Theorem.

Theorem 3.1 *Under the assumptions (H₁)-(H₇), we have, as $n \rightarrow \infty$,*

$$\widehat{r}_\theta(x, h_n) - r_\theta(x) = O_{a.co.} \left(h_n^\beta + \sqrt{\frac{\log n}{n^{1-\gamma} G_\theta(x, h_n)}} \right). \quad (3.6)$$

1. We say that the sequence $(\Theta_n)_n$ converges a.co. to zero, if and only if

$$\forall \tau > 0, \sum_{n \geq 1} \mathbb{P}(|\Theta_n| > \tau) < \infty.$$

Furthermore, we say that $\Theta_n = O_{a.co.}(\theta_n)$, if there exists $\tau_0 > 0$, such that

$$\sum_{n \geq 1} \mathbb{P}(|\Theta_n| > \tau_0 \theta_n) < \infty.$$

Let

$$\widehat{r}_{\theta,0}(x) := \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n K_i(x) \quad \text{and} \quad \widehat{r}_{\theta,1}(x) := \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n Y_i K_i(x). \quad (3.7)$$

Let us consider the following decomposition :

$$\begin{aligned} \widehat{r}_{\theta}(x, h_n) - r_{\theta}(x) &= \frac{\widehat{r}_{\theta,1}(x)}{\widehat{r}_{\theta,0}(x)} - r_{\theta}(x) \\ &= \frac{1}{\widehat{r}_{\theta,0}(x)} [(\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1})) - (r_{\theta}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x)))] \\ &\quad - \frac{r_{\theta}(x)}{\widehat{r}_{\theta,0}(x)} (\widehat{r}_{\theta,0}(x) - 1) \\ &= \frac{1}{\widehat{r}_{\theta,0}(x)} [(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) + (\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x)))] \\ &\quad + \frac{1}{\widehat{r}_{\theta,0}(x)} [(\mathbb{E}(\widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) - (r_{\theta}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x)))] \\ &\quad - \frac{r_{\theta}(x)}{\widehat{r}_{\theta,0}(x)} (\widehat{r}_{\theta,0}(x) - 1), \end{aligned}$$

where

$$\widehat{r}_{\theta,2}(x) := \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n \widehat{Y}_i K_i(x) \quad (3.8)$$

The real variable Y response which is not necessarily bounded. For this, we introduce the truncated random variable \widehat{Y} defined by $\widehat{Y}_i = Y_i \mathbb{1}_{\{|Y_i| \leq \delta_n\}}$ with $\delta_n = n^{\gamma/6}$. The proof of the Theorem 3.1 is based on the following Lemmas. We start by a Bernstein-type inequality for weakly dependent random variables introduced by [Kallabis and Neumann \(2006\)](#) in Theorem 2.1.

Lemme 3.1 (See [Kallabis and Neumann \(2006\)](#)) *Let X_1, \dots, X_n the real random variables such that $\mathbb{E}X_i = 0$ and $\mathbb{P}(|X_i| \leq M) = 1$, for all $i = 1, \dots, n$ and some $M < \infty$. Let $\sigma_n^2 = \text{Var}(\sum_{i=1}^n X_i)$. Assume, furthermore, that there exist $K < \infty$ and $\beta > 0$ such that, for all u -tuplets (s_1, \dots, s_u) and all v -tuplets (t_1, \dots, t_v) with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$, the following inequality is fulfilled :*

$$\left| \text{Cov}(X_{s_1} \dots X_{s_u}, X_{t_1} \dots X_{t_v}) \right| \leq K^2 M^{u+v-2} v e^{-\beta(t_1 - s_u)}.$$

Then,

$$\mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| \geq t\right) \leq \exp\left\{-\frac{t^2/2}{A_n + B_n^{\frac{1}{3}} t^{\frac{5}{3}}}\right\},$$

for $A_n \leq \sigma_n^2$ and

$$B_n = \left(\frac{16nK^2}{9A_n(1 - e^{-\beta})} \vee 1\right) \left(\frac{2(K \vee M)}{1 - e^{-\beta}}\right).$$

Lemme 3.2 *Under the assumptions (H_1) - (H_7) , we have, as $n \rightarrow \infty$,*

$$|\widehat{r}_{\theta,2}(x) - \mathbb{E}\widehat{r}_{\theta,2}(x)| = O_{a.co.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} G_{\theta}(x, h_n)}} \right). \quad (3.9)$$

Lemme 3.3 Under the assumptions (H_1) - (H_6) , we have, as $n \rightarrow \infty$,

$$|\widehat{r}_{\theta,0}(x) - 1| = O_{a.co.} \left(\sqrt{\frac{\log n}{nG_{\theta}(x, h_n)}} \right). \quad (3.10)$$

Lemme 3.4 Under the assumptions of Lemma 3.3, we have, as $n \rightarrow \infty$,

$$\exists \eta > 0 \text{ such that } \sum_{i=1}^n \mathbb{P}(|\widehat{r}_{\theta,0}(x)| < \eta) < \infty. \quad (3.11)$$

Lemme 3.5 Under the assumptions (H_1) - (H_4) , we have, as $n \rightarrow \infty$,

$$|r_{\theta}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))| = O(h_n^{\beta}). \quad (3.12)$$

Lemme 3.6 Under the assumptions (H_1) , (H_2) , (H_3) , (H_5) , and H_7 we have, as $n \rightarrow \infty$,

$$|\mathbb{E}(\widehat{r}_{\theta,2}) - \mathbb{E}(\widehat{r}_{\theta,1}(x))| = O \left(\sqrt{\frac{\log n}{n^{1-\gamma}G_{\theta}(x, h_n)}} \right). \quad (3.13)$$

Lemme 3.7 Under the assumptions (H_5) and (H_7) , we have, as $n \rightarrow \infty$,

$$|\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)| = O_{a.co.} \left(\sqrt{\frac{\log n}{n^{1-\gamma}G_{\theta}(x, h_n)}} \right). \quad (3.14)$$

3.4.2 The asymptotic normality

Now, we study the asymptotic normality of $\widehat{r}_{\theta}(x)$. The main result of this section is summarized in the following Theorem.

Theorem 3.2 Under the assumptions (H_1) - (H_6) , (B_1) - (B_2) and in addition we assume that $\psi(h_n) = O(h_n^b)$ with $b > 0$, $n\psi^{1+2c}(h_n) \rightarrow \infty$ and $\log^2(n)\psi^{-(1+\frac{2}{b})}(h_n) \rightarrow 0$. We have, as $n \rightarrow \infty$,

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) \xrightarrow{\mathcal{D}} {}^2N(0, \sigma_{\theta,1}^2(x)), \quad (3.15)$$

and

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta,1}(x) - r_{\theta}(x)\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - r_{\theta}(x)\widehat{r}_{\theta,0}(x))) \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,2}^2(x)), \quad (3.16)$$

where $\sigma_{\theta,1}^2(x) = \frac{C_2}{C_1^2} \frac{\varphi(x)}{f_1(x)}$ and $\sigma_{\theta,2}^2(x) = \frac{C_2}{C_1^2} \frac{(\varphi(x) - r_{\theta}^2(x))}{f_1(x)}$ with $\varphi(z) := \mathbb{E}(Y_1^2 | X_1 = z)$ and $f_1(\cdot)$ is a probability density of X_1 such that $f_1(x) > 0$.

2. $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution

Consider the following decomposition :

$$\begin{aligned}
 \widehat{r}_\theta(x, h_n) - r_\theta(x) &= \frac{\widehat{r}_{\theta,1}(x)}{\widehat{r}_{\theta,0}(x)} - r_\theta(x) \\
 &= [\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x)] - \mathbb{E}[\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x)] \\
 &\quad - (r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) + (r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) (\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x))) \\
 &\quad - (\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) (\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x))) + \widehat{r}_\theta(x) (\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x)))^2 \\
 &= J_1 + J_2 + J_3 + J_4 + J_5,
 \end{aligned} \tag{3.17}$$

where $\widehat{r}_{\theta,i}(x)$ for $i = 0, 1, 2$ and \widehat{Y}_i (with $\delta_n = \delta_0 \log(n)$, $\delta_0 > 0$) are defined in the previous section.

Corollary 3.1 *Under the hypotheses of Theorem 3.2 and in addition we assume that $nh_n^{2\beta}\psi(h_n) \rightarrow 0$ we have, as $n \rightarrow \infty$,*

$$\sqrt{n\psi(h_n)} (\widehat{r}_\theta(x, h_n) - r_\theta(x)) \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,2}^2(x)). \tag{3.18}$$

Lemme 3.8 *Under the hypotheses of Theorem 3.2 and in addition we assume that $n^{2\delta_0}\psi(h_n) \rightarrow \infty$, we have, as $n \rightarrow \infty$,*

$$n\psi(h_n)\text{Var}(\widehat{r}_{\theta,1}(x)) = \sigma_{\theta,1}^2(x) + o(1), \tag{3.19}$$

and

$$n\psi(h_n)\text{Var}(\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x)) = \sigma_{\theta,2}^2(x) + o(1). \tag{3.20}$$

3.4.3 Application to the confidence intervals

The purpose of a confidence interval is to supplement the functional estimate at a point with information about the uncertainty in this estimate. It is a direct application of the Central Limit Theorem (CLT). In order to provide a confidence interval for the regression function in single functional model, we need first to propose a consistent estimator of the variance $\sigma_{\theta,2}^2(x)$. A natural consistent estimator of this variance is then

$$\widehat{\sigma}_{\theta,2}^2(x) = \frac{\widehat{C}_2 (\widehat{\varphi}_n(x) - \widehat{r}_\theta^2(x))}{\widehat{C}_1^2 \widehat{f}_n(x)}, \tag{3.21}$$

where we recall that the coefficients $(C_j)_{j=1,2}$ are given in equation (3.5). We have now to give estimates of $(C_j)_{j=1,2}$ and $\varphi(\cdot)$. Since $r_\theta(x)$ and $f_1(x)$ are estimated by $\widehat{r}_\theta(x, h_n)$ and $\widehat{f}_{n,\theta}(x)$ respectively, then C_j can be estimated empirically as follows :

$$\widehat{C}_j = \frac{1}{nG_\theta(x, h)} \sum_{i=1}^n K^j \left(\frac{\langle \theta, x - X_i \rangle}{h_n} \right), \quad j = 1, 2, \tag{3.22}$$

where

$$G_\theta(x, h_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{|\langle \theta, x - X_i \rangle| < h_n\}}.$$

We use the Nadaraya-Watson type estimator to estimate $\varphi(\cdot)$, $r_\theta(x)$ and $f_1(x)$ by $\widehat{\varphi}_n(x)$, $\widehat{r}_\theta(x, h_n)$ and $\widehat{f}_{n,\theta}(x)$ respectively, then we obtain a consistent estimator of $\sigma_{\theta,2}^2(x)$:

$$\frac{\widehat{C}_2}{\widehat{C}_1^2} \frac{(\widehat{\varphi}_n(x) - \widehat{r}_\theta^2(x, h_n))}{\widehat{f}_n(x)},$$

with

$$\widehat{\varphi}_n(x) = \frac{\sum_{i=1}^n Y_i^2 K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right)}, \quad (3.23)$$

and

$$\widehat{f}_{n,\theta}(x) = \frac{1}{n\mathbb{E}(K_1(x))} \sum_{i=1}^n K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right). \quad (3.24)$$

Consequently, by combining the equations (3.22), (3.23) and (3.24), we obtain the consistent estimator of $\widehat{\sigma}_{\theta,2}^2(x)$ given by (3.21). An application of Theorem 3.2 and Corollary 3.1 gives that

$$\mathbb{P} \left[-\mathcal{U}_{1-\frac{\alpha}{2}} \leq \left(\frac{n\psi(h_n)}{\sigma_{\theta,2}^2(x)} \right)^{\frac{1}{2}} (\widehat{r}_\theta(x, h_n) - r_\theta(x)) \leq +\mathcal{U}_{1-\frac{\alpha}{2}} \right] = 1 - \alpha \quad (3.25)$$

where $(1 - \alpha)$ is the confidence level and $\mathcal{U}_{1-\frac{\alpha}{2}}$ denotes the quantile of order $(1 - \frac{\alpha}{2})$ of the standard normal law. The asymptotic confidence band at asymptotic level $1 - \alpha$ for $r_\theta(x)$ is provided by

$$\widehat{r}_\theta(x, h_n) \pm \mathcal{U}_{1-\frac{\alpha}{2}} \left(\frac{\sigma_{\theta,2}^2(x)}{n\psi(h_n)} \right)^{\frac{1}{2}}.$$

The problem is that $\sigma_{\theta,2}^2(x)$ is unknown, and the best we can do is to replace $\sigma_{\theta,2}^2(x)$ by our consistent estimator $\widehat{\sigma}_{\theta,2}^2(x)$. Finally, we have the confidence interval

$$\widehat{r}_\theta(x, h_n) \pm \mathcal{U}_{1-\frac{\alpha}{2}} \left(\frac{\widehat{\sigma}_{\theta,2}^2(x)}{n\psi(h_n)} \right)^{\frac{1}{2}}. \quad (3.26)$$

3.5 Proofs of the intermediate results

Proof of Lemma 3.1

See the proof of this Lemma in [Kallabis and Neumann \(2006\)](#).

Proof of Lemma 3.2

The proof of this Lemma is based on inequality given in Lemma 3.1 on the variables

$$\widehat{\Delta}_i(x) := \frac{1}{n\mathbb{E}(K_1(x))} \left[\widehat{Z}_i - \mathbb{E}(\widehat{Z}_i) \right], \quad i = 1, \dots, n,$$

where $\widehat{Z}_i = \widehat{Y}_i K_i(x)$ and $\widehat{Y}_i = Y_i \mathbf{1}_{\{|Y_i| \leq \delta_n\}}$, we have

$$\begin{aligned} \mathbb{E}(\widehat{\Delta}_i(x)) &= 0, \\ \|\widehat{\Delta}_i(x)\|_\infty &\leq \frac{2\delta_n}{nG_\theta(x, h_n)} \|K\|_\infty, \\ \text{Lip}(\widehat{\Delta}_i(x)) &\leq 2\text{Lip}(K) \frac{\delta_n}{nG_\theta(x, h_n)h_n}, \\ \widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x)) &= \sum_{i=1}^n \widehat{\Delta}_i(x). \end{aligned}$$

We start by evaluating the covariance term $\text{Cov}(\widehat{\Delta}_{s_1}, \dots, \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1}, \dots, \widehat{\Delta}_{t_v})$, for all $(s_1, \dots, s_u) \in \mathbb{N}^u$ and $(t_1, \dots, t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$. If $m = t_1 - s_u = 0$, using the fact that $\mathbb{E}(K_1) = O(G_\theta(x, h_n))$ and under (H_5) , we readily obtain

$$\begin{aligned} \left| \text{Cov}(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v}) \right| &\leq \left(\frac{1}{n\mathbb{E}(K_1)} \right)^{u+v} \mathbb{E} \left(\left| \widehat{Z}_{s_1} \dots \widehat{Z}_{s_u}^2, \dots, \widehat{Z}_{t_v} \right| \right) \\ &\leq \left(\frac{C\delta_n \|K\|_\infty}{nG_\theta(x, h_n)} \right)^{u+v} \mathbb{E}(Y_{s_u}^2 K_{s_u}^2) \\ &\leq \left(\frac{C\delta_n}{nG_\theta(x, h_n)} \right)^{u+v} G_\theta(x, h_n). \end{aligned}$$

If $m = t_1 - s_u > 0$, by quasi-association of the sequence (\widehat{Z}_n) in connection with (H_6) , we infer that

$$\begin{aligned} \left| \text{Cov}(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v}) \right| &\leq 4 \left(\frac{\delta_n \text{Lip}(K)}{nG_\theta(x, h_n)h_n} \right)^2 \left(\frac{2\delta_n \|K\|_\infty}{nG_\theta(x, h_n)} \right)^{u+v-2} \sum_{i=1}^u \sum_{j=1}^v \lambda_{s_i, t_j} \\ &\leq C^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \left(\frac{\delta_n}{nG_\theta(x, h_n)} \right)^{u+v} (u \wedge v) \lambda_{t_1 - s_u} \\ &\leq C^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \left(\frac{\delta_n}{nG_\theta(x, h_n)} \right)^{u+v} v e^{-am}. \end{aligned} \tag{3.27}$$

Making use of the condition (H_3) , it follows that

$$\begin{aligned} \left| \text{Cov}(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v}) \right| &\leq \left(\frac{C\delta_n \|K\|_\infty}{nG_\theta(x, h_n)} \right)^{u+v-2} \left| \text{Cov}(\widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1}) \right| \\ &\leq \left(\frac{C\delta_n \|K\|_\infty}{nG_\theta(x, h_n)} \right)^{u+v-2} \left(\left| \mathbb{E}(\widehat{\Delta}_{s_u} \widehat{\Delta}_{t_1}) \right| + \mathbb{E}|\widehat{\Delta}_{s_u}| \mathbb{E}|\widehat{\Delta}_{t_1}| \right) \\ &\leq \left(\frac{C\delta_n \|K\|_\infty}{nG_\theta(x, h_n)} \right)^{u+v-2} \left(\frac{C}{nG_\theta(x, h_n)} \right)^2 \delta_n^2 M. \end{aligned}$$

where

$$M = \sup_{i \neq j} \mathbb{P} \{ (X_i, X_j) \in B(x, h_n) \times B(x, h_n) \} + \{ \mathbb{P}^2 (X_1 \in B(x, h_n)) \}.$$

Then we obtain a second bound of $\mathbf{Cov}(\cdot, \cdot)$

$$\left| \mathbf{Cov} \left(\widehat{\Delta}_{s_1}, \dots, \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1}, \dots, \widehat{\Delta}_{t_v} \right) \right| \leq C^{u+v} \left(\frac{\delta_n}{nG_\theta(x, h_n)} \right)^{u+v} G_\theta^2(x, h_n). \quad (3.28)$$

Moreover, by multiplying a τ -power of (3.27) and $(1 - \tau)$ -power of (3.28) for some $\frac{1}{4} < \tau < \frac{1}{2}$ we obtain an upper-bound of the covariance as follows for $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$:

$$\left| \mathbf{Cov} \left(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v} \right) \right| \leq C^{u+v} \left(\frac{\delta_n}{nG_\theta(x, h_n)} \right)^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^{2\tau} G_\theta^{2(1-\tau)}(x, h_n) v e^{-a\tau m}.$$

So, by Lemma 3.1, we have

$$\left| \mathbf{Cov} \left(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v} \right) \right| \leq \left(\frac{C\delta_n}{nG_\theta(x, h_n)} \right)^{u+v-2} \left(\frac{C\delta_n}{nG_\theta(x, h_n)} \right)^2 G_\theta(x, h_n) v e^{-a\tau m},$$

where $M_n = \frac{C\delta_n}{nG_\theta(x, h_n)}$ and $K_n = \frac{C\delta_n}{n\sqrt{G_\theta(x, h_n)}}$. It remains to calculate $\text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right)$:

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right) &= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 \sum_i \sum_j \mathbf{Cov} \left(\widehat{Z}_i, \widehat{Z}_j \right) \\ &= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 \left[n\text{Var} \left(\widehat{Z}_1 \right) + \sum_i \sum_{j \neq i} \mathbf{Cov} \left(\widehat{Z}_i, \widehat{Z}_j \right) \right] \\ &= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 [nT_1 + T_{ij}]. \end{aligned}$$

Now, under (H_5) , we obtain for the first term :

$$\begin{aligned} T_1 = \text{Var} \left(\widehat{Z}_1 \right) &= \mathbb{E} \left(\widehat{Y}_1^2 K_1^2(x) \right) - \left(\mathbb{E} \left(\widehat{Y}_1 K_1(x) \right) \right)^2 \\ &\leq \mathbb{E} \left(Y_1^2 K_1^2(x) \right) \\ &\leq \mathbb{E} \left(K_1^2(x) \mathbb{E} \left(Y_1^2 \mid X \right) \right) \leq C \mathbb{E} \left(K_1^2(x) \right). \end{aligned}$$

For all $j \geq 1$

$$\mathbb{E} \left(K_1^j(x) \right) = O \left(G_\theta(x, h_n) \right), \quad (3.29)$$

we have

$$T_1 = \text{Var} \left(\widehat{Z}_1 \right) = O \left(G_\theta(x, h_n) \right).$$

We readily obtain that

$$\frac{1}{n \left(\mathbb{E} \left(K_1(x) \right) \right)^2} T_1 \leq \frac{C G_\theta(x, h_n)}{n O \left(G_\theta^2(x, h_n) \right)}$$

$$\leq \frac{C}{nG_\theta(x, h_n)}.$$

For the second term, we have the following decomposition

$$\begin{aligned} T_{ij} &= \sum_i \sum_{0 < |i-j| \leq u_n} \text{Cov}(\widehat{Z}_i, \widehat{Z}_j) + \sum_i \sum_{|i-j| > u_n} \text{Cov}(\widehat{Z}_i, \widehat{Z}_j) \\ &= J_1 + J_2 \end{aligned}$$

where (u_n) is a sequence of positive integer and

$$\lim_{n \rightarrow \infty} u_n = \infty.$$

Now, under the assumptions (H_2) , (H_3) and (H_5) , we have

$$\begin{aligned} |J_1| &= \sum_i \sum_{0 < |i-j| \leq u_n} \left| \text{Cov}(\widehat{Z}_i, \widehat{Z}_j) \right| \leq nu_n \left[\max_{i \neq j} |\mathbb{E}(K_i K_j)| + (\mathbb{E}(K_1))^2 \right] \\ &\leq Cnu_n \delta_n^2 G_\theta^2(x, h_n). \end{aligned}$$

Making use of the condition (H_6) , we infer that

$$\begin{aligned} |J_2| &= \sum_i \sum_{|i-j| > u_n} \left| \text{Cov}(\widehat{Z}_i, \widehat{Z}_j) \right| \leq C\delta_n^2 \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \sum_i \sum_{|i-j| > u_n} \lambda_{i,j} \\ &\leq Cn\delta_n^2 h_n^{-2} e^{-au_n}. \end{aligned}$$

This implies that

$$\begin{aligned} |T_{ij}| &\leq \sum_{i=1}^n \sum_{i \neq j} \left| \text{Cov}(\widehat{Z}_i, \widehat{Z}_j) \right| \\ &\leq C(nu_n \delta_n^2 G_\theta^2(x, h_n) + n\delta_n^2 h_n^{-2} e^{-au_n}). \end{aligned}$$

We then obtain that

$$\frac{1}{nG_\theta(x, h_n)} |T_{ij}| \leq Cg(u_n), \tag{3.30}$$

where

$$g(u_n) := u_n G_\theta(x, h_n) + G_\theta^{-1}(x, h_n) h_n^{-2} e^{-au_n}.$$

Taking $u_n = \frac{1}{a} \log\left(\frac{a}{h_n^2 G_\theta^2(x, h_n)}\right)$ with $a > 0$. From (3.30) and (H_7) we get

$$\frac{1}{nG_\theta(x, h_n)} |T_{ij}| \rightarrow 0, \text{ as } n \rightarrow \infty. \tag{3.31}$$

Then we get

$$\text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right) = \left[O \left(\frac{1}{nG_\theta(x, h_n)} \right) + \frac{1}{nG_\theta(x, h_n)} o(1) \right]$$

$$= O\left(\frac{1}{nG_\theta(x, h_n)}\right)$$

The conditions of Lemma 3.1 are verified for

$$\begin{aligned} K_n &= \frac{C\delta_n}{n\sqrt{G_\theta(x, h_n)}}, \\ M_n &= \frac{C\delta_n}{nG_\theta(x, h_n)}, \\ A_n &= \text{Var}\left(\sum_{i=1}^n \widehat{\Delta}_i\right) = O\left(\frac{1}{nG_\theta(x, h_n)}\right), \\ B_n &= \left(\frac{16nK^2}{9A_n(1-e^{-\beta})} \vee 1\right) \left(\frac{2(K \vee M)}{1-e^{-\beta}}\right) = \frac{\delta_n^3}{nG_\theta(x, h_n)}. \end{aligned}$$

So, we apply the inequality in Kallabis and Neumann (2006) to the random variables $\widehat{\Delta}_i$ to infer that

$$\begin{aligned} \mathbb{P}\left(|\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x))| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma}G_\theta(x, h_n)}}\right) &= \mathbb{P}\left(\left|\sum_{i=1}^n \widehat{\Delta}_i\right| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma}G_\theta(x, h_n)}}\right) \\ &\leq \exp\left(\frac{-\varepsilon^2 \log n}{n^{1-\gamma}G_\theta(x, h_n)L_\theta(n)}\right), \end{aligned}$$

where

$$L_\theta(n) = \left(O\left(\frac{1}{nG_\theta(x, h_n)}\right) + \left(\frac{\delta_n^3}{nG_\theta(x, h_n)}\right)^{\frac{1}{3}} \left(\frac{\log n}{n^{1-\gamma}G_\theta(x, h_n)}\right)^{\frac{5}{6}}\right).$$

Then we finally obtain that

$$\begin{aligned} \mathbb{P}\left(|\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2})| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma}G_\theta(x, h_n)}}\right) &\leq \exp\left(\frac{-\varepsilon^2 \log n}{C + \left(\frac{\delta_n^6 n^{-\gamma} \log^5 n}{nG_\theta(x, h_n)}\right)^{\frac{1}{6}}}\right) \\ &\leq C_1 \exp(-\varepsilon^2 \log(n)) \end{aligned}$$

The proof is achieved by a suitable choice of ε . □

Proof of Lemma 3.3

The proof of this Lemma is similar to the proof of the previous Lemma 3.2. Since $\widehat{Y}_i = 1$, it suffices to replace $\widehat{\Delta}_i$ by

$$\widetilde{\Delta}_i = \frac{1}{n\mathbb{E}(K_1(x))} [K_i(x) - \mathbb{E}(K_i(x))], \quad i = 1, \dots, n.$$

So, we apply the exponential inequality given by Corollary A8(ii) in Ferraty and Vieu (2006) with $a^2 = G_\theta^{-1}(x, h_n)$. Then we obtain, under $(H_1) - (H_6)$,

$$\begin{aligned} \mathbb{P}\left(|\widehat{r}_{\theta,0}(x) - 1| > \varepsilon \sqrt{\frac{\log n}{nG_\theta(x, h_n)}}\right) &= \mathbb{P}\left(\left|\sum_{i=1}^n \widetilde{\Delta}_i\right| > \varepsilon \sqrt{\frac{\log n}{nG_\theta(x, h_n)}}\right) \\ &\leq C'_1 \exp(-\varepsilon^2 \log(n)). \end{aligned}$$

Thus the proof is complete. □

Proof of Lemma 3.4

Notice that we have

$$\left\{ |\widehat{r}_{\theta,0}(x)| \leq \frac{1}{2} \right\} \subset \left\{ |\widehat{r}_{\theta,0}(x) - 1| > \frac{1}{2} \right\},$$

that implies that

$$\mathbb{P} \left(|\widehat{r}_{\theta,0}(x)| \leq \frac{1}{2} \right) \leq \mathbb{P} \left(|\widehat{r}_{\theta,0}(x) - 1| > \frac{1}{2} \right).$$

Under the hypothesis $(H_1) - (H_6)$ and by applying Lemma 3.3, we deduce that

$$\sum_n \mathbb{P} \left(|\widehat{r}_{\theta,0}(x)| \leq \frac{1}{2} \right) \leq \sum_n \mathbb{P} \left(|\widehat{r}_{\theta,0}(x) - 1| > \frac{1}{2} \right) < \infty$$

Then, for $\eta = \frac{1}{2}$, we have

$$\sum_n \mathbb{P} (|\widehat{r}_{\theta,0}(x)| \leq \epsilon) < \infty.$$

Thus the proof is complete. □

Proof of Lemma 3.5

One can see easily that we have

$$\begin{aligned} |r_{\theta}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))| &= \left| r_{\theta}(x) - \mathbb{E} \left(\frac{1}{n\mathbb{E}(K_1)} \sum_{i=1}^n Y_i K_i \right) \right| \\ &= \frac{1}{\mathbb{E}(K_1)} [|r_{\theta}(x)\mathbb{E}(K_1) - \mathbb{E}(Y_1 K_1)|] \\ &\leq \frac{1}{\mathbb{E}(K_1)} \mathbb{E} [(|r_{\theta}(x) - r_{\theta}(X_1)|) K_1]. \\ &\leq Ch_n^{\beta}, \end{aligned}$$

This readily implies that we have

$$r_{\theta}(x) - \mathbb{E}(\widehat{r}_{\theta,1}) = O(h_n^{\beta}).$$

Thus the proof is complete. □

Proof of Lemma 3.6

We first observe that we have

$$\begin{aligned} |\mathbb{E}(\widehat{r}_{\theta,2}) - \mathbb{E}(\widehat{r}_{\theta,1})| &= \frac{1}{n\mathbb{E}K_1} \left| \mathbb{E} \left(\sum_{i=1}^n Y_i \mathbf{1}_{\{|Y_i| > \delta_n\}} K_i \right) \right| \\ &\leq \mathbb{E} (|Y_1| \mathbf{1}_{|Y_1| > \delta_n} K_1) (\mathbb{E}(K_1))^{-1} \\ &\leq \mathbb{E} (\exp(|Y_1|/4) \mathbf{1}_{|Y_1| > \delta_n} K_1) (\mathbb{E}(K_1))^{-1}. \end{aligned}$$

The Hölder's inequality allows to write that

$$\begin{aligned} |\mathbb{E}(\widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x))| &\leq (\mathbb{E}(\exp(|Y_1|/2) \mathbf{1}_{|Y_1| > \delta_n}))^{\frac{1}{2}} (\mathbb{E}(K_1^2))^{\frac{1}{2}} (\mathbb{E}(K_1))^{-1} \\ &\leq \exp(-\delta_n/4) (\mathbb{E}(\exp(|Y_1|)))^{\frac{1}{2}} (\mathbb{E}(K_1^2))^{\frac{1}{2}} (\mathbb{E}(K_1))^{-1} \\ &\leq C \exp(-\delta_n/4) (G_\theta(x, h))^{\frac{-1}{2}}. \end{aligned}$$

Hence we obtain that

$$\frac{|\mathbb{E}(\widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x))|}{\sqrt{\frac{\log(n)}{n^{1-\gamma} G_\theta(x, h)}}} \leq C \sqrt{n^{1-\gamma}} \exp(-\delta_n/4) \log^{\frac{-1}{2}}(n). \quad (3.32)$$

So, we can write

$$|\mathbb{E}(\widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x))| = o\left(\sqrt{\frac{\log n}{n^{1-\gamma} G_\theta(x, h)}}\right).$$

Thus the proof is complete. □

Proof of Lemma 3.7

By (H_5) and we apply the Markov's inequality to show that, $\forall \epsilon > 0$,

$$\begin{aligned} \mathbb{P}(|\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)| > \epsilon) &= \mathbb{P}\left(\frac{1}{n\mathbb{E}K_1} \left| \sum_{i=1}^n Y_i \mathbf{1}_{|Y_i| > \delta_n} K_i \right| > \epsilon\right) \\ &\leq n\mathbb{P}(|Y_1| > \delta_n) \\ &\leq n \exp(-\delta_n) \mathbb{E}(\exp(|Y|)) \\ &\leq Cn \exp(-\delta_n). \end{aligned}$$

This, in turn, implies that

$$\sum_{n \geq 1} \mathbb{P}(|\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)| > \epsilon) \leq C \sum_{n \geq 1} n \exp(-\delta_n) < \infty. \quad (3.33)$$

which completes the proof of Lemma. □

Proof of Lemma 3.8

In order to handle with the variance given in the limiting distribution, we will use the following decomposition of the variance of $\widehat{r}_{\theta,1}$

$$\begin{aligned} \text{Var}(\widehat{r}_{\theta,1}(x)) &= \mathbb{E}[(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) + (\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x)))]^2 \\ &= \mathbb{E}[(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x))]^2 + \text{Var}(\widehat{r}_{\theta,2}(x)) \\ &+ 2\mathbb{E}[(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) (\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x)))] \\ &= M_1 + M_2 + M_3. \end{aligned} \quad (3.34)$$

By using Hölder inequality and for $\delta_n = \delta_0 \log(n)$, $\delta_0 > 0$, it follows that

$$\begin{aligned}
 & (n\psi(h_n)) \mathbb{E} |(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x))|^2 \\
 & \leq \frac{C^2(n\psi(h_n))}{\psi^2(h_n)} \mathbb{E} [|Y_1|^2 \mathbb{1}_{\{|Y_1| > \delta_n\}} K_1^2(x)] \\
 & \leq C \frac{n}{\psi(h_n)} \left[\mathbb{E} |Y_1|^4 K_1^4(x) \right]^{\frac{1}{2}} \left(\frac{\mathbb{E}(\exp(|Y_1|))}{\exp(\delta_n)} \right)^{\frac{1}{2}} \\
 & \leq C \frac{n}{\psi(h_n)} \left(\mathbb{E} \left[\mathbb{E}(|Y_1|^4 | X_1) K_1^4(x) \right] \right)^{\frac{1}{2}} (\mathbb{E}(\exp(|Y_1|)))^{\frac{1}{2}} (\exp(\delta_n))^{-\frac{1}{2}} \\
 & \leq C n^{1-\frac{\delta_0}{2}} \psi^{-\frac{1}{2}}(h_n) \rightarrow 0.
 \end{aligned} \tag{3.35}$$

We have

$$\text{Var}(\widehat{r}_{\theta,2}(x)) = \frac{1}{n(\mathbb{E}K_1(x))^2} \text{Var}(\widehat{Y}_1 K_1) + \frac{1}{(n\mathbb{E}K_1(x))^2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{Cov}(\widehat{Y}_i K_i, \widehat{Y}_j K_j). \tag{3.36}$$

Notice that

$$\begin{aligned}
 \text{Var}(\widehat{Y}_1 K_1) &= \mathbb{E}(Y_1^2 \mathbb{1}_{\{|Y_1| \leq \delta_n\}} K_1^2) - (\mathbb{E}(Y_1 \mathbb{1}_{\{|Y_1| \leq \delta_n\}} K_1(x)))^2 \\
 &= \mathbb{E}(Y_1^2 K_1^2) - \mathbb{E}(Y_1^2 \mathbb{1}_{\{|Y_1| > \delta_n\}} K_1^2) \\
 &\quad - [\mathbb{E}(Y_1 K_1) - \mathbb{E}(Y_1 \mathbb{1}_{\{|Y_1| > \delta_n\}} K_1)]^2.
 \end{aligned} \tag{3.37}$$

By using the condition (B_1) , we have

$$\begin{aligned}
 \mathbb{E}(Y_1^2 K_1^2) &= \mathbb{E}(\mathbb{E}(Y_1^2 | X_1) K_1^2) \\
 &= \mathbb{E}((\varphi(X_1) - \varphi(x)) K_1^2) + \varphi(x) \mathbb{E}(K_1^2) \\
 &\leq \sup_{\{z: |\theta, x-z| \leq h_n\}} |\varphi(z) - \varphi(x)| \mathbb{E}(K_1^2(x)) + \varphi(x) \mathbb{E}(K_1^2) \\
 &= (\varphi(x) + o(1)) \mathbb{E}[K_1^2(x)],
 \end{aligned} \tag{3.38}$$

this when combined with the fact that the function $r_\theta(\cdot)$ is a Lipschitzian function, implies that

$$\begin{aligned}
 \mathbb{E}(Y_1 K_1) &= r_\theta(x) \mathbb{E}(K_1(x)) + \mathbb{E}((r_\theta(X_1) - r_\theta(x)) K_1(x)) \\
 &= \left(r_\theta(x) + O(h_n^\beta) \right) \mathbb{E}(K_1(x)).
 \end{aligned} \tag{3.39}$$

We use the Hölder inequality, and by $(H_5)(i)$ we get that

$$\begin{aligned}
 |\mathbb{E}(Y_1 \mathbf{1}_{\{|Y_1| > \delta_n\}} K_1(x))| &\leq \mathbb{E}(|Y_1| \mathbf{1}_{\{|Y_1| > \delta_n\}} K_1(x)) \\
 &\leq \mathbb{E}\left(\exp\left(\frac{|Y_1|}{4}\right) \mathbf{1}_{\{|Y_1| > \delta_n\}} K_1(x)\right) \\
 &\leq \left[\mathbb{E}\left(\exp\left(\frac{|Y_1|}{2}\right) \mathbf{1}_{\{|Y_1| > \delta_n\}}\right)\right]^{\frac{1}{2}} (\mathbb{E}(K_1^2(x)))^{\frac{1}{2}} \\
 &\leq \left[(\mathbb{E}(\exp(|Y_1|)))^{\frac{1}{2}} \left(\frac{\mathbb{E}(\exp(|Y_1|))}{\exp(\delta_n)}\right)^{\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbb{E}(K_1^2(x)))^{\frac{1}{2}} \\
 &\leq (\mathbb{E}(\exp(|Y_1|)))^{\frac{1}{2}} \left(\exp\left(-\frac{\delta_n}{4}\right)\right)^{\frac{1}{2}} (\mathbb{E}(K_1^2(x)))^{\frac{1}{2}} \\
 &\leq C\psi^{\frac{1}{2}}(h_n)n^{-\frac{\delta_0}{4}} \\
 &= o\left(\sqrt{\frac{1}{n^{\frac{\delta_0}{2}}\psi(h_n)}}\right), \tag{3.40}
 \end{aligned}$$

and

$$\begin{aligned}
 |\mathbb{E}(Y_1^2 \mathbf{1}_{\{|Y_1| > \delta_n\}} K_1^2(x))| &\leq \mathbb{E}(|Y_1|^2 \mathbf{1}_{\{|Y_1| > \delta_n\}} K_1^2(x)) \\
 &\leq (\mathbb{E}(|Y_1|^4 K_1^4(x)))^{\frac{1}{2}} \left(\frac{\mathbb{E}(\exp(|Y_1|))}{\exp(\delta_n)}\right)^{\frac{1}{2}} \\
 &\leq [\mathbb{E}(\mathbb{E}(|Y_1|^4 | X) K_1^4(x))]^{\frac{1}{2}} (\mathbb{E}(\exp(|Y_1|)))^{\frac{1}{2}} \exp\left(-\frac{\delta_n}{2}\right) \\
 &\leq C\psi^{\frac{1}{2}}(h_n)n^{-\frac{\delta_0}{2}} \\
 &= o\left(\sqrt{\frac{1}{n^{\delta_0}\psi(h_n)}}\right). \tag{3.41}
 \end{aligned}$$

Once more, an application of the Hölder inequality in connection with equation (3.35) and the third term of equation (3.34), we have

$$\begin{aligned}
 &(n\psi(h_n)) \mathbb{E}[(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x))(\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2}(x)))] \\
 &\leq \left[(n\psi(h_n)) \mathbb{E}(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x) \mathbb{E}[(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x))]^2)\right]^{\frac{1}{2}} \times [(n\psi(h_n)) \text{Var}(\widehat{r}_{\theta,2}(x))]^{\frac{1}{2}} \\
 &\longrightarrow 0.
 \end{aligned}$$

For $j = 1, 2$, we have

$$\frac{1}{\psi(h_n)} \mathbb{E}\left[K\left(\frac{\langle \theta, x - X_1 \rangle}{h_n}\right)\right] = \frac{1}{\psi(h_n)} \int_0^{h_n} K^j\left(\frac{z}{h_n}\right) dG_\theta(z, x) \longrightarrow C_j f_1(x) \tag{3.42}$$

Using the equations (3.37), (3.38), (3.39), (3.40), (3.41), (3.42), we get

$$\frac{\psi(h_n)}{(\mathbb{E}K_1(x))^2} \text{Var}(\widehat{Y}_1 K_1(x)) = \varphi(x) \frac{C_2}{C_1^2 f_1(x)} + o(1). \tag{3.43}$$

By combination of equations (3.30), (3.31), (3.34), (3.36) and (3.43) we obtain the first assertion of Lemma 3.19 :

$$\frac{\psi(h_n)}{(\mathbb{E}K_1(x))^2} \text{Var}(\widehat{r}_{\theta,1}(x)) \longrightarrow \varphi(x) \frac{C_2}{C_1^2 f_1(x)}.$$

We will use in the second assertion of Lemma 3.8 with the following notation

$$\widetilde{Y}_i = Y_i - r_\theta(x).$$

By replacing Y_i with \widetilde{Y}_i , the expression of the variance is written as follows :

$$\begin{aligned} \text{Var}(\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x)) &= \text{Var}[\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x) - r_\theta(x)\widehat{r}_{\theta,0}(x))] \\ &= \text{Var}\left[\frac{1}{n\mathbb{K}_1(x)} \sum_{i=1}^n Y_i K_i(x) - r_\theta(x) \frac{1}{n\mathbb{K}_1(x)} \sum_{i=1}^n K_i(x)\right] \\ &= \text{Var}\left[\frac{1}{n\mathbb{K}_1(x)} \sum_{i=1}^n \widetilde{Y}_i K_i(x)\right] \\ &= \frac{1}{n(\mathbb{K}_1(x))^2} \left[\mathbb{E}(\widetilde{Y}_1^2 K_1^2(x)) - \left(\mathbb{E}(\widetilde{Y}_1 K_1(x))\right)^2 \right]. \end{aligned} \quad (3.44)$$

Hence by the condition (H_4) , we obtain that

$$\begin{aligned} \left| \mathbb{E}(\widetilde{Y}_1 K_1(x)) \right| &= |\mathbb{E}((Y_1 - r_\theta(x)) K_1(x))| \\ &= |\mathbb{E}((r_\theta(X_1) - r_\theta(x)) K_1(x))| \\ &\leq Ch_n^\beta \psi(h_n) \longrightarrow 0. \end{aligned} \quad (3.45)$$

By condition (B_1) , we have

$$\begin{aligned} \left| \mathbb{E}(\widetilde{Y}_1^2 K_1^2(x)) \right| &= \left| \mathbb{E}(\mathbb{E}(\widetilde{Y}_1^2 | X_1) K_1^2(x)) \right| \\ &= \left| \mathbb{E}((\varphi(X_1) - \varphi(x) - r_\theta^2(x) + \varphi(x)) K_1^2(x)) \right| \\ &= (\varphi(x) - r_\theta^2(x) + o(1)) \mathbb{E}(K_1^2(x)). \end{aligned} \quad (3.46)$$

We finally obtain

$$\frac{\psi(h_n)}{(\mathbb{E}(K_1(x)))^2} \text{Var}(\widetilde{Y}_1 K_1(x)) = \frac{C_2}{C_1^2} \frac{(\varphi(x) - r_\theta^2(x))}{f_1(x)} + o(1)$$

Thus the proof is complete. □

Proof of Theorem 3.2

Let us introduce the following sum

$$S_n = \sum_{i=1}^n L_{ni}(x),$$

where

$$L_{ni}(x) = \frac{\sqrt{\psi(h_n)}}{\sqrt{n\mathbb{E}[K_1(x)]}} \left(\widehat{Y}_i K_i(x) - \mathbb{E}(\widehat{Y}_i K_i(x)) \right).$$

Therefore S_n can be written as follows

$$S_n = \sqrt{n\psi(h_n)} (\hat{r}_{\theta,2}(x) - \widehat{r}_{\theta,2}(x)),$$

where

$$\widehat{r}_{\theta,2}(x) = \frac{1}{n\mathbb{E}K_1} \sum_{i=1}^n \widehat{Y}_i K_i(x) \text{ and } \widehat{Y}_i = Y_i \mathbf{1}_{\{|Y_i| \leq \delta_n\}}.$$

In order to complete the proof of Theorem 3.2, we need first to prove the following result

$$\sqrt{n\psi(h_n)} (\widehat{r}_{\theta,2}(x) - \widehat{r}_{\theta,2}(x)) - S_n \xrightarrow{\mathbb{L}^2} 0^3. \quad (3.47)$$

The proof of this equation (3.47) is easy to deduce from (3.35) and by using Hölder inequality and by choosing $\delta_n = \delta_0 \log(n)$. Thus, it suffices now to prove

$$S_n \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,1}^2(x)). \quad (3.48)$$

In order to demonstrate the equation (3.48), we use the basic technique of Doob (1953) consists in splitting the set $\{1, \dots, n\}$ in two sequences of natural numbers tending to infinity, such that $p = p_n$, $q = q_n$ and $k = k_n$. We decompose S_n into

$$S_n = \Lambda_n + \Lambda'_n + \Lambda''_n,$$

where

$$\begin{aligned} \Lambda_n &= \sum_{j=1}^k \eta_j, \\ \Lambda'_n &= \sum_{j=1}^k \xi_j, \\ \Lambda''_n &:= \zeta_k := \sum_{i=k(p+q)+1}^n L_{ni}(x), \\ \eta_j &:= \sum_{i \in I_j} L_{ni}(x), \\ \xi_j &:= \sum_{i \in J_j} L_{ni}(x). \end{aligned}$$

For $j = 1, \dots, k$, I_j and J_j denoted by p -blocks and q -blocks respectively as follows :

$$\begin{aligned} I_j &= \{(j-1)(p+q) + 1, \dots, (j-1)(p+q) + p\} \\ J_j &= \{(j-1)(p+q) + p + 1, \dots, j(p+q) + p\}. \end{aligned}$$

Let k be defined by $\lfloor \frac{n}{(p+q)} \rfloor$ (where $\lfloor x \rfloor$ stands for the integral part of x). We suppose that $\frac{qk}{n} \rightarrow 0$ and $\frac{pk}{n} \rightarrow 1$. The convergence in (3.48) will be established by showing the following asymptotic result

$$\Lambda_n \xrightarrow{\mathcal{D}} N(0, \sigma_{\theta,1}^2(x)), \quad (3.49)$$

3. $\xrightarrow{\mathbb{L}^2}$ denotes the convergence in quadratic mean.

and

$$\mathbb{E} (\Lambda'_n)^2 + \mathbb{E} (\Lambda''_n)^2 \longrightarrow 0. \quad (3.50)$$

Proof of (3.50). By stationarity, we get

$$\begin{aligned} \mathbb{E} (\Lambda'_n)^2 &= \mathbb{E} \left(\left(\sum_{j=1}^k \xi_j \right)^2 \right) \\ &\leq k \text{Var}(\xi_1) + 2 \sum_{1 \leq i < j \leq k} |\text{Cov}(\xi_i, \xi_j)|, \end{aligned} \quad (3.51)$$

and

$$\begin{aligned} k \text{Var}(\xi_1) &= k \text{Var} \left(\sum_{i \in J_1} L_{ni}(x) \right) \\ &\leq qk \text{Var}(L_{n1}(x)) + 2k \sum_{1 \leq i < j \leq q} |\text{Cov}(L_{ni}(x), L_{nj}(x))|. \end{aligned} \quad (3.52)$$

Using the fact that $\frac{kq}{n} \longrightarrow 0$, $\frac{q}{n} \longrightarrow 0$ and the convergence in (3.43), we readily obtain the first term in the right-hand of inequality (3.52). The second term in inequality (3.52), by choosing u_n such that

$$u_n = \frac{1}{a} \log \left(\frac{a}{\psi^{2(1+b)}(h_n)} \right),$$

we will use the same natural decomposition given in (3.30). We then have

$$\begin{aligned} k \sum_{1 \leq i < j \leq q} |\text{Cov}(L_{ni}(x), L_{nj}(x))| &= \frac{k\psi(h_n)}{n (\mathbb{E}(K_1(x)))^2} \sum_{1 \leq i < j \leq q} \left| \text{Cov}(\widehat{Y}_i K_i(x), \widehat{Y}_j K_j(x)) \right| \\ &\leq C \frac{kq}{n} [\delta_n^2 \psi(h_n) u_n + \delta_n^2 \psi^{-1}(h_n) h_n^{-2} e^{-au_n}], \end{aligned} \quad (3.53)$$

where $\frac{kq}{n} \longrightarrow 0$ and

$$\frac{1}{\psi^2(h_n)} (\mathbb{E}(K_1))^2 \longrightarrow C_1^2 f_1^2(x) \text{ as } n \longrightarrow \infty.$$

Then we obtain as in (1.15) the following result

$$k \sum_{1 \leq i < j \leq q} |\text{Cov}(L_{ni}(x), L_{nj}(x))| \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

By stationarity, the second term of inequality (3.51) gives

$$\begin{aligned} \sum_{1 \leq i < j \leq k} |\text{Cov}(\xi_i, \xi_j)| &= \sum_{l=1}^{k-1} (k-l) |\text{Cov}(\xi_1, \xi_{l+1})| \\ &\leq k \sum_{l=1}^{k-1} |\text{Cov}(\xi_1, \xi_{l+1})|. \end{aligned} \quad (3.54)$$

For all $(i, j) \in J_i \times J_j$, $|i - j| \geq p + 1 > p$, we have

$$\begin{aligned}
 \sum_{1 \leq i < j \leq k} |\mathbf{Cov}(\xi_i, \xi_j)| &\leq k \sum_{i=1}^{k-1} \sum_{(l,j) \in J_i \times J_{i+1}} |\mathbf{Cov}(L_{ni}(x), L_{nj}(x))| \\
 &\leq k \frac{C\psi(h_n)\delta_n^2}{nh_n^2 (\mathbb{E}(K_1(x)))^2} \sum_{l=1}^p \sum_{j=2p+q+1; |i-j|>p}^k \lambda_{i,j} \\
 &\leq C \frac{kp}{n} \frac{\psi(h_n)\delta_n^2}{h_n^2 (\mathbb{E}(K_1(x)))^2} \lambda_p.
 \end{aligned} \tag{3.55}$$

The term $\mathbb{E}(\Lambda_n'')^2$ can be written as follows, where

$$\Lambda_n'' := \zeta_k := \sum_{i=k(p+q)+1}^n L_{ni}(x).$$

We have

$$\begin{aligned}
 \mathbb{E}(\Lambda_n'')^2 &\leq (n - k(p+q)) \mathbf{Var}(L_{n1}(x)) + 2 \sum_{1 \leq i < j \leq n} |\mathbf{Cov}(L_{ni}(x), L_{nj}(x))| \\
 &\leq p \mathbf{Var}(L_{n1}(x)) + 2 \sum_{1 \leq i < j \leq n} |\mathbf{Cov}(L_{ni}(x), L_{nj}(x))| \\
 &\leq \frac{p\psi(h_n)}{n (\mathbb{E}(K_1(x)))^2} \mathbf{Var}(\widehat{Y}_1 K_1(x)) + \frac{C}{n\psi(h_n)} \\
 &\quad \times \sum_{1 \leq i < j \leq n} \left| \mathbf{Cov}(\widehat{Y}_i K_i(x), \widehat{Y}_j K_j(x)) \right|.
 \end{aligned} \tag{3.56}$$

Using that the fact $k = \left\lfloor \frac{n}{p+q} \right\rfloor$, $\frac{p}{n} \rightarrow 0$ and by the convergence (3.31), (3.30), (3.43) we readily obtain the right hand side of (3.56) tends to 0 when n tends to infinity.

Proof of (3.49). The proof of convergence in (3.49) is based in the following two results

$$\left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k \mathbb{E} (e^{it\eta_j}) \right| \rightarrow 0, \tag{3.57}$$

and

$$k \mathbf{Var}(\eta_1) \rightarrow \sigma_{\theta,1}^2(x), \quad k \mathbb{E} \left(\eta_1^2 \mathbf{1}_{\{\eta_1 > \varepsilon \sigma_{\theta,1}(x)\}} \right) \rightarrow 0. \tag{3.58}$$

Convergence in (3.57) to show the independence of η_j through using characteristic functions, and the convergence in (3.58) is the condition of Lindeberg-Feller for asymptotic normality of Λ_n .

Proof of (3.57). Let us introduce the following decomposition

$$\begin{aligned}
 \left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k \mathbb{E} (e^{it\eta_j}) \right| &\leq \left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \mathbb{E} \left(e^{it \sum_{j=1}^{k-1} \eta_j} \right) \mathbb{E} (e^{it\eta_k}) \right| \\
 &+ \left| \mathbb{E} \left(e^{it \sum_{j=1}^{k-1} \eta_j} \right) - \prod_{j=1}^{k-1} \mathbb{E} (e^{it\eta_j}) \right| \\
 &= \left| \text{Cov} \left(e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it\eta_k} \right) \right| \\
 &+ \left| \mathbb{E} \left(e^{it \sum_{j=1}^{k-1} \eta_j} \right) - \prod_{j=1}^{k-1} \mathbb{E} (e^{it\eta_j}) \right|. \tag{3.59}
 \end{aligned}$$

We get successively

$$\begin{aligned}
 \left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k \mathbb{E} (e^{it\eta_j}) \right| &\leq \left| \text{Cov} \left(e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it\eta_k} \right) \right| \\
 &+ \left| \text{Cov} \left(e^{it \sum_{j=1}^{k-2} \eta_j}, e^{it\eta_{k-1}} \right) \right| + \dots + |\text{Cov} (e^{it\eta_2}, e^{it\eta_1})|. \tag{3.60}
 \end{aligned}$$

For the right hand side of (3.60), we use the quasi-associated property to get

$$\begin{aligned}
 &\left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k \mathbb{E} (e^{it\eta_j}) \right| \\
 &\leq \frac{Ct^2 \delta_n^2 \psi(h_n)}{nh_n^2 (\mathbb{E}[K_1(x)])^2} \left[\sum_{i \in I_1} \sum_{j \in I_2} \lambda_{i,j} + \dots + \sum_{i \in (I_1 \cup I_2 \cup \dots \cup I_{k-1})} \sum_{j \in I_k} \lambda_{i,j} \right]. \tag{3.61}
 \end{aligned}$$

By stationarity and for every $2 \leq l \leq k-1$, $(i, j) \in I_l \times I_{l+1}$, we have $|i - j| \geq q + 1 > q$, then we obtain

$$\begin{aligned} \left| \mathbb{E} \left(e^{it \sum_{j=1}^k \eta_j} \right) - \prod_{j=1}^k \mathbb{E} (e^{it\eta_j}) \right| &\leq \frac{Ckpt^2 \delta_n^2 \psi(h_n)}{nh_n^2 (\mathbb{E}[K_1(x)])^2} \lambda_q \\ &\leq \frac{Ckpt^2 \delta_n^2 \psi(h_n)}{nh_n^2 (\mathbb{E}[K_1(x)])^2} e^{-aq} \rightarrow 0. \end{aligned} \quad (3.62)$$

Proof of (3.58). By the similar arguments in (3.31), (3.30) and (3.43), we get the following result

$$\begin{aligned} k\text{Var}(\eta_1) &= k\text{Var} \left(\sum_{i \in I_1} L_{ni}(x) \right) \\ &= kp\text{Var}(L_{n1}(x)) + 2k \sum_{1 \leq i < j \leq p} |\text{Cov}(L_{ni}(x), L_{nj}(x))| \\ &= \frac{kp\psi(h_n)}{n(\mathbb{E}(K_1(x)))^2} \text{Var}(\widehat{Y}_1 K_1(x)) \\ &\quad + \frac{2k}{n\psi(h_n)} \sum_{1 \leq i < j \leq p} \left| \text{Cov}(\widehat{Y}_i K_i(x), \widehat{Y}_j K_j(x)) \right| \\ &\rightarrow \sigma_{\theta,1}^2(x). \end{aligned} \quad (3.63)$$

Using the fact that

$$\begin{aligned} |\eta_1| &= \left| \sum_{i=1}^p L_{ni}(x) \right| \\ &\leq \frac{\sqrt{\psi(h_n)}}{\sqrt{n}\mathbb{E}[K_1(x)]} \sum_{i=1}^p \left| \widehat{Y}_i K_i(x) - \mathbb{E}(\widehat{Y}_i K_i(x)) \right| \\ &\leq C \frac{\delta_n p}{\sqrt{n\psi(h_n)}}, \end{aligned} \quad (3.64)$$

and by Tchebyshev's inequality we get

$$\begin{aligned} k\mathbb{E} \left(\eta_1^2 \mathbf{1}_{\eta_1 > \varepsilon \sigma_{\theta,1}(x)} \right) &\leq C \frac{k\delta_n^2 p^2}{n\psi(h_n)} \times \mathbb{P}(\eta_1 > \varepsilon \sigma_{\theta,1}(x)) \\ &\leq C \frac{\delta_n^2 p^2}{n\psi(h_n)} \frac{k\text{Var}(\eta_1)}{\varepsilon^2 \sigma_{\theta,1}^2(x)} \\ &\leq C \frac{\log^2(n)}{n\psi^{1+2\delta}(h_n)} \frac{k\text{Var}(\eta_1)}{\varepsilon^2 \sigma_{\theta,1}^2(x)} \rightarrow 0. \end{aligned} \quad (3.65)$$

Similarly to find the second convergence in (3.16) is an application of the first one (3.15), it suffices to replace Y_i by $\tilde{Y}_i = Y_i - r_\theta(x)$. \square

Proof of Corollary 3.1

By combining Theorem 3.2 and the decomposition (3.17), we establish the normality asymptotic of $\widehat{r}_\theta(x, h_n)$ to $r_\theta(x)$. Thus, to achieve this proof, it suffices to just see when n tends to infinity

$$\mathbb{E} \left[(n\psi(h_n))^{\frac{1}{2}} J_1 \right]^2 \longrightarrow \sigma_{\theta,2}^2(x).$$

By Lemma 3.5 we obtain

$$\begin{aligned} (n\psi(h_n))^{\frac{1}{2}} \mathbb{E}[J_2] &= (n\psi(h_n))^{\frac{1}{2}} [r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))] \\ &= O\left(\left(n\psi(h_n)h_n^{2\beta}\right)^{\frac{1}{2}}\right). \end{aligned}$$

We have

$$\begin{aligned} (n\psi(h_n))^{\frac{1}{2}} \mathbb{E}[J_3] &= (n\psi(h_n))^{\frac{1}{2}} (r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) \mathbb{E}[\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x))] \\ &= O\left(\left(n\psi(h_n)h_n^{2\beta}\right)^{\frac{1}{2}}\right). \end{aligned}$$

Notice that

$$\begin{aligned} (n\psi(h_n))^{\frac{1}{2}} \mathbb{E}|J_4| &= (n\psi(h_n))^{\frac{1}{2}} \mathbb{E}|(\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x))) (\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x)))| \\ &\leq (n\psi(h_n))^{\frac{1}{2}} \left(\mathbb{E}(\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1}(x)))^2\right)^{\frac{1}{2}} \left(\mathbb{E}(\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x)))^2\right)^{\frac{1}{2}} \\ &\leq \frac{1}{(n\psi(h_n))^{\frac{1}{2}}} [(n\psi(h_n) \text{Var}(\widehat{r}_{\theta,1}(x)))^{\frac{1}{2}}] [(n\psi(h_n) \text{Var}(\widehat{r}_{\theta,0}(x)))^{\frac{1}{2}}] \\ &= O\left(\frac{1}{(n\psi(h_n))^{\frac{1}{2}}}\right). \end{aligned}$$

In a similar way, we have

$$\begin{aligned} \left((n\psi(h_n))^{\frac{1}{2}} \mathbb{E}[J_5]\right)^{\frac{1}{2}} &= (n\psi(h_n))^{\frac{1}{4}} (\mathbb{E}|\widehat{r}_\theta(x, h_n)|)^{\frac{1}{2}} \left[\mathbb{E}(\widehat{r}_{\theta,0}(x) - \mathbb{E}(\widehat{r}_{\theta,0}(x)))^2\right]^{\frac{1}{2}} \\ &= O\left(\left(\frac{\log(n)}{(n\psi(h_n))^{\frac{1}{2}}}\right)^{\frac{1}{2}}\right). \end{aligned}$$

Hence the proof is complete. □

Bibliographie

- Ait-Saïdi, A., Ferraty, F., and Kassa, R. (2005). Single functional index model for a time series. *Rev. Roumaine Math. Pures Appl.*, **50**(4), 321–330.
- Ait-Saïdi, A., Ferraty, F., Kassa, R., and Vieu, P. (2008). Cross-validated estimations in the single-functional index model. *Statistics*, **42**(6), 475–494.
- Attaoui, S. and Ling, N. (2016). Asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling for time series data with applications. *Metrika*, **79**(5), 485–511.
- Attaoui, S., Laksaci, A., and Ould Said, E. (2011). A note on the conditional density estimate in the single functional index model. *Statist. Probab. Lett.*, **81**(1), 45–53.
- Attaoui, S., Laksaci, A., and Saïd, E. O. (2015). Asymptotic results for an M -estimator of the regression function for quasi-associated processes. In *Functional statistics and applications*, Contrib. Stat., pages 3–28. Springer, Cham.
- Bulinski, A. and Suquet, C. (2001). Normal approximation for quasi-associated random fields. *Statistics and Probability Letters*, **54**, 215–226.
- Cardot, H., Ferraty, F., and Sarda, P. (2003). Spline estimators for the functional linear model. *Statist. Sinica*, **13**(3), 571–591.
- Chen, D., Hall, P., and Müller, H.-G. (2011). Single and multiple index functional regression models with nonparametric link. *The Annals of Statistics*, **39**(3), 1720–1747.
- Delecroix, M., Härdle, W., and Hristache, M. (2003). Efficient estimation in conditional single-index regression. *Journal of Multivariate Analysis*, **86**(2), 213–226.
- Ding, H., Liu, Y., Xu, W., and Zhang, R. (2017). A class of functional partially linear single-index models. *J. Multivariate Anal.*, **161**, 68–82.
- Doob, J. L. (1953). *Stochastic processes*. John Wiley & Sons, Inc., New York ; Chapman & Hall, Limited, London.
- Douge, L. (2010). Théorèmes limites pour des variables quasi-associées hilbertiennes. *Ann. I.S.U.P.*, **54**(1-2), 51–60.
- Douge, L. (2018). Nonparametric regression estimation for quasi-associated hilbertian processes.

- Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, **84**(2), 313–342.
- Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.*, **38**, 1466–1474.
- Ferraty, F. and Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist.*, **17**(4), 545–564.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Peuch, A., and Vieu, P. (2003). Modèle à indice fonctionnel simple. *C. R. Math. Acad. Sci. Paris*, **336**(12), 1025–1028.
- Ferraty, F., Park, J., and Vieu, P. (2011). Estimation of a functional single index model. In *Recent advances in functional data analysis and related topics*, Contrib. Statist., pages 111–116. Physica-Verlag/Springer, Heidelberg.
- Gasser, T., Hall, P., and 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.
- Härdle, W., Hall, P., and Ichimura, H. (1993). Optimal smoothing in single-index models. *Ann. Statist.*, **21**(1), 157–178.
- Hristache, M., Juditsky, A., and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Ann. Statist.*, **29**(3), 595–623.
- Huang, Z. (2010). Empirical likelihood for single-index varying-coefficient models with right-censored data. *J. Korean Statist. Soc.*, **39**(4), 533–544.
- Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *J. Econometrics*, **58**(1-2), 71–120.
- Kallabis, R. S. and Neumann, M. H. (2006). An exponential inequality under weak dependence. *Bernoulli*, **12**(2), 333–350.
- Li, T.-T., Yang, H., Wang, J.-L., Xue, L.-G., and Zhu, L.-X. (2011). Estimation for a partial-linear single-index model [correction on mr2589322]. *Ann. Statist.*, **39**(6), 3441–3443.
- Ling, N. and Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Statist. Probab. Lett.*, **82**(12), 2235–2243.
- Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stochastic Process. Appl.*, **115**(1), 155–177.
- Mechab, W. and Laksaci, A. (2016). Nonparametric relative regression for associated random variables. *Metron*, **74**(1), 75–97.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer Series in Statistics. Springer, New York, second edition.

-
- Richard E. Barlow, F. P. (1975). *Statistical theory of reliability and life testing : probability models*. International series in decision processes Series in quantitative methods for decision making. Holt, Rinehart and Winston, 1975.
- Shang, H. L. (2018). Estimation of a functional single index model with dependent errors and unknown error density.
- Strzalkowska-Kominiak, E. and Cao, R. (2013). Maximum likelihood estimation for conditional distribution single-index models under censoring. *J. Multivariate Anal.*, **114**, 74–98.
- Ye, Z. and Hooker, G. (2018). Local quadratic estimation of the curvature in a functional single index model.

Chapitre 4

La méthode des k plus proches voisins dans le modèle à indice fonctionnel simple pour les données de séries chronologiques quasi-associées.

Ce chapitre présente la méthode des k -plus proches voisins (k -Nearest Neighbors ou k -NN) dont l'objectif est d'estimer, sous la condition de dépendance de quasi-association, la fonction de régression pour une variable réponse réelle et une variable explicative à valeurs dans un espace semi-métrique fonctionnel. Cette méthode des k -NN a de nombreux avantages, principalement celui de respecter la structure locale des données, essentielle en dimension infinie. Le principal résultat de ce chapitre est l'obtention de la convergence presque complète de notre estimateur de régression et sa vitesse de convergence en utilisant le même modèle MIFS.

Le chapitre fait l'objet d'un article soumis dans une revue internationale.

The k -Nearest Neighbors method in single index regression model for functional quasi-associated times series data.

Salim Bouzebda¹, Ali Laksaci² and Mustapha Mohammedi³

¹Alliance Sorbonne Université, Université de Technologie de Compiègne,
L.M.A.C., Compiègne, France

²Department of Mathematics, College of Science,
King Khalid University, Abha, 61413, Saudi Arabia

³Université Djillali Liabès, BP 89, 22000, Sidi Bel Abbès, Algérie,
L.S.P.S., Sidi Bel Abbès, Algérie

Abstract

In the present paper, we consider the k -Nearest Neighbors method in single index regression model when the explanatory variable is valued in functional space, and under the quasi-association dependence condition. The main result of this work is the establishment of the almost complete convergence rates. The results are obtained rely on the classical functional kernel estimate.

Key words : Almost complete convergence rates, Functional Hilbert space, k -Nearest Neighbors (k -NN), Kernel regression estimation, Quasi-associated variables, Single functional index model, Weak dependence.

AMS subject classification : 62G05 ; 62G08 ; 62L12 ; 62G20.

4.1 Introduction

The framework of this work is the estimation of the regression function via the k Nearest Neighbors (k -NN) method in single index nonparametric regression model under the quasi-association dependence condition. The single index are very popular in the econometric community and respond to two important preoccupations. The first concerns dimension reduction since this type of model makes it possible to provide a solution to the problem of the curse of dimensionality. The second is related to the interpretability of the index θ introduced in these models, one can refer to [Härdle *et al.* \(1993\)](#) and [Hristache *et al.* \(2001\)](#) in the case of a vector explanatory variable and [Ferraty *et al.* \(2003\)](#) in the infinite dimension. Therefore, the single functional index model accumulate the advantages of single index model, and inherits the potential of the functional linear model in terms of applications, refer [Ramsay and Silverman \(2005\)](#), [Ferraty and Vieu \(2002\)](#), [Ferraty and Vieu \(2006\)](#) and [Ferraty *et al.* \(2011\)](#).

The k -NN method consists in considering the k variables X_i closest to x with respect to some distance $d(\cdot, \cdot)$. Although the local bandwidth of the k nearest neighbors is random and depends on the data X_i , the method has many advantages, mainly that of respecting the local structure of the data, essential in infinite dimension. It is commonly used in practice (see [Ferraty and Vieu \(2006\)](#)) and is simple to handle because

the user has only one parameter to control (the number k of nearest neighbors), this parameter k taking its values in a finite set. In addition, it allows to build a neighborhood adapted to the data at any point. This method of k -NN is a widely studied if the explanatory variable is an element of a finite-dimensional space, we can quote for example [Collomb \(1980\)](#), [Mack \(1981\)](#), [Györfi \(1981\)](#), [Bhattacharya and Mack \(1987\)](#) and [Devroye *et al.* \(1994\)](#). In an infinite dimensional space, i.e., a functional framework, there are three different approaches for the k -NN regression estimation. The first, published by [Laloë \(2008\)](#), examines a k -NN kernel estimate when the functional variable is an element of a separable Hilbert space H . In this approach, [Laloë \(2008\)](#) establishes a weak consistency result. The strategy of Laloë is to reduce the infinite dimension of H by using a projection on a finite dimension subspace which considering only the first m coefficients of an expansion of X in an orthonormal system of H , and then applying multivariate techniques on the projected data to perform the k -NN regression. The second approach, combines between the k -NN procedure and the functional local linear estimation to show the consistency and the rate of convergence (see e.g. [Attouch *et al.* \(2017\)](#), [Chikr-Elmezouar *et al.* \(2019\)](#) and [Ling *et al.* \(2020\)](#)). The third approach published by [Burba *et al.* \(2009\)](#), is rather based on a pure functional approach. For recent results in infinite dimensions (see e.g. [Ferraty and Vieu \(2006\)](#), [Cérou, Frédéric and Guyader, Arnaud \(2006\)](#) and [Chen and Shah \(2018\)](#)). [Burba *et al.* \(2009\)](#) extended a lemma of [Collomb \(1980\)](#), that we will also use in our proofs, and examine the problem on a semi-metric functional space. They proved almost complete convergence and almost complete convergence rates for independent data. [Lian \(2011\)](#) introduce the rates of almost sure convergence of the k -nearest neighbors estimate when the response is an element of a Hilbert space. The most recent results in k -NN nonparametric functional data analysis can be found in [Kudraszow and Vieu \(2013\)](#) and [Kara *et al.* \(2017\)](#) and references therein.

In this paper we focus in some more general correlation, that is the quasi-associated condition. This correlation can be viewed as a particular case of the weak dependence condition which allows treating the mixing and association correlation in a unified approach. For more references and properties of this correlation (see e.g. [Esary *et al.* \(1967\)](#), [Doukhan and Louhichi \(1999\)](#), [Bulinski and Suquet \(2001\)](#) and [Douge \(2010\)](#)). Furthermore, it should be noted that the dependence condition considered here allow to avoid the widely used strong mixing condition which is very easy to verified in practice.

In dependent functional data analysis, [Müller and Dippon \(2011\)](#) examines the k -nearest neighbours method in functional non-parametric regression for α -mixing data. They prove almost complete convergence and give the almost complete convergence rate of the k -NN kernel estimator. Recently, [Ling *et al.* \(2019\)](#) consider the nonparametric regression model and they established the uniform almost complete convergence rate of the k -NN estimator for strong mixing functional times series data.

In the present work we establish the almost complete convergence rates for the k -NN kernel single index estimator. This investigation is far from being trivial, which form a basically unsolved open problem in the literature. We aim at filling this gap in the literature by combining results [Burba *et al.* \(2009\)](#) with techniques handling the associated functional data given in [Kallabis and Neumann \(2006\)](#) and [Douge \(2010\)](#). However, as will be seen later, the problem requires much more than “simply” combining ideas from the existing results. In fact, delicate mathematical derivations will be required to cope with the regular functional data in our context. This requires the effective application of large sample theory techniques.

The rest of this work is organized as follows. In Section 4.2, we describe the nonparametric model and estimator of the k -NN. In Section 4.3, we will establish our main results of the almost complete convergence of the Nadaraya-Watson kernel estimator and k -NN method under some mild conditions. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the Section 4.4.

4.2 Model and estimator

Let the symbols $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the inner product and the associated norm on a Hilbert space H , respectively, and let $(e_p)_{p \geq 1}$ be a complete orthonormal system of H . To define the quasi-association dependence in a Hilbert space H we define, in a first, this dependence in the real valued random fields.

Definition 4.1 (see [Bulinski and Suquet \(2001\)](#)) A sequence $(X_n)_{n \in \mathbb{N}}$ of random variables is said to be quasi-associated, if for any disjoint subsets I and J of \mathbb{N} and all bounded Lipschitz functions $f_1 : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ and $f_2 : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$ satisfying :

$$|\text{Cov}(f_1(X_i, i \in I), f_2(X_j, j \in J))| \leq \text{Lip}(f_1) \text{Lip}(f_2) \sum_{i \in I} \sum_{j \in J} |\text{Cov}(X_i, X_j)|, \quad (4.1)$$

where $|I|$ denotes cardinality of a finite set I , and the Lipschitz of a function $f(\cdot)$ defined by

$$\text{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1},$$

with

$$\|(x_1, \dots, x_k)\|_1 = \sum_{k=1}^n |x_k|.$$

Definition 4.2 (see [Douge \(2010\)](#)) A sequence $(X_i)_{i \in \mathbb{N}}$ of random variables taking values in a Hilbert space H is called quasi-associated relative to an orthonormal basis $\{e_p : p \geq 1\}$ of H , if for any $p \geq 1$, $(\langle X_i, e_1 \rangle, \dots, \langle X_i, e_p \rangle)_{i \in \mathbb{N}}$ is a sequence of random vectors quasi-associated.

Now, let $(X_i, Y_i)_{i=1}^n$ be a quasi-associated random variables sequence identically distributed as (X, Y) , the latter being a random pair with values in the measurable space $(H \times \mathbb{R}, \mathcal{A}_d \otimes \mathcal{B}_{\mathbb{R}})$ where H is a separable real Hilbert space, (\mathcal{A}, d) is semi-metric space and \mathcal{A}_d is the σ -algebra generated by the topology of H that is defined by the semi-metric $d(\cdot, \cdot)$ ¹ and $\mathcal{B}_{\mathbb{R}}$ is the usual Borel, σ -algebra of \mathbb{R} . We consider the semi-metric $d_\theta(\cdot, \cdot)$ associated to the single-index $\theta \in H$ defined by $\forall u, v \in H$:

$$d_\theta(u, v) := |\langle \theta, u - v \rangle|.$$

We suppose that there exists a fixed $\theta \in \Theta \subset H$ where the observations $(X_i, Y_i)_{i=1, \dots, n}$ are related by the following relation

$$Y_i = r(\langle \theta, X_i \rangle) + \varepsilon_i, \quad \forall i = 1, \dots, n, \quad (4.2)$$

where $r(\cdot)$ is a real function, and for $i = 1, \dots, n$, ε_i is a real random variable such that

$$\mathbb{E}(\varepsilon_i | X_i) = 0.$$

The k -NN estimator of regression operator $r_\theta(\cdot) = r(\langle \theta, \cdot \rangle)$ is defined by

$$\hat{r}_{k, \theta}(x, H_{n, k}) = \frac{\sum_{i=1}^n Y_i K_i(x)}{\sum_{i=1}^n K_i(x)} \quad \text{for all } x \in H, \quad (4.3)$$

1. A semi-metric (sometimes called pseudo-metric) $d(\cdot, \cdot)$ is a metric which allows $d(x_1, x_2) = 0$ for some $x_1 \neq x_2$.

where $K_i(x) := K\left(\frac{\langle \theta, x - X_i \rangle}{H_{n,k}}\right)$ is a kernel function, $H_{n,k}$ is a random bandwidth depend on x and is defined as follows :

$$H_{n,k}(x) = \min \left\{ h_n \in \mathbb{R}^+ : \sum_{i=1}^n \mathbf{1}_{B(x, h_n)}(x_i) = k \right\}, \quad (4.4)$$

where $B(x, h_n)$ is a ball of center x with radius $h_n > 0$ defined as follows

$$B(x, h_n) = \{x' \in H, d_\theta(x, x') \leq h_n\},$$

and $\mathbf{1}_C$ denotes the indicator function of the set C .

4.3 Main result

To prove the almost complete convergence² of the k -NN kernel estimate, we need some results of the Nadaraya-Watson (NW) kernel estimate. Therefore, we present this kernel estimate in the next subsection.

4.3.1 Consistency of NW method

We establish the almost complete convergence of $\widehat{r}_\theta(h_n, x)$ to $r_\theta(x)$. To do that, we consider the following hypotheses, and we will denote by C_i with $i = 1, 2, \dots$ some strictly positive generic constants.

Assumptions and notation

(H₁) The kernel function $K(\cdot)$ is a bounded continuous Lipschitz function such that

$$C_1 \mathbf{1}_{[0,1]}(t) < K(t) < C_2 \mathbf{1}_{[0,1]}(t).$$

(H₂) Let $E_i(x) := \langle \theta, x - X_i \rangle$ so that $E_i(x)$ is a real-valued random variable,

$$G_{x,\theta}(h_n) = \mathbb{P}(|E_i(x)| \leq h_n) > 0.$$

(H₃) The joint distribution function of the couple (X_i, X_j) satisfy

$$\begin{aligned} 0 &< \sup_{i \neq j} \mathbb{P}\left((X_i, X_j) \in B(x, h_n) \times B(x, h_n)\right) \\ &= \max_{i \neq j} \left\{ \mathbb{P}\left(|E_i| \leq h_n\right), \mathbb{P}\left(|E_j| \leq h_n\right) \right\} = O(G_{x,\theta}^2(h_n)). \end{aligned}$$

2. We say that the sequence $(\Theta_n)_n$ converges a.co. to zero, if and only if

$$\forall \tau > 0, \sum_{n \geq 1} \mathbb{P}(|\Theta_n| > \tau) < \infty.$$

Furthermore, we say that $\Theta_n = O_{a.co.}(\theta_n)$, if there exists $\tau_0 > 0$, such that

$$\sum_{n \geq 1} \mathbb{P}(|\Theta_n| > \tau_0 \theta_n) < \infty.$$

(H₄) For all $u, v \in H$ for some $\alpha > 0$,

$$|r_\theta(u) - r_\theta(v)| \leq C_3 |\langle \theta, u - v \rangle|^\alpha.$$

(H₅) (i) $\mathbb{E}(\exp(|Y|)) \leq C_4 < \infty$,

(ii) $\mathbb{E}(|Y|^\nu | X = x) \leq C_5 < \infty$ for all $\nu \geq 2$,

(iii) $\forall i \neq j \mathbb{E}(|Y_i Y_j| | X_i, X_j) \leq C_6 < \infty$.

(H₆) Denote λ_r the covariance coefficient defined as :

$$\lambda_r := \sup_{s \geq r} \sum_{|i-j| \geq s} \lambda_{i,j},$$

where

$$\begin{aligned} \lambda_{i,j} &= \sum_{k \geq 1} \sum_{l \geq 1} \left| \text{Cov}(X_i^k, X_j^l) \right| \\ &\quad + \sum_{k \geq 1} \left| \text{Cov}(X_i^k, Y_j) \right| + \sum_{l \geq 1} \left| \text{Cov}(Y_i, X_j^l) \right| + |\text{Cov}(Y_i, Y_j)|, \end{aligned}$$

and $X_i^p := \langle X_i, e_p \rangle$. We assume that the covariance coefficient λ_r , $r \in \mathbb{N}$ satisfying :

$$\exists a > 0, \exists C_7 > 0, \text{ such that } \lambda_r \leq C_7 e^{-ar}.$$

Some comments on the assumptions of NW method. Assumption (H₁) is standard condition on the kernel $K(\cdot)$. The concentration property of the explanatory variable in small balls under single index topological structure is defined in the assumption (H₂). This assumption replaces the condition of a strict positivity of the density of explanatory variable usually assumed in finite-dimensional case. Assumption (H₃) gives the behavior of the joint distribution of $(E_i(x), E_j(x))$ and to establish the asymptotic normality of our model in the quasi-associated variables setting. The hypothesis (H₄) is used to control the regularity condition which characterizes the functional space of our model and is needed to evaluate the bias in the asymptotic results of this paper. Assumptions ((H₅) (i), (ii), (iii)) are a mild smoothness assumptions on the regression functional. The hypothesis (H₆) used for the quasi-associated data, more precisely in the control of the stochastic terms.

The main results for NW estimates satisfying the above assumptions are the following. Let us note the estimators $\hat{r}_{\theta,0}(h, x)$ and $\hat{r}_{\theta,1}(h, x)$ defined as follows

$$\begin{aligned} \hat{r}_{\theta,0}(h, x) &= \left(n \mathbb{E} [K(h^{-1} \langle \theta, x - X \rangle)] \right)^{-1} \sum_{i=1}^n K(h^{-1} \langle \theta, x - X_i \rangle), \\ \hat{r}_{\theta,1}(h, x) &= \left(n \mathbb{E} [K(h^{-1} \langle \theta, x - X \rangle)] \right)^{-1} \sum_{i=1}^n Y_i K(h^{-1} \langle \theta, x - X_i \rangle). \end{aligned}$$

Theorem 4.1 *Under the assumptions [(H₁) - (H₆)] and, in addition, if there exist $\gamma \in (0, 1)$ and $\xi_1, \xi_2 > 0$ such that*

$$\frac{\log n}{n^{1-\gamma-\xi_2}} \leq G_{x,\theta}(h_n) \leq \frac{C}{(\log n)^{1+\xi_1}}.$$

We have, as $n \rightarrow \infty$,

$$\hat{r}_\theta(h_n, x) - r_\theta(x) = O(h_n^\alpha) + O_{a.co.} \left(\sqrt{\frac{\log(n)}{n^{1-\gamma} G_{x,\theta}(h_n)}} \right) = O_{a.co.}(\mathcal{V}_n). \quad (4.5)$$

4.3.2 Consistency of k -NN method

In this subsection, we establish the almost complete convergence of k -NN kernel regression estimator $\widehat{r}_{k,\theta}(H_{n,k}, x)$ to $r_\theta(x)$. The main results for k -NN estimates satisfying the followings assumptions.

- (B₁) The derivative of $K(\cdot)$ exist on $[0, 1]$ with $-\infty < C'_1 < K'(t) < C'_2 < \infty$
 and $\forall t \in [0, 1]$ $K(1) > 0$ and the function $G_{x,\theta}(\cdot)$ has to fulfill the additional condition :

$$\exists C > 0, \exists \epsilon_0 > 0, \forall \epsilon < \epsilon_0 : \int_0^\epsilon G_{x,\theta}(u) du > C\epsilon G_{x,\theta}(\epsilon).$$

- (B₂) The concentration function $G_{x,\theta}(\cdot)$ is continuous and strictly increasing on a neighbourhood of 0 with $G_{x,\theta}(0) = 0$.

- (B₃) There exist a nonnegative function $\phi(\cdot)$ tends to zero as ϵ goes to zero, a bounded and positive function $\mathcal{L}(\cdot)$, and a function $\zeta_0(u) = u^b$, for some $\alpha > 0$, $\epsilon > 0$ and $b > 0$ such that

(i) $\frac{G_{x,\theta}(\epsilon)}{\phi(\epsilon)} - \mathcal{L}(x) = O(\epsilon^\alpha)$, as $\epsilon \rightarrow 0$,

(ii) for all $u > 0$, $\lim_{\epsilon \rightarrow 0} \frac{\phi(u\epsilon)}{\phi(\epsilon)} = \zeta_0(u)$,

(iii) $0 < C'_3\phi(\epsilon) \leq G_{x,\theta}(\epsilon) \leq C'_4\phi(\epsilon) < \infty$.

Some comments on the assumptions of k -NN method. Similar to (H₁), assumption (B₁) is standard condition on the kernel $K(\cdot)$. The second condition in hypothesis (B₁) allows us to control more precisely the behaviour of the small ball probability function $G_{x,\theta}(\epsilon)$ around zero, and that it is always possible to choose a semi-metric $d(\cdot, \cdot)$ for which this condition is fulfilled (see p.44 in [Ferraty and Vieu \(2006\)](#)). Assumptions (B₂) and (B₃) are the quite usual conditions on the small ball probability, (B₃(i)) shows that the small ball probability can be written approximately as the product of two independent function $\phi(\cdot)$ and $\mathcal{L}(\cdot)$ which has been employed by [Masry \(2005\)](#). The function $\zeta_0(u)$ defined in (B₃(ii)) intervenes in all asymptotic behaviour, with a simple algebra, it is possible to specify this function, for example, by u^b for some $b > 0$. For other examples we refer to the paper of [Ferraty et al. \(2007\)](#). The upper and lower bound on $\mathbb{P}(|\langle \theta, x - X_i \rangle| \leq h_n)$ of the form (B₃(iii)) is assumed by [Ferraty and Vieu \(2004\)](#), for controlling the concentration of the measure of probability of the functional variable on a small ball.

Theorem 4.2 *Under the assumptions of Theorem 4.1 and suppose that conditions [(B₁) - (B₃)] hold, in addition, we suppose that a sequence of positive real numbers $k = k_n$ satisfies $\frac{k}{n} \rightarrow 0$, there exist $\gamma \in (0, 1)$ such that $\frac{n^\gamma \log(n)}{k} \rightarrow 0$ when n tends to infinity and for all $\epsilon_1, \epsilon_2 > 0$ such that*

$$\frac{n \log(n)}{n^{1-\gamma-\epsilon_2}} \leq k \leq \frac{n}{(\log(n))^{1+\epsilon_1}}.$$

Then we have

$$\widehat{r}_{k,\theta}(H_{n,k}, x) - r_\theta(x) = O\left(\left(\phi^{-1}\left(\frac{k}{n}\right)\right)^\alpha\right) + O_{a.co.}\left(\sqrt{\frac{n^\gamma \log(n)}{k}}\right) = O_{a.co.}(\mathcal{U}_n). \quad (4.6)$$

4.4 Proofs of the intermediate results

A general Lemma. Let $(A_i, B_i)_{i=1,\dots,n}$ be a sequence of quasi-associated random variables identically distributed as (A, B) , which are valued in $(H \times \mathbb{R}, \mathcal{A} \times \mathcal{B}_\mathbb{R})$, where (H, \mathcal{A}) is a measurable space and $(\mathbb{R}, \mathcal{B}_\mathbb{R})$

is the usual one dimensional Borel measure space. Let $L(\cdot, \cdot) : \mathbb{R}^+ \times (H \times H) \longrightarrow \mathbb{R}^+$ be a function such that, $\forall x \in H, L(\cdot, (x, \cdot))$ is measurable and $\forall t_1, t_2 \in \mathbb{R}^+$ with $t_1 < t_2$,

$$L(t_1, z) \leq L(t_2, z) \text{ for } \forall z \in H \times H.$$

For all $x \in H$ and $n \geq 1$, we define

$$R_n(t) = \frac{\sum_{i=1}^n B_i L(t, (x, A_i))}{\sum_{i=1}^n L(t, (x, A_i))}. \quad (4.7)$$

Lemme 4.1 (See [Burba et al. \(2009\)](#)) Let $(D_n(x))_{n \geq 1}$ be a sequence of real random variable and $(\mathcal{U}_n)_{n \in \mathbb{N}}$ a decreasing positive sequence (with $\lim_{n \rightarrow \infty} \mathcal{U}_n = 0$), and $R(\cdot) : H \longrightarrow \mathbb{R}$ be a nonrandom function. If for all increasing sequence $\beta_n \in (0, 1)$ with limit 1 ($\beta_n - 1 = O(\mathcal{U}_n)$), there exist two sequences of real random variable $(D_n^-(\beta_n, x))_{n \geq 1}$ and $(D_n^+(\beta_n, x))_{n \geq 1}$ such that

C1. $\forall n \geq 1, \forall x \in H, D_n^-(\beta_n, x) \leq D_n^+(\beta_n, x)$ and

$$\mathbb{1}_{\{D_n^-(\beta_n, x) \leq D_n(x) \leq D_n^+(\beta_n, x)\}} \xrightarrow{a.co.} 1,$$

C2.

$$\frac{\sum_{i=1}^n L(D_n^-(\beta_n, x), (x, A_i))}{\sum_{i=1}^n L(D_n^+(\beta_n, x), (x, A_i))} - \beta_n = O_{a.co.}(\mathcal{U}_n),$$

C3.

$$R_{n,x}(D_n^-(\beta_n, x)) - R(x) = O_{a.co.}(\mathcal{U}_n),$$

and

$$R_{n,x}(D_n^+(\beta_n, x)) - R(x) = O_{a.co.}(\mathcal{U}_n).$$

Then, we have

$$R_{n,x}(D_n(x)) - R(x) = O_{a.co.}(\mathcal{U}_n). \quad (4.8)$$

Lemme 4.2 (See [Ezzahrioui and Ould Saïd \(2008\)](#)) Suppose that Assumptions (H_1) , (B_1) and (B_3) are holds. For all fixed x , we have

$$(\phi(h_n))^{-1} \mathbb{E} [K^j(h_n \langle \theta, x - X \rangle)] \longrightarrow v_j \mathcal{L}(x) \text{ as } n \longrightarrow \infty, \text{ for } j = 1, 2, \quad (4.9)$$

where, for $j=1, 2$,

$$v_j = K^j(1) - \int_0^1 (K^j(u))' \zeta_0(u) du.$$

Lemme 4.3 *Under the assumptions (H_1) - (H_6) , we have, as $n \rightarrow \infty$,*

$$|\widehat{r}_{\theta,0}(h_n, x) - 1| = O_{a.co.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(h_n)}} \right). \quad (4.10)$$

Furthermore, we have

$$\exists \eta > 0 \text{ such that } \sum_{i=1}^n \mathbb{P}(|\widehat{r}_{\theta,0}(h_n, x)| < \eta) < \infty. \quad (4.11)$$

Proof of Lemma 4.3.

The proof of (4.10) is based on inequality given by Theorem 2.1 in [Kallabis and Neumann \(2006\)](#) on the variables

$$\widehat{\Delta}_i(x) := \frac{1}{n\mathbb{E}(K_1(x))} [K_i(x) - \mathbb{E}(K_i)], \quad i = 1, \dots, n,$$

we have

$$\begin{aligned} \mathbb{E}(\widehat{\Delta}_i) &= 0, \\ \|\widehat{\Delta}_i\|_\infty &\leq \frac{2}{nG_{x,\theta}(h_n)} \|K\|_\infty, \\ \text{Lip}(\widehat{\Delta}_i) &\leq 2\text{Lip}(K) \frac{1}{nG_{x,\theta}(h_n)h_n}, \\ \widehat{r}_{\theta,0}(h_n, x) - \mathbb{E}(\widehat{r}_{\theta,0}(h_n, x)) &= \sum_{i=1}^n \widehat{\Delta}_i(x). \end{aligned}$$

We start by evaluating the covariance term $\text{Cov}(\widehat{\Delta}_{s_1}, \dots, \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1}, \dots, \widehat{\Delta}_{t_v})$, for all $(s_1, \dots, s_u) \in \mathbb{N}^u$ and $(t_1, \dots, t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$. If $m = t_1 - s_u = 0$, using the fact that

$$\mathbb{E}(K_1) = O(G_{x,\theta}(h_n)),$$

and under (H_5) , we readily obtain

$$\begin{aligned} \left| \text{Cov}(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v}) \right| &\leq \left(\frac{1}{n\mathbb{E}(K_1)} \right)^{u+v} \mathbb{E}(|K_{s_1} \dots K_{s_u}^2 \dots K_{t_v}|) \\ &\leq \left(\frac{\|K\|_\infty}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \mathbb{E}(K_{s_u}^2) \\ &\leq \left(\frac{C}{nG_{x,\theta}(h_n)} \right)^{u+v} G_{x,\theta}(h_n). \end{aligned}$$

If $m = t_1 - s_u > 0$, by quasi-association of the variables $K_i(x) = K\left(\frac{\langle \theta, x - X_i \rangle}{h_n}\right)$ in connection with (H_6) , we infer that

$$\begin{aligned} &\left| \text{Cov}(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v}) \right| \\ &\leq 4 \left(\frac{\text{Lip}(K)}{nG_{x,\theta}(h_n)h_n} \right)^2 \left(\frac{2\|K\|_\infty}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \sum_{i=1}^u \sum_{j=1}^v \lambda_{s_i, t_j} \end{aligned}$$

$$\begin{aligned}
&\leq C^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \left(\frac{1}{nG_{x,\theta}(h_n)} \right)^{u+v} (u \wedge v) \lambda_{t_1 - s_u} \\
&\leq C^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \left(\frac{1}{nG_{x,\theta}(h_n)} \right)^{u+v} v e^{-am}.
\end{aligned} \tag{4.12}$$

Making use of the condition (H_3) , it follows that

$$\begin{aligned}
&\left| \text{Cov} \left(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v} \right) \right| \\
&\leq \left(\frac{\|K\|_\infty}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \left| \text{Cov} \left(\widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \right) \right| \\
&\leq \left(\frac{\|K\|_\infty}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \left(\left| \mathbb{E} \left(\widehat{\Delta}_{s_u} \widehat{\Delta}_{t_1} \right) \right| + \mathbb{E} \left| \widehat{\Delta}_{s_u} \right| \mathbb{E} \left| \widehat{\Delta}_{t_1} \right| \right) \\
&\leq \left(\frac{\|K\|_\infty}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \left(\frac{C}{nG_{x,\theta}(h_n)} \right)^2 M,
\end{aligned}$$

where

$$M = \sup_{i \neq j} \mathbb{P} \{ (X_i, X_j) \in B(x, h_n) \times B(x, h_n) \} + \{ \mathbb{P}^2 (X_1 \in B(x, h_n)) \}.$$

Then, we obtain a second bound of $\text{Cov}(\cdot, \cdot)$

$$\left| \text{Cov} \left(\widehat{\Delta}_{s_1}, \dots, \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1}, \dots, \widehat{\Delta}_{t_v} \right) \right| \leq C^{u+v} \left(\frac{1}{nG_{x,\theta}(h_n)} \right)^{u+v} G_{x,\theta}^2(h_n). \tag{4.13}$$

Moreover, by multiplying a τ -power of (4.12) and $(1 - \tau)$ -power of (4.13) for some $\frac{1}{4} < \tau < \frac{1}{2}$ we obtain an upper-bound of the covariance as follows for $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$

$$\begin{aligned}
&\left| \text{Cov} \left(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v} \right) \right| \\
&\leq C^{u+v} \left(\frac{1}{nG_{x,\theta}(h_n)} \right)^{u+v} \left(\frac{\text{Lip}(K)}{h_n} \right)^{2\tau} G_{x,\theta}^{2(1-\tau)}(h_n) v e^{-a\tau m}.
\end{aligned}$$

So, by the inequality given in Theorem 2.1 in [Kallabis and Neumann \(2006\)](#), we have

$$\begin{aligned}
&\left| \text{Cov} \left(\widehat{\Delta}_{s_1} \dots \widehat{\Delta}_{s_u}, \widehat{\Delta}_{t_1} \dots \widehat{\Delta}_{t_v} \right) \right| \\
&\leq \left(\frac{C}{nG_{x,\theta}(h_n)} \right)^{u+v-2} \left(\frac{C}{nG_{x,\theta}(h_n)} \right)^2 G_{x,\theta}(h_n) v e^{-a\tau m},
\end{aligned} \tag{4.14}$$

where $M_n = \frac{C}{nG_{x,\theta}(h_n)}$ and $K_n = \frac{C}{n\sqrt{G_{x,\theta}(h_n)}}$. It remains to calculate $\text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right)$:

$$\begin{aligned}
\text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right) &= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 \sum_i \sum_j \text{Cov}(K_i, K_j) \\
&= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 \left[n\text{Var}(K_1) + \sum_i \sum_{j \neq i} \text{Cov}(K_i, K_j) \right] \\
&= \left(\frac{1}{n\mathbb{E}(K_1(x))} \right)^2 [nT_1 + T_{ij}].
\end{aligned} \tag{4.15}$$

Now, under (H_5) and for all $j \geq 1$, $\mathbb{E} \left(K_1^j(x) \right) = O \left(G_{x,\theta}(h_n) \right)$, we obtain for the first term of (4.15)

$$T_1 = \text{Var} (K_1) = O \left(G_{x,\theta}(h_n) \right). \quad (4.16)$$

We readily obtain that

$$\begin{aligned} \frac{1}{n \left(\mathbb{E} (K_1(x)) \right)^2} T_1 &\leq \frac{C G_{x,\theta}(h_n)}{n O \left(G_{x,\theta}^2(h_n) \right)} \\ &\leq \frac{C}{n G_{x,\theta}(h_n)}. \end{aligned} \quad (4.17)$$

For the second term of (4.15), we have the following decomposition

$$\begin{aligned} T_{ij} &= \sum_i \sum_{0 < |i-j| \leq u_n} \text{Cov} (K_i, K_j) + \sum_i \sum_{|i-j| > u_n} \text{Cov} (K_i, K_j) \\ &= J_1 + J_2, \end{aligned} \quad (4.18)$$

where (u_n) is a sequence of positive integer and

$$\lim_{n \rightarrow \infty} u_n = \infty.$$

Now, under the assumptions (H_2) , (H_3) and (H_5) , we have

$$\begin{aligned} |J_1| &= \sum_i \sum_{0 < |i-j| \leq u_n} |\text{Cov} (K_i, K_j)| \\ &\leq n u_n \left[\max_{i \neq j} |\mathbb{E} (K_i K_j)| + \left(\mathbb{E} (K_1) \right)^2 \right] \\ &\leq C n u_n G_{x,\theta}^2(h_n). \end{aligned} \quad (4.19)$$

Making use of the condition (H_6) , we infer that

$$\begin{aligned} |J_2| &= \sum_i \sum_{|i-j| > u_n} |\text{Cov} (K_i, K_j)| \leq C \left(\frac{\text{Lip}(K)}{h_n} \right)^2 \sum_i \sum_{|i-j| > u_n} \lambda_{i,j} \\ &\leq C n h_n^{-2} e^{-a u_n}. \end{aligned} \quad (4.20)$$

This implies that

$$\begin{aligned} |T_{ij}| &\leq \sum_{i=1}^n \sum_{i \neq j} |\text{Cov} (K_i, K_j)| \\ &\leq C \left(n u_n G_{x,\theta}^2(h_n) + n h_n^{-2} e^{-a u_n} \right). \end{aligned} \quad (4.21)$$

We then obtain that

$$\frac{1}{n G_{x,\theta}(h_n)} |T_{ij}| \leq C g(u_n), \quad (4.22)$$

where

$$g(u_n) := u_n G_{x,\theta}(h_n) + G_{x,\theta}^{-1}(h_n) h_n^{-2} e^{-a u_n}.$$

Taking $u_n = \frac{1}{a} \log \left(\frac{a}{h_n^2 G_{x,\theta}^2(h_n)} \right)$ with $a > 0$. From (4.15), (4.18), (4.22) and (H7) we get

$$\frac{1}{nG_{x,\theta}(h_n)} |T_{ij}| \longrightarrow 0, \text{ as } n \longrightarrow \infty. \quad (4.23)$$

Then, we get from (4.15), (4.17) and (4.23)

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right) &= \left[O \left(\frac{1}{nG_{x,\theta}(h_n)} \right) + \frac{1}{nG_{x,\theta}(h_n)} o(1) \right] \\ &= O \left(\frac{1}{nG_{x,\theta}(h_n)} \right). \end{aligned}$$

The conditions of Theorem 2.1 in Kallabis and Neumann (2006) are verified for

$$\begin{aligned} K_n &= \frac{C}{n\sqrt{G_{x,\theta}(h_n)}}, \\ M_n &= \frac{C}{nG_{x,\theta}(h_n)}, \\ A_n &= \text{Var} \left(\sum_{i=1}^n \widehat{\Delta}_i \right) = O \left(\frac{1}{nG_{x,\theta}(h_n)} \right), \\ B_n &= \left(\frac{16nK_n^2}{9A_n(1-e^{-\beta})} \vee 1 \right) \left(\frac{2(K_n \vee M)}{1-e^{-\beta}} \right) = \frac{1}{nG_{x,\theta}(h_n)}. \end{aligned}$$

So, we apply the inequality in Kallabis and Neumann (2006) to the random variables $\widehat{\Delta}_i$ to infer that

$$\begin{aligned} &\mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x) - \mathbb{E}(\widehat{r}_{\theta,0}(h_n, x))| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(h_n)}} \right) \\ &= \mathbb{P} \left(\left| \sum_{i=1}^n \widehat{\Delta}_i \right| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(h_n)}} \right) \\ &\leq \exp \left(\frac{-\varepsilon^2 \log n}{n^{1-\gamma} G_{x,\theta}(h_n) L_\theta(n)} \right), \end{aligned}$$

where

$$L_\theta(n) = \left(O \left(\frac{1}{nG_{x,\theta}(h_n)} \right) + \left(\frac{1}{nG_{x,\theta}(h_n)} \right)^{\frac{1}{3}} \left(\frac{\log n}{n^{1-\gamma} G_{x,\theta}(h_n)} \right)^{\frac{5}{6}} \right).$$

We finally obtain that

$$\begin{aligned} &\mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x) - \mathbb{E}(\widehat{r}_{\theta,0}(h_n, x))| > \varepsilon \sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(h_n)}} \right) \\ &\leq \exp \left(\frac{-\varepsilon^2 \log n}{C + \left(\frac{n^{-\gamma} \log^5 n}{nG_{x,\theta}(h_n)} \right)^{\frac{1}{6}}} \right) \\ &\leq C_1 \exp(-\varepsilon^2 \log(n)). \end{aligned}$$

The proof of (4.10) is achieved by a suitable choice of ε .

Proof of (4.11). Notice that we have

$$\left\{ |\widehat{r}_{\theta,0}(h_n, x)| \leq \frac{1}{2} \right\} \subset \left\{ |\widehat{r}_{\theta,0}(h_n, x) - 1| > \frac{1}{2} \right\},$$

that implies that

$$\mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x)| \leq \frac{1}{2} \right) \leq \mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x) - 1| > \frac{1}{2} \right).$$

Under the hypothesis $(H_1) - (H_6)$ and by applying the equation (??), we deduce that

$$\sum_n \mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x)| \leq \frac{1}{2} \right) \leq \sum_n \mathbb{P} \left(|\widehat{r}_{\theta,0}(h_n, x) - 1| > \frac{1}{2} \right) < \infty.$$

Then, for $\eta = \frac{1}{2}$, we have

$$\sum_n \mathbb{P} (|\widehat{r}_{\theta,0}(h_n, x)| \leq \eta) < \infty.$$

Thus the proof of Lemma 4.3 is complete. □

Proof of Theorem 4.2

In order to demonstrate the consistency of $\widehat{r}_{k,\theta}(H_{n,k}, x)$ presented in Theorem 4.2, we need to check the conditions (C1.), (C2.) and (C3.) of Lemma 4.1. For that, we first identify the variables as follows : $A_i = X_i$, $B_i = Y_i$,

$$\begin{aligned} L(T, (x, A_i)) &= K(T^{-1}(\theta, x - X_i)), \\ D_n(x) &= H_{n,k}(x), \\ R_{n,x}(H_{n,k}(x)) &= \widehat{r}_{k,\theta}(H_{n,k}, x), \\ R(x) &= r_\theta(x). \end{aligned}$$

Choosing $D_n^-(\beta_n, x)$ and $D_n^+(\beta_n, x)$ such that

$$G_{x,\theta}(D_n^-(\beta_n, x)) = G_{x,\theta}(h_n(x))\beta_n^{\frac{1}{2}}, \tag{4.24}$$

and

$$G_{x,\theta}(D_n^+(\beta_n, x)) = G_{x,\theta}(h_n(x))\beta_n^{-\frac{1}{2}}. \tag{4.25}$$

Let we put $h_n(x) = \phi^{-1} \left(\frac{k}{n} \right)$ and $\beta_n \in (0, 1)$ be an increasing sequence such that $\beta_n - 1 = O(\mathcal{U}_n)$.

Checking the condition (C3.) of Lemma 4.1

To show the first convergence of the condition (4.1), we can use Theorem 4.1 for $D_n^- = D_n^-(\beta_n, x)$ and $D_n^+ = D_n^+(\beta_n, x)$. The fact that β_n is bounded by 1, and the local bandwidth D_n^- satisfies together with the

fixed bandwidth h_n the conditions of Theorem 4.1, then we have

$$\begin{aligned}
 R_{n,x}(D_n^-(\beta_n, x)) - R(x) &= \frac{\sum_{i=1}^n Y_i K((D_n^-)^{-1}(\beta_n, x)\langle\theta, x - X_i\rangle)}{\sum_{i=1}^n K((D_n^-)^{-1}(\beta_n, x)\langle\theta, x - X_i\rangle)} - r_\theta(x) \\
 &= \widehat{r}_\theta(D_n^-(\beta_n, x), x) - r_\theta(x) \\
 &= O_{a.co.} \left(\left(\phi^{-1} \left(\frac{k}{n} \right) \right)^\alpha + \sqrt{\frac{n^\gamma \log(n)}{k}} \right) \\
 &= O_{a.co.}(\mathcal{U}_n). \tag{4.26}
 \end{aligned}$$

For the second convergence of the condition (C3.), we will use the same natural reasoning to the case $D_n^+(\beta_n, x)$, we readily obtain

$$R_{n,x}(D_n^+(\beta_n, x)) - R(x) = O_{a.co.}(\mathcal{U}_n). \tag{4.27}$$

Hence the condition (C3.) is checked.

Checking the condition (C2.) of Lemma 4.1

We consider the following quantities :

$$\begin{aligned}
 \mu_\theta(h_n, x) &:= \mathbb{E} [K(h_n^{-1}\langle\theta, x - X\rangle)], \\
 L_{n1}(x) &:= \frac{\mu_\theta(D_n^-, x)}{\mu_\theta(D_n^+, x)}, \\
 L_{n2}(x) &:= \frac{\widehat{r}_{\theta,0}(D_n^-, x)}{\widehat{r}_{\theta,0}(D_n^+, x)} - 1, \\
 L_{n3}(x) &:= \frac{\mu_\theta(D_n^+, x)}{\mu_\theta(D_n^-, x)} \beta_n - 1.
 \end{aligned}$$

The condition (C2.) can be written as follows :

$$\begin{aligned}
 L_n(x) - \beta_n &= \frac{\mu_\theta(D_n^-, x) \widehat{r}_{\theta,0}(D_n^-, x)}{\mu_\theta(D_n^+, x) \widehat{r}_{\theta,0}(D_n^+, x)} - \beta_n \\
 &= \frac{\mu_\theta(D_n^-, x)}{\mu_\theta(D_n^+, x)} \left(\frac{\widehat{r}_{\theta,0}(D_n^-, x)}{\widehat{r}_{\theta,0}(D_n^+, x)} - 1 \right) + \frac{\mu_\theta(D_n^-, x)}{\mu_\theta(D_n^+, x)} \\
 &\quad - \frac{\mu_\theta(D_n^-, x) \mu_\theta(D_n^+, x)}{\mu_\theta(D_n^+, x) \mu_\theta(D_n^-, x)} \beta_n \\
 &= L_{n1}(x) L_{n2}(x) - L_{n1}(x) L_{n3}(x). \tag{4.28}
 \end{aligned}$$

Hence

$$|L_n(x) - \beta_n| \leq |L_{n1}| (|L_{n2}(x)| + |L_{n3}(x)|), \tag{4.29}$$

where

$$L_n(x) = \frac{\sum_{i=1}^n K((D_n^-)^{-1}(\beta_n, x), (x, X_i))}{\sum_{i=1}^n K((D_n^+)^{-1}(\beta_n, x), (x, X_i))}.$$

Hence, by the fact that $\beta_n \rightarrow 1$, our claimed result is

$$|L_n(x) - \beta_n| = O_{a.co.}(\mathcal{U}_n). \quad (4.30)$$

The proof of (4.30) is based on the following three results

$$|L_{n1}(x)| \leq C, \quad (4.31)$$

$$|L_{n2}(x)| = O_{a.co.} \left(\sqrt{\frac{n^\gamma \log(n)}{k}} \right), \quad (4.32)$$

$$|L_{n3}(x)| = O \left(\left(\beta_n^{\frac{1}{2}} \phi^{-1} \left(\frac{k}{n} \right) \right)^\alpha \right). \quad (4.33)$$

Proof of (4.31). Using the equations (4.24), (4.25) and suppose that the assumptions (H_1) and (B_1) are holds and Lemma 1 in Ezzahrioui and Ould Saïd (2008), we have

$$\begin{aligned} |L_{n1}(x)| &= \left| \frac{\mu_\theta(D_n^-, x)}{\mu_\theta(D_n^+, x)} \right| \\ &= \left| \frac{(\phi(D_n^-))^{-1} \mathbb{E} [K((D_n^-)^{-1}(\theta, x - X))]}{(\phi(D_n^+))^{-1} \mathbb{E} [K((D_n^+)^{-1}(\theta, x - X))]} \right| \frac{\phi(D_n^-)}{\phi(D_n^+)} \\ &\leq C. \end{aligned} \quad (4.34)$$

Proof of (4.32). Remark that we have

$$\begin{aligned} |L_{n2}(x)| &= \left| \frac{\widehat{r}_{\theta,0}(D_n^-, x)}{\widehat{r}_{\theta,0}(D_n^+, x)} - 1 \right| \\ &\leq \frac{1}{|\widehat{r}_{\theta,0}(D_n^+, x)|} \left| \widehat{r}_{\theta,0}(D_n^-, x) - 1 \right| + \left| \widehat{r}_{\theta,0}(D_n^+, x) - 1 \right|. \end{aligned} \quad (4.35)$$

The same choice of D_n^- and D_n^+ as above, by the assumptions of Lemma 4.10 and (B_3) , we have

$$\begin{aligned} |\widehat{r}_{\theta,0}(D_n^-, x) - 1| &= O_{a.co.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(D_n^-)}} \right) \\ &= O_{a.co.} \left(\sqrt{\frac{n^\gamma \log n}{k \beta_n^{\frac{1}{2}}}} \right), \end{aligned} \quad (4.36)$$

and

$$\begin{aligned} |\widehat{r}_{\theta,0}(D_n^+, x) - 1| &= O_{a.co.} \left(\sqrt{\frac{\log n}{n^{1-\gamma} G_{x,\theta}(D_n^+)}} \right) \\ &= O_{a.co.} \left(\sqrt{\frac{\beta_n^{\frac{1}{2}} n^\gamma \log n}{k}} \right). \end{aligned} \quad (4.37)$$

Hence, by the fact that $\beta_n \rightarrow 1$ and combining (4.11) and ((4.35)-(4.37)), it follows that

$$|L_{n2}(x)| = O_{a.co.} \left(\sqrt{\frac{n^\gamma \log n}{k}} \right). \quad (4.38)$$

Proof of (4.33). By (H_1) , (B_1) and (B_3) and following the proof of Lemma 1 in Ezzahrioui and Ould Saïd (2008), we can easily see by $(B_3(i))$ that for some $\delta > 0$

$$\begin{aligned} \mu_\theta(h_n, x) &= \phi(h_n)\delta\mathcal{L}(x) + O(\phi(h_n)h_n^\alpha) \\ &= \delta G_{x,\theta}(h_n) + O(\phi(h_n)h_n^\alpha), \forall x \in H. \end{aligned} \quad (4.39)$$

In view of the fact that $\frac{G_{x,\theta}(D_n^-)}{G_{x,\theta}(D_n^+)} = \beta_n$ and $(B_3[iii])$ one gets

$$\begin{aligned} |L_{n3}(x)| &:= \left| \frac{\mu_\theta(D_n^+, x)}{\mu_\theta(D_n^-, x)} \beta_n - 1 \right| \\ &= \left| \frac{\delta G_{x,\theta}(D_n^+) + O(\phi(D_n^+)(D_n^+)^\alpha) G_{x,\theta}(D_n^-)}{\delta G_{x,\theta}(D_n^-) + O(\phi(D_n^-)(D_n^-)^\alpha) G_{x,\theta}(D_n^+)} - 1 \right| \\ &= \left| \frac{\delta + \frac{O(\phi(D_n^+)(D_n^+)^\alpha)}{G_{x,\theta}(D_n^+)}}{\delta + \frac{O(\phi(D_n^-)(D_n^-)^\alpha)}{G_{x,\theta}(D_n^-)}} - 1 \right| \\ &\leq \left| \frac{\delta + O((D_n^+)^\alpha)}{\delta + O((D_n^-)^\alpha)} - 1 \right| \\ &\leq \left| \frac{O((D_n^+)^\alpha) - O((D_n^-)^\alpha)}{\delta + O((D_n^-)^\alpha)} \right| \\ &\leq C \left| \frac{O((D_n^-)^\alpha)}{\delta + O((D_n^-)^\alpha)} \right| \\ &= O((D_n^-)^\alpha). \end{aligned} \quad (4.40)$$

Hence, by the fact that $h_n(x) = \phi^{-1}\left(\frac{k}{n}\right)$ and $D_n^- = h_n(x)\beta_n^{\frac{1}{2}}$,

$$|L_{n3}(x)| = O\left(\left(\beta_n^{\frac{1}{2}}\phi^{-1}\left(\frac{k}{n}\right)\right)^\alpha\right). \quad (4.41)$$

Thus, the condition (C2.) of Lemma 4.1 is verified because $\beta_n \rightarrow 1$ and the results (4.29) and ((4.34)-(4.41)) which implies

$$\left| \frac{\sum_{i=1}^n K((D_n^-)^{-1}(\beta_n, x), (x, X_i))}{\sum_{i=1}^n K((D_n^+)^{-1}(\beta_n, x), (x, X_i))} - \beta_n \right| = O_{a.co.}(\mathcal{U}_n).$$

Checking the condition (C1.) of Lemma 4.1

Using the assumptions (H_2) , (B_2) and by the equations (4.24) and (4.25), and the fact that β_n is an increasing sequence tends to 1, we show

$$D_n^-(\beta_n, x) \leq H_{n,k}(x) \leq D_n^+(\beta_n, x), \quad (4.42)$$

and

$$\forall \epsilon > 0, \sum_{n \geq 1} \mathbb{P} \left(\left| \mathbb{1}_{D_n^-(\beta_n, x) \leq H_{n,k}(x) \leq D_n^+(\beta_n, x)} - 1 \right| > \epsilon \right) < \infty. \quad (4.43)$$

Indeed, by (4.24) and (4.25) we have, $G_{x,\theta}(D_n^-) \leq G_{x,\theta}(h_n) \leq G_{x,\theta}(D_n^+)$. Then, by the equation (4.4) and the property of $G_{x,\theta}(\cdot)$, it follows that

$$D_n^-(\beta_n, x) \leq H_{n,k}(x) \leq D_n^+(\beta_n, x).$$

Proof of (4.43). In the case of associates random variables, we cannot use the Chernoff-type exponential inequality to check the condition (C1.). Instead we use the exponential inequality in Douge (2007).

Let $\varepsilon > 0$ be fixed. We have

$$\mathbb{P} \left(\left| \mathbb{1}_{D_n^- \leq H_{n,k} \leq D_n^+} - 1 \right| > \varepsilon \right) \leq \mathbb{P} (H_{n,k} < D_n^-) + \mathbb{P} (H_{n,k} > D_n^+). \quad (4.44)$$

For the two terms in (4.44), we mention that, for $i = 1, \dots, n$,

$$\Lambda_{x,\theta}(h_n, X_i) = \mathbb{1}_{B(x, h_n)}(X_i) - G_{x,\theta}(h_n)$$

are binary and centred associates random variables and we write

$$\begin{aligned} \mathbb{P} (H_{n,k} < D_n^-) &\leq \mathbb{P} \left(\sum_{i=1}^n \mathbb{1}_{B(x, D_n^-)}(X_i) > k \right) \\ &\leq \mathbb{P} \left(\sum_{i=1}^n \left(\mathbb{1}_{B(x, D_n^-)}(X_i) - G_{x,\theta}(D_n^-) \right) > k - nG_{x,\theta}(D_n^-) \right) \\ &\leq \mathbb{P} \left(\sum_{i=1}^n \Lambda_{x,\theta}(D_n^-, X_i) > k - nG_{x,\theta}(D_n^-) \right) \\ &=: A_n, \end{aligned} \quad (4.45)$$

and

$$\begin{aligned} \mathbb{P} (H_{n,k} > D_n^+) &\leq \mathbb{P} \left(\sum_{i=1}^n \mathbb{1}_{B(x, D_n^+)}(X_i) < k \right) \\ &\leq \mathbb{P} \left(\sum_{i=1}^n \left(\mathbb{1}_{B(x, D_n^+)}(X_i) - G_{x,\theta}(D_n^+) \right) < k - nG_{x,\theta}(D_n^+) \right) \\ &\leq \mathbb{P} \left(\sum_{i=1}^n \Lambda_{x,\theta}(D_n^+, X_i) < k - nG_{x,\theta}(D_n^+) \right) \\ &=: B_n. \end{aligned} \quad (4.46)$$

Let us introduce the following sum of random variables

$$\mathbb{S}_n = \sum_{i=1}^n \Lambda_{x,\theta}(h_n, X_i),$$

From Lemma 2 in [Lehmann \(1966\)](#) we get

$$\text{Cov}(\mathbb{1}_{B(x,h)}(X_i), \mathbb{1}_{B(x,h)}(X_j)) \leq C \text{Cov}(X_i, X_j). \quad (4.47)$$

By applying Theorem 2.1 in [Douge \(2007\)](#) for bounded associated variables $\mathbb{1}_{B(x,h)}(X_1), \mathbb{1}_{B(x,h)}(X_2), \dots, \mathbb{1}_{B(x,h)}(X_n)$, we obtain for all $n \geq 2$ and $\epsilon > 6/\sqrt{n}$

$$\begin{aligned} A_n &= \mathbb{P} \left(\frac{1}{n} \sum_{i=1}^n \Lambda_{x,\theta}(D_n^-, X_i) > \frac{k}{n} - G_{x,\theta}(D_n^-) \right) \\ &\leq 4C_0 \exp \left(-\frac{\theta \wedge 1}{12} \sqrt{n} \left(\frac{k}{n} - G_{x,\theta}(D_n^-) \right) \right), \end{aligned} \quad (4.48)$$

where $C_0 = \exp \left(\frac{\theta_0}{4(1-e^{-\theta})} \right)$, $\theta > 0$ and $\theta_0 > 0$.

By similar arguments as for A_n , we obtain for all $n \geq 2$ and $\epsilon > 6/\sqrt{n}$

$$\begin{aligned} B_n &= \mathbb{P} \left(\frac{1}{n} \sum_{i=1}^n \Lambda_{x,\theta}(D_n^+, X_i) < - \left(G_{x,\theta}(D_n^+) - \frac{k}{n} \right) \right) \\ &\leq \mathbb{P} \left(\frac{1}{n} \left| \sum_{i=1}^n \Lambda_{x,\theta}(D_n^+, X_i) \right| > G_{x,\theta}(D_n^+) - \frac{k}{n} \right) \\ &\leq 8C_0 \exp \left(-\frac{\theta \wedge 1}{12} \sqrt{n} \left(G_{x,\theta}(D_n^+) - \frac{k}{n} \right) \right). \end{aligned} \quad (4.49)$$

Thus, by ((4.44)-(4.49)), the first condition of Lemma 4.1 is verified. Finally, by Lemma 4.1, the proof of Theorem 4.2 is completed. \square

Bibliographie

- Attouch, M., Laksaci, A., and Rafea, F. (2017). Estimation locale linéaire de la régression non paramétrique fonctionnelle par la méthode des k plus proches voisins. *Comptes Rendus Mathématique*, **355**(7), 824 – 829.
- Bhattacharya, P. K. and Mack, Y. P. (1987). Weak convergence of k -nn density and regression estimators with varying k and applications. *Ann. Statist.*, **15**(3), 976–994.
- Bulinski, A. and Suquet, C. (2001). Normal approximation for quasi-associated random fields. *Statistics & Probability Letters*, **54**, 215–226.
- Burba, F., Ferraty, F., and Vieu, P. (2009). k -nearest neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics*, **21**(4), 453–469.
- Cérou, Frédéric and Guyader, Arnaud (2006). Nearest neighbor classification in infinite dimension. *ESAIM : PS*, **10**, 340–355.
- Chen, G. H. and Shah, D. (2018). Explaining the success of nearest neighbor methods in prediction. *Foundations and Trends in Machine Learning*, **10**(5-6), 337–588.
- Chikr-Elmezouar, Z., Almanjahie, I. M., Laksaci, A., and Rachdi, M. (2019). Fda : strong consistency of the k nn local linear estimation of the functional conditional density and mode. *Journal of Nonparametric Statistics*, **31**(1), 175–195.
- Collomb, G. (1980). Estimation de la régression par la méthode des k points les plus proches avec noyau : quelques propriétés de convergence ponctuelle. In *Nonparametric asymptotic statistics (Proc. Conf., Rouen, 1979) (French)*, volume 821 of *Lecture Notes in Math.*, pages 159–175. Springer, Berlin.
- Devroye, L., Györfi, L., Krzyzak, A., and Lugosi, G. (1994). On the strong universal consistency of nearest neighbor regression function estimates. *Ann. Statist.*, **22**(3), 1371–1385.
- Douge, L. (2007). Vitesses de convergence dans la loi forte des grands nombres et dans l’estimation de la densité pour des variables aléatoires associées. *Comptes Rendus Mathématique*, **344**(8), 515 – 518.
- Douge, L. (2010). Théorèmes limites pour des variables quasi-associées hilbertiennes. *Ann. I.S.U.P.*, **54**(1-2), 51–60.
- Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, **84**(2), 313–342.

- Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.*, **38**, 1466–1474.
- Ezzahrioui, M. and Ould Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *Journal of Nonparametric Statistics*, **20**, 3–18.
- Ferraty, F. and Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist.*, **17**(4), 545–564.
- Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination. volume 16, pages 111–125. The International Conference on Recent Trends and Directions in Nonparametric Statistics.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Peuch, A., and Vieu, P. (2003). Modèle à indice fonctionnel simple. *C. R. Math. Acad. Sci. Paris*, **336**(12), 1025–1028.
- Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data : Inference and practical aspects. *Australian & New Zealand Journal of Statistics*, **49**, 267 – 286.
- Ferraty, F., Park, J., and Vieu, P. (2011). Estimation of a functional single index model. In *Recent advances in functional data analysis and related topics*, Contrib. Statist., pages 111–116. Physica-Verlag/Springer, Heidelberg.
- Györfi, L. (1981). The rate of convergence of k-nn regression estimation and classification. *IEEE Trans. Inform. Theory.*, **27**(3), 500–509.
- Härdle, W., Hall, P., and Ichimura, H. (1993). Optimal smoothing in single-index models. *Ann. Statist.*, **21**(1), 157–178.
- Hristache, M., Juditsky, A., and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Ann. Statist.*, **29**(3), 595–623.
- Kallabis, R. S. and Neumann, M. H. (2006). An exponential inequality under weak dependence. *Bernoulli*, **12**(2), 333–350.
- Kara, L.-Z., Laksaci, A., Rachdi, M., and Vieu, P. (2017). Data-driven knn estimation in nonparametric functional data analysis. *Journal of Multivariate Analysis*, **153**, 176 – 188.
- Kudraszow, N. L. and Vieu, P. (2013). Uniform consistency of knn regressors for functional variables. *Statistics & Probability Letters*, **83**(8), 1863 – 1870.
- Laloë, T. (2008). A k-nearest neighbor approach for functional regression. *Statistics & Probability Letters*, **78**(10), 1189 – 1193.
- Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, **37**(5), 1137–1153.
- Lian, H. (2011). Convergence of functional k -nearest neighbor regression estimate with functional responses. *Electron. J. Stat.*, **5**, 31–40.

- Ling, N., Meng, S., and Vieu, P. (2019). Uniform consistency rate of knn regression estimation for functional time series data. *Journal of Nonparametric Statistics*, **31**(2), 451–468.
- Ling, N., Aneiros, G., and Vieu, P. (2020). knn estimation in functional partial linear modeling. *Statistical Papers*, **61**(1), 423–444.
- Mack, Y. P. (1981). Local properties of k-nn regression estimates. *SIAM Journal on Algebraic Discrete Methods*, **2**(3), 311–323.
- Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stochastic Process. Appl.*, **115**(1), 155–177.
- Müller, S. and Dippon, J. (2011). k-nn kernel estimate for nonparametric functional regression in time series analysis.
- Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer Series in Statistics. Springer, New York, second edition.

Chapitre 5

Les propriétés asymptotiques de l'estimateur à noyau des expectiles conditionnels pour les données fonctionnelles.

Le but de ce chapitre est d'estimer de manière non paramétrique la régression expectile dans le cas d'un prédicteur fonctionnel et d'une réponse scalaire. Plus précisément, nous construisons un estimateur à noyau de la fonction de régression expectile (RE). La principale contribution est l'obtention des propriétés asymptotiques de l'estimateur de Nadaraya-Watson de la RE, à savoir, la convergence presque complète avec spécification de la vitesse de convergence. De plus, nous obtenons la normalité asymptotique de l'estimateur proposé sous certaines conditions générales. Ainsi, nous expliquons comment appliquer nos résultats pour construire les intervalles de confiance. Une discussion riche et intéressante sur les impacts potentiels de la régression expectile fonctionnelle en NFDA est proposée avec une attention particulière sur les problèmes de classification, de prédiction et d'analyse des risques financiers. Enfin, les performances de la méthodologie sont illustrées à l'aide de l'analyse de données simulées et des données réelles d'analyse des risques financiers.

Le chapitre fait l'objet de deux articles publiés dans des revues internationales : «*Compte Rendus Mathématique. Académie des Sciences. Paris* » et «*Journal of Multivariate Analysis* ».

The Consistency and Asymptotic Normality of the Kernel type Expectile Regression Estimator for Functional Data.

Mustapha Mohammedi¹, Salim Bouzebda² and Ali Laksaci³

¹Université Djillali Liabès, BP 89, 22000, Sidi Bel Abbès, Algérie,
L.S.P.S., Sidi Bel Abbès, Algérie

²Alliance Sorbonne Université, Université de Technologie de Compiègne,
L.M.A.C., Compiègne, France

³Department of Mathematics, College of Science,
King Khalid University, Abha, 61413, Saudi Arabia

Abstract

The aim of this paper is to nonparametrically estimate the expectile regression in the case of a functional predictor and a scalar response. More precisely, we construct a kernel-type estimator of the expectile regression function. The main contribution of this study is the establishment of the asymptotic properties of the expectile regression estimator. Precisely, we establish the almost complete convergence with rate. Furthermore, we obtain the asymptotic normality of the proposed estimator under some mild conditions. We provide how to apply our results to construct the confidence intervals. The case of functional predictor is of particular interest and challenge, both from theoretical as well as practical point of view. We discuss the potential impacts of functional expectile regression in NFDA with a particular focus on the supervised classification, prediction and financial risk analysis problems. Finally, the finite-sample performances of the model and the estimation method are illustrated using the analysis of simulated data and real data coming from the financial risk analysis.

Key words : Almost consistency, Asymmetric least squares regression, Asymptotic normality, Expectiles, Functional data, Kernel type function estimator, Probability convergence, Nonparametric estimation, Risk measure, Strong mixing process.

AMS subject classification : 62G08 ; 62E20 ; 62F12 ; 62G10 ; 62G20 ; 62M10.

5.1 Introduction

Risk measures such as quantiles and expected shortfall have been studied and developed by several authors from a practical and a theoretical point of view, both risk measures have their own merits and defects. In this paper, we investigate an alternative way based on the least asymmetrically weighted squares estimation, borrowed from the econometrics literature, that is one of the basic tools in statistical applications. This method often involves [Newey and Powell \(1987\)](#) concept of expectiles, a least squares analogue of the

traditional quantiles. The proposed risk measure allows to overcome some drawbacks of the classical measures. It is expressed by

$$\boldsymbol{\theta}_p(Y) = \arg \min_{t \in \mathbb{R}} \mathbb{E} [\rho_p(Y, t)], \quad \text{for } 0 < p < 1, \quad (5.1)$$

such that

$$\begin{aligned} \rho_p(Y, t) &= |p - \mathbb{1}_{\{(Y-t) \leq 0\}}| (Y - t)^2 \\ &= p(Y - t)^2 \mathbb{1}_{\{(Y-t) > 0\}} + (1 - p)(Y - t)^2 \mathbb{1}_{\{(Y-t) \leq 0\}}, \end{aligned} \quad (5.2)$$

with $\rho_p(Y, t)$ is the scoring function, the level p is called the asymmetry parameter and $\mathbb{1}_A$ stands for the indicator function of the event A . Such risk model was introduced by [Aigner *et al.* \(1976\)](#) and [Newey and Powell \(1987\)](#), in the context of linear regression models. It is worth noticing that (5.2) generalizes the conditional *expectation* of Y given $X = x$, which coincides with $\boldsymbol{\theta}_p(Y)$ when specifically $p = 1/2$. On the other hand, (5.2) is similar to the conditional p -*quantile* of Y given $X = x$, which can be obtained by replacing $(Y - t)^2$ by $|Y - t|$ in (5.2). This motivates the name the conditional p -*expectile*. Hence the expectile regression is constructed by combining the ideas of least square regression to those of the conditional quantile allowing to accumulate the advantages of both models. In particular, its high sensitivity to the extreme value permits for more prudent and reactive risk management, and can be considered as an important advantage, since if we are measuring potential losses, we want that our measure to be sensitive to the extreme tail losses. The main aim of this work is to analyze the asymptotic behavior of this model using the functional nonparametric approach.

Recall that, for multivariate parametric context, the expectile model has received growing interest. We cite for instance [Bellini and Bernardino \(2017\)](#) for dapper discussion on the advantages of this model over the classical risk measures such as the quantiles and the expected shortfall. They indicate that the expectiles are perfectly reasonable alternatives to the value at risk (VaR) and the expected shortfall (ES) risk measures. The expectile based risk measure is coherent for $0.5 \leq p < 1$. Moreover, according to [Bellini *et al.* \(2014\)](#) and [Bellini and Bigozzi \(2015\)](#), the expectile is the only coherent as well as elicitable risk measure. For more details about the statistical significance of the property of elicibility and coherence, we refer the interested reader to the papers of [Gneiting \(2011\)](#), [Emmer *et al.* \(2015\)](#), [Artzner *et al.* \(1999\)](#) and [Ziegel \(2016\)](#). It turns out that, similar to the conditional mean or the percentile, the expectile regression of order p (or the conditional p -expectile) can be uniquely be defined by

$$\boldsymbol{\theta}(p; x) = \arg \min_{t \in \mathbb{R}} \mathbb{E} [\rho_p(Y, t) \mid X = x], \quad \text{for } 0 < p < 1. \quad (5.3)$$

Although they present differences in their construction, both quantiles and expectiles share similar properties. The main reason for this, as shown in [Jones \(1994\)](#), is the fact that expectiles are precisely quantiles but for a transformation of the original distribution. [Abdous and Rémillard \(1995\)](#) established an important feature is that quantiles and expectiles of the same distribution coincide under the hypothesis of weighted symmetry and pointed out that inference on expectiles is much easier than inference on quantiles. Notice that the quantiles are not always satisfactory and can be criticized for being somewhat difficult to compute as the corresponding loss function is not continuously differentiable. The key advantage of the expectile over the quantile is its efficiency and computing expedience, although it has not a direct interpretation as the quantile in terms of the relative frequency, see [Daouia *et al.* \(2019\)](#). Another substantial difference is that the conditional expectile is based on fundamentally different information and relies on more complete information to measure the risk in the sense that it depends on the form of the entire distribution, and it is easy to see that the conditional expectiles are characterized by the tail expectations in the same way that

the conditional quantiles are characterized by the conditional distribution function as shown by the following equation, for more details refer to [Newey and Powell \(1987\)](#) and [Abdous *et al.* \(2003\)](#) :

$$\frac{p}{1-p} = \left(\int_{(-\infty, \theta(p;x))} (y - \theta(p;x))^2 dF_Y^x(y) \right) \left(\int_{(\theta(p;x), +\infty)} (y - \theta(p;x))^2 dF_Y^x(y) \right)^{-1}, \quad (5.4)$$

where $F_Y^x(y)$ represents the conditional distribution function of Y given $\{X = x\}$ and $\theta(p;x)$ is the conditional expectile of order p .

The conditional expectile has been widely studied in the applied area such as econometrics, finance and actuarial science, see, for instance [Kuan *et al.* \(2009\)](#) as pioneer work, [Ehm *et al.* \(2016\)](#). Despite this importance, the expectile regression is unexplored compared to both competitive regressions (conditional mean and quantile). [Bellini *et al.* \(2018\)](#) have generalized the conditional expectation to the conditional expectile by means of the minimization of an asymmetric quadratic loss function and expose their main properties. [Koenker and Bassett \(1978\)](#) have defined the regression quantile estimators via an asymmetric absolute loss, the same authors in [Koenker and Bassett \(1982\)](#) have studied a new approach of regression quantiles based upon the estimates of the coefficients of the regression equation. Making use of the expectiles, [Daouia *et al.* \(2018\)](#) have proposed the estimations of the VaR and the ES or the CVaR. More details and results on regression are given by many references see [Efron \(1991\)](#), [Holzmann and Klar \(2016\)](#), [Maume-Deschamps *et al.* \(2017\)](#) and [Krättschmer and Zähle \(2017\)](#). While, these cited works consider the finite dimensional case, in the present work we use this model to analyze the effect of a functional covariate on a scalar response variable. Therefore, we treat in this paper the functional case. Note that analogously to the quantile regression, the conditional expectile allows to get an exhaustive information on the impact of explanatory variable on the response one by exploring its conditional distribution. On the other hand, the expectile regression can be employed for other standard regression model such as the curves discrimination or the prediction problems. Such questions of functional data analysis are particularly interesting, for many applied areas. Some basic materials in this modern branch of statistics can be found in the monographs of [Ramsay and Silverman \(2005\)](#), [Ferraty and Vieu \(2006\)](#), [Horváth and Kokoszka \(2012\)](#), [Kokoszka and Reimherr \(2017\)](#), [Hsing and Eubank \(2015\)](#) or in the special issues [Aneiros *et al.* \(2019a\)](#), [Aneiros *et al.* \(2019b\)](#) and [Goia and Vieu \(2016\)](#). It should be noted that the principal motivation of the functional statistics is the recent technological development of the measuring instruments allowing the data recording over thinner discretization grids. In this context, the economics or financial data, which are the principal applied areas of the expectile model, constitute a natural source of functional data. Accordingly, the inference statistic by the functional expectile is motivated by the diversity of its application areas including the risk analysis, the discriminant analysis and the forecasting problem.

Notice that the nonparametric estimation in the conditional models when the regressor takes its values in an infinite dimensional space, is an interesting subject and further details can be found in [Ferraty and Vieu \(2006\)](#), with more details. [Abdous *et al.* \(2003\)](#) study the nonparametric weighted symmetry tests, while the almost complete convergence of the conditional quantile kernel estimators is proved by [Ferraty and Vieu \(2006\)](#) in the independent and identically distributed observations (i.i.d.). The dependent case is considered by many authors, see for example [Ferraty *et al.* \(2005\)](#), [Laksaci *et al.* \(2009\)](#) and [Laksaci *et al.* \(2011\)](#). The papers of [Ezzahrioui and Ould-Saïd \(2008\)](#) and [Dabo-Niang and Laksaci \(2012\)](#) investigated the asymptotic normality of this estimator in both cases, i.i.d. and strong mixing variables. Moreover, the kernel method based on the locally linear fit is adapted to the nonparametric estimation of regression expectiles and percentiles by using the ALS approach, and under the assumptions of stationarity and ρ -mixing, the asymptotic normality for the estimators of conditional expectiles is thoroughly discussed in the paper of [Yao](#)

and Tong (1996).

In this contribution we establish the asymptotic properties of the kernel-type estimator of the expectile regression in the case of a functional predictor and a scalar response. To our best knowledge, this problem was open up to present, and it gives the main motivation to our paper. We focus on the almost complete consistency by giving its convergence rate, and we will establish the asymptotic normality of the estimator. The obtained results are stated under standard assumptions in nonparametric functional statistics. Nevertheless, a lot of attention has been paid to others nonparametric functional regression models such as the conditional expectation and the conditional quantiles. Among the wide literature concerning the nonparametric treatment of these functional models we only refer to Ferraty and Vieu (2006), Laksaci *et al.* (2009), Laksaci *et al.* (2011), Al-Awadhi *et al.* (2019) and the references therein. Finally, we refer the reader to the survey paper Ling and Vieu (2018) for a state of the art as well as some important future tracks on this topic of the nonparametric functional data analysis.

The layout of the article is as follows. In Section 5.2, we first present and describe the nonparametric model of the expectile regression. In Section 5.3, we establish our main results of the almost complete convergence of the kernel estimators and the asymptotic normality under some mild conditions in the i.i.d case. In Section 5.3.2, we extend our results to the dependent setting by showing the consistency in probability for the α -mixing sequences. In Section 5.3.3, we discuss an application of our results for the construction of the confidence interval. In Section 5.4, we discuss some connections of the conditional expectile in the supervised classification, prediction and financial risk analysis problems. The implementation of the expectile regression model in practice is given in Section 5.4.3 and Section 5.4.4, where a comparison with the quantile regression by illustration on real data is provided. Some concluding remarks are given in Section 5.5. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the Section 5.6. For a better understanding of our proofs, we recall some relevant results in the Appendix.

5.2 Model and estimators

5.2.1 The Conditional expectile

Let (X, Y) be a pair of random variable valued in $\mathcal{F} \times \mathbb{R}$, where (\mathcal{F}, d) is a semi-metric space equipped with a semi-metric $d(\cdot, \cdot)$ ¹, x is a fixed point in \mathcal{F} , defining a topology to measure the proximity between two elements of \mathcal{F} and which is disconnected of the definition of X in order to avoid measurability problems. For $0 < p < 1$, we define the conditional expectile $\theta(p; x)$ of order p of the increasing function $G(\cdot; x)$, defined below, by the unique solution with respect to (w.r.t.) t of the following optimization problem

$$\min_{t \in \mathbb{R}} \{ \mathbb{E} [p(Y - t)^2 \mathbf{1}_{\{(Y-t) > 0\}} \mid X = x] + \mathbb{E} [(1 - p)(Y - t)^2 \mathbf{1}_{\{(Y-t) \leq 0\}} \mid X = x] \}. \quad (5.5)$$

By a simple manipulation, we show that $\theta(p; x)$ is the unique solution w.r.t. t of

$$\zeta(p; x) = G(t; x) := \frac{G^1(t; x)}{G^2(t; x)}, \quad (5.6)$$

1. A semi-metric (sometimes called pseudo-metric) $d(\cdot, \cdot)$ is a metric which allows $d(x_1, x_2) = 0$ for some $x_1 \neq x_2$.

where

$$\begin{cases} \zeta(p; x) &= \frac{p}{1-p}, \\ G^1(t; x) &= -\mathbb{E}[(Y-t)\mathbf{1}_{\{(Y-t)\leq 0\}} | X=x], \\ G^2(t; x) &= \mathbb{E}[(Y-t)\mathbf{1}_{\{(Y-t)> 0\}} | X=x]. \end{cases}$$

We define the conditional expectile $\theta(p; x)$ of order p of the function $G(\cdot; x)$ as follows

$$\theta(p; x) = \inf\{t \in \mathbb{R} : G(t; x) \geq \zeta(p; x)\}. \quad (5.7)$$

Remark 1 For all $t \in \mathbb{R}$, the function $G(\cdot; x)$ is an increasing function.

Remark 2 In the real economical and financial problems, the conditional expectile measure can be considered and interpreted as the threshold that provides a ratio between profit and loss by the quantity $\zeta(p; x)$.

5.2.2 Nonparametric estimation of the conditional expectile

Let $\{(X_i, Y_i) : i \geq 1\}$ be a sequence of stationary random variables with the same distribution as the couple (X, Y) . Let $K(\cdot)$ be a kernel function and $h = h_n$ is a sequence of positive real numbers tending to zero as n tends to infinity and is called the smoothing parameter. The kernel estimator of the function $G(\cdot; x)$ is given by

$$\widehat{G}_{n, h_n}(t; x) = \frac{-\sum_{i=1}^n W_{ni}(x)(Y_i - t)\mathbf{1}_{\{(Y_i - t)\leq 0\}}}{\sum_{i=1}^n W_{ni}(x)(Y_i - t)\mathbf{1}_{\{(Y_i - t)> 0\}}}, \quad \text{for } t \in \mathbb{R}, \quad (5.8)$$

where

$$W_{ni}(x) = \frac{K_i(x)}{\sum_{i=1}^n K_i(x)} \quad \text{and} \quad K_i(x) = K(h^{-1}d(x, X_i)).$$

Using the functional kernel local weighted variables, we can write

$$\widehat{G}_{n, h_n}(t; x) = \frac{\widehat{G}_{n, h_n; -}(t; x)}{\widehat{G}_{n, h_n; +}(t; x)}, \quad (5.9)$$

where

$$\begin{cases} \widehat{G}_{n, h_n; -}(t; x) &= -\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)(Y_i - t)\mathbf{1}_{\{(Y_i - t)\leq 0\}}, \\ \widehat{G}_{n, h_n; +}(t; x) &= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)(Y_i - t)\mathbf{1}_{\{(Y_i - t)> 0\}}. \end{cases}$$

Consequently, we estimate the conditional expectile $\theta(p; x)$ by $\widehat{\theta}_{n, h_n}(p; x)$ that is

$$\widehat{\theta}_{n, h_n}(p; x) = \arg \min_{t \in \mathbb{R}} \widehat{\Psi}_p(t; x), \quad (5.10)$$

where $\widehat{\Psi}_p(t; x)$ is the estimator of $\Psi_p(t; x)$ defined by

$$\widehat{\Psi}_p(t; x) = \sum_{i=1}^n W_{ni}(x) \rho_p(Y_i, t), \text{ for } t \in \mathbb{R},$$

which yields to

$$\widehat{\theta}_{n, h_n}(p; x) = \inf\{t \in \mathbb{R} : \widehat{G}_{n, h_n}(t; x) \geq \zeta(p; x)\}. \quad (5.11)$$

5.2.3 Hypotheses and notation

In our analysis, the following conditions are needed to study the asymptotic properties of the estimator $\widehat{\theta}_{n, h_n}(p; x)$. We will denote by C_i , for $i = 1, 2, \dots$, some strictly positive constants.

- (A1) The function $G(\cdot; x)$ is differentiable in \mathbb{R} and satisfies the following Lipschitz's condition : $\exists a > 0$, $\forall t \in [\theta(p; x) - a, \theta(p; x) + a]$, $\forall x_1, x_2 \in \mathcal{F}$,

$$\left| G^i(t; x_1) - G^i(t; x_2) \right| \leq C_i d^{k_i}(x_1, x_2), \text{ for } k_i > 0 \text{ and } i = 1, 2.$$

Furthermore, we assume that $\forall x \in \mathcal{F}$, $G^{1'}(\theta(p; x); x)$ and $G^{2'}(\theta(p; x); x)$ are not identical to zero, with $G^{i'}(t; x)$ denoting the derivative of the function $G^i(t; x)$ with respect to t ;

- (A2) For each $m \geq 2$, $\varphi_m(Y^-) = \mathbb{E} \left[\left| Y^- \right|^m \mid X \right] \leq C_3 < \infty$, *a.s.*; with $Y^- = (Y - t) \mathbf{1}_{\{(Y-t) \leq 0\}}$;
 (A3) For all $x \in \mathcal{F}$,

$$\mathbb{P}(X \in B(x, h_n)) = \phi_x(h_n) > 0,$$

where $\phi_x(h_n)$ the small ball probability function, characterizes the classical concentration property (see [Ferraty and Vieu \(2006\)](#)) and can be written approximately as the product of two independent functions. $B(x, h_n)$ denotes the ball of the center x and radius $h_n > 0$ with

$$\lim_{n \rightarrow \infty} h_n = 0 \text{ and } \lim_{n \rightarrow \infty} \phi_x(h_n) = 0;$$

- (A4) Assume that

$$\lim_{n \rightarrow \infty} \frac{n \phi_x(h_n)}{\log n} = \infty;$$

- (A5) $K(\cdot)$ is a measurable function with a support $[0, 1]$ and there exist two constants, C_4 and C_5 such that $0 < C_4 < C_5 < \infty$, the kernel satisfies

$$0 < C_4 \leq K(\cdot) \leq C_5 < \infty;$$

- (B1) The function

$$R(t; x) = \mathbb{E}((Y - t)^2 \mid X = x),$$

is differentiable with respect to t in \mathbb{R} and satisfies the following Lipschitz's condition : $\exists c > 0$, $\forall t \in [\theta(p; x) - c, \theta(p; x) + c]$, $\forall x_1, x_2 \in \mathcal{F}$,

$$\left| R(t; x_1) - R(t; x_2) \right| \leq C_6 d^r(x_1, x_2), \text{ for } r > 0;$$

(B2) For all $\delta > 0$, and $C_7 > 0$,

$$\mathbb{E} \left[\left| Y - t \right|^{2+\delta} \mid X \right] \leq C_7 < \infty, \quad a.s.$$

(B3) The function $K(\cdot)$ satisfies (A5) and is differentiable on $(0, 1)$ with derivative $K'(\cdot)$ such that there exist two constants, C_8 and C_9 fulfilling $0 < C_8 < C_9 < \infty$, and

$$C_8 \mathbb{1}_{(0,1)}(\cdot) \leq K'(\cdot) \leq C_9 \mathbb{1}_{(0,1)}(\cdot).$$

Moreover, we assume that $n\phi_x(h_n) \rightarrow \infty$ and for $\eta > \min(k_1, k_2)$, $nh^{2\eta}\phi_x(h_n) \rightarrow 0$ as $n \rightarrow \infty$;

(B4) The concentration property given in (A3) holds and there exists a function $\beta_x(\cdot)$ such that

$$\forall s \in [0, 1], \quad \lim_{t \rightarrow 0} \frac{\phi_x(st)}{\phi_x(t)} = \beta_x(s).$$

Some comments

- Recall that the preceding conditions are not very restrictive. Our main result is stated under standard conditions of the almost complete consistency in the nonparametric functional statistics. In particular, the structure of the considered assumptions are the same as in the monograph of [Ferraty and Vieu \(2006\)](#). They cover the three structural axes of this study (data, model and estimator). Specifically, (A1) and (A2) allow to explore the dimensionality of the data and the model, respectively.
- The hypothesis (A1) is a regularity condition which characterizes the functional space of our model and is needed to evaluate the bias in the asymptotic results of this paper. The hypotheses (A2), (A4) and (A5) are technical conditions and are, also, similar to those considered in [Ferraty et al. \(2010\)](#).
- The bridge between local weighting and the notion of small ball probabilities is given by the hypothesis (A3), which is a simple normalization of the concentration property of the probability measure on the small balls.
- The hypotheses (B1), (B2) and (B3) are technical conditions on the functions $R(\cdot, \cdot)$ and $K(\cdot)$, and the hypothesis (B4) permits to give the variance term explicitly in the asymptotic normality, and is satisfied for a large family of random functional variables. Indeed, in many examples, the small ball probability function $\phi_x(h_n)$ can be written approximately as the product of two independent functions in terms of x and h , as in the following examples, which can be found in Proposition 1 of [Ferraty et al. \(2007\)](#) :
 1. If $\phi_x(h) \sim Ch^v$ for some $v > 0$ then $\beta_x(s) = s^v$;
 2. If $\phi_x(h) \sim C_1 h^v \exp(-C_2 h^{-p})$ for some $v > 0$ and $p > 0$ then $\beta_x(s)$ is the Dirac's function ;
 3. If $\phi_x(h) \sim C |\ln(h)|^{-1}$ then $\beta_x(s) = \mathbb{1}_{]0,1]}(s)$ the indicator function in $]0, 1]$.

5.3 The main results

In this section, we present our main results concerning the asymptotic properties of the estimator $\hat{\theta}_{n,h_n}(p; x)$ in both situations : independent and dependent cases.

5.3.1 Independent case

In this paragraph, we suppose that the $\{(X_i, Y_i) : i \geq 1\}$ are i.i.d. The first main result is given by the following proposition and theorem. We establish the almost complete convergence (a.co.)² of $\widehat{\boldsymbol{\theta}}_{n, h_n}(p; x)$ to $\boldsymbol{\theta}(p; x)$ in the i.i.d. case.

Proposition 5.1 *Under the hypotheses (A1)-(A5), we have, for all $t \in \mathbb{R}$, as $n \rightarrow \infty$,*

$$\left| \widehat{G}_{n, h_n}(t; x) - G(t; x) \right| = O\left(h_n^{k_1}\right) + O\left(h_n^{k_2}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \quad (5.12)$$

The proof of Proposition 5.1 is postponed until Section 5.6.

Theorem 5.1 *Under the hypotheses of Proposition 5.1, then, for all $\epsilon > 0$, we obtain*

$$\sum_{n \geq 1} \mathbb{P}\left(\left|\widehat{\boldsymbol{\theta}}_{n, h_n}(p; x) - \boldsymbol{\theta}(p; x)\right| > \epsilon\right) < \infty. \quad (5.13)$$

The proof of Theorem 5.1 is postponed until Section 5.6.

Now, we establish the asymptotic normality of $\widehat{\boldsymbol{\theta}}_{n, h_n}(p; x)$. Recall that the notation $Z \stackrel{\mathcal{D}}{=} N(\mu, \sigma^2)$ is used whenever the random variable Z follows a normal law with expectation μ and variance σ^2 , $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution and $\xrightarrow{\mathbb{P}}$ denotes the convergence in probability. Our second main result is presented in the following theorem.

Theorem 5.2 *Under the hypotheses (A1)-(A5)-(B1)-(B4), we have, as $n \rightarrow \infty$,*

$$Z_n = \left(\frac{n\phi_x(h_n)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\widehat{\boldsymbol{\theta}}_{n, h_n}(p; x) - \boldsymbol{\theta}(p; x)\right) \xrightarrow{\mathcal{D}} N(0, 1), \quad (5.14)$$

where

$$\sigma_p^2(x) = \frac{\alpha_2(x)\gamma_p(\boldsymbol{\theta}(p; x); x)}{\alpha_1^2(x)\Lambda_p^2(\boldsymbol{\theta}(p; x); x)}, \quad (5.15)$$

$$\alpha_j(x) = K^j(1) - \int_0^1 (K^j)'(s)\beta_x(s)ds \text{ for } j = 1, 2, \quad (5.16)$$

2. We say that the sequence $(\boldsymbol{\theta}_n)_n$ converges a.co. to zero, if and only if

$$\forall \tau > 0, \sum_{n \geq 1} \mathbb{P}(|\boldsymbol{\theta}_n| > \tau) < \infty.$$

Furthermore, we say that $\boldsymbol{\theta}_n = O_{a.co.}(\boldsymbol{\theta}_n)$, if there exists $\tau_0 > 0$, such that

$$\sum_{n \geq 1} \mathbb{P}(|\boldsymbol{\theta}_n| > \tau_0 \boldsymbol{\theta}_n) < \infty.$$

and

$$\gamma_p(\boldsymbol{\theta}(p; x); x) = \left(\frac{p}{1-p} \right)^2 R_+(\boldsymbol{\theta}(p; x); x) + R_-(\boldsymbol{\theta}(p; x); x),$$

where

$$\begin{aligned} R_+(\boldsymbol{\theta}(p; x); x) &= \mathbb{E} \left((Y_1 - \boldsymbol{\theta}(p; x))^2 \mathbb{1}_{\{Y_1 > \boldsymbol{\theta}(p; x)\}} \mid X = x \right), \\ R_-(\boldsymbol{\theta}(p; x); x) &= \mathbb{E} \left((Y_1 - \boldsymbol{\theta}(p; x))^2 \mathbb{1}_{\{Y_1 \leq \boldsymbol{\theta}(p; x)\}} \mid X = x \right), \end{aligned}$$

and

$$\Lambda_p(\boldsymbol{\theta}(p; x); x) = G^{1'}(\boldsymbol{\theta}(p; x); x) - \left(\frac{p}{1-p} \right) G^{2'}(\boldsymbol{\theta}(p; x); x).$$

The proof of Theorem 5.2 is postponed until Section 5.6.

Remark 3 • *It is clear that the convergence rate of the Proposition 5.1 is usual in nonparametric functional data analysis. In the sense that the nonparametric path of the model is explored in the bias term. While the functional nature of the model is explored in the variance term through the small ball probability function $\phi_x(h_n)$.*

- *Recall that the result of Theorem 5.1 can be viewed as generalisation of Theorem 6.1 in Ferraty and Vieu (2006) on the regression operator, by the fact that the expectile regression of order $p = 0.5$ reduces to the classical regression $\mathbb{E}[Y \mid X = x]$.*
- *Once again the asymptotic normality in Theorem 5.2 generalizes the asymptotic normality of the regression operator obtained in Theorem 2 of Ferraty et al. (2007). It should be noted that the asymptotic normality is of particular interest of some various asymptotic studies. In particular, it can be used as preliminary result to prove the mean quadratic consistency or to prove the uniform integrability of the estimator, one can refer for more details and discussion to Crambes et al. (2008).*

5.3.2 Dependent case

Often statistical observations are not independent but are not far from being independent. If not taken into account, dependence can have disastrous effects on statistical inference. The notion of mixing quantifies how close to independence a sequence of random variables is, and it can help us to extend classical results for independent sequences to weakly dependent or mixing sequences. Let us specify the dependence that we will consider in this section.

Definition 5.1 *We call the stationary $\{(X_t, Y_t)\}$ strong mixing if*

$$\alpha(j) = \sup_{A \in \widehat{F}_{-\infty}^0, B \in \widehat{F}_j^{-\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| = 0 \text{ as } j \rightarrow \infty,$$

where \widehat{F}_s^t denotes σ -algebra generated by $\{(X_t, Y_t), s \leq i \leq t\}$. We use the term *geometrically strong mixing* if : $\alpha(j) \leq aj^{-\beta}$ for some $a > 0$ and $\beta > 1$, and *exponentially strong mixing* if : $\alpha(k) \leq b\gamma^k$ for some $b > 0$ and $0 < \gamma < 1$.

The interested reader may refer to [Volkonskiĭ and Rozanov \(1959\)](#) and [Rosenblatt \(1956\)](#). [Ferraty and Vieu \(2006\)](#) provided some properties of this kind of functional process. To obtain the consistency in probability of $\widehat{\theta}_{n,h_n}(p;x)$, in addition to the assumptions **(A1)**-**(A3)**, **(A5)**, we need to the following conditions in our analysis.

(C1) The sequence $\{(X_i, Y_i) : i \geq 1\}$ satisfies : $\exists a > 2, \exists c > 0 : \forall n \in \mathbb{N}, \alpha(n) \leq cn^{-a}$ and

$$\begin{cases} \mathbb{E}[Y_i Y_j | X_i, X_j] \leq C < \infty, & \forall i \neq j, \\ \mathbb{P}((X_i, X_j) \in B(x, r) \times B(x, r)) = O(\phi_x^{1+\epsilon_1}(r)) > 0, & \text{for some } \epsilon_1 \in (2a^{-1}, 1]. \end{cases}$$

(C2) Assume that

$$\exists \epsilon_2 \in ((a\epsilon_1 - 1)^{-1}, 1), \quad \phi_x(h_n) = O(n^{-\epsilon_2}).$$

The main result of this section to be proved here may now be stated precisely as follows.

Theorem 5.3 *Under the hypotheses **(A1)**-**(A3)**, **(A5)** and **(C1)**-**(C2)**, we have, for all $\epsilon > 0$, as $n \rightarrow \infty$,*

$$\mathbb{P}\left(\left|\widehat{\theta}_{n,h_n}(p;x) - \theta(p;x)\right| > \epsilon\right) \rightarrow 0.$$

The proof of this results based on the same ideas of [Theorem 5.1](#). The proof of [Theorem 5.3](#) is postponed until [Section 5.6](#).

Remark 4 *Let's point out that [Theorem 5.3](#) states the asymptotic property of the estimator $\widehat{\theta}_{n,h_n}(p;x)$ in functional time series case using the strong mixing condition. The same result can be obtained using the geometric strong mixing condition for which the coefficient $\alpha(n)$ fulfils the following specific condition*

$$\exists C > 0, \exists t \in (0, 1), \quad \alpha(n) \leq Ct^n.$$

5.3.3 Application to the confidence intervals

The purpose of a confidence interval is to supplement the functional estimate at a point with information about the uncertainty in this estimate. It is a direct application of the central limit theorem (CLT). In order to provide a confidence interval for the conditional expectile measure, we need first to propose a consistent estimator of the variance given in [equation \(5.15\)](#). A natural consistent estimator of $\sigma_p^2(x)$ is then

$$\widehat{\sigma}_p^2(x) = \frac{\widehat{\alpha}_2(x) \widehat{\gamma}_p(\widehat{\theta}_{n,h_n}(p;x); x)}{\widehat{\alpha}_1^2(x) \widehat{\Lambda}_p^2(\widehat{\theta}_{n,h_n}(p;x); x)}, \quad (5.17)$$

where $(\widehat{\alpha}_i)_{i=1,2}$ are the empirical estimators of the coefficients $(\alpha_i)_{i=1,2}$ given in [equation \(5.15\)](#). Similarly to [Ferraty et al. \(2007\)](#), $(\widehat{\alpha}_i)_{i=1,2}$ are expressed by

$$\widehat{\alpha}_1(x) = \frac{1}{n\phi_x(h_n)} \sum_{i=1}^n K_i(x) \quad \text{and} \quad \widehat{\alpha}_2(x) = \frac{1}{n\phi_x(h_n)} \sum_{i=1}^n K_i^2(x).$$

We have now to give estimates of $\gamma_p(\cdot, \cdot)$ and $\Lambda_p(\cdot, \cdot)$ given in (5.17). Since $G^1(\cdot; x)$, $G^2(\cdot; x)$ and $\theta(p; x)$ are estimated by $\widehat{G}_{n, h_n; -}(\cdot; x)$, $\widehat{G}_{n, h_n; +}(\cdot; x)$ and $\widehat{\theta}_{n, h_n}(p; x)$ respectively, then $\Lambda_p(\cdot, \cdot)$ can be consistently estimated by

$$\widehat{\Lambda}_p\left(\widehat{\theta}_{n, h_n}(p; x); x\right) = \widehat{G}'_{n, h_n; -}\left(\widehat{\theta}_{n, h_n}(p; x); x\right) - \left(\frac{p}{1-p}\right) \widehat{G}'_{n, h_n; +}\left(\widehat{\theta}_{n, h_n}(p; x); x\right). \quad (5.18)$$

For $\gamma_p(\cdot, \cdot)$, we use the Nadaraya-Watson type estimator to estimate $R_-(\cdot; x)$ and $R_+(\cdot; x)$ by $\widetilde{R}_-(\cdot; x)$ and $\widetilde{R}_+(\cdot; x)$ respectively, then we obtain a consistent estimator of $\gamma_p(\cdot, \cdot)$:

$$\widehat{\gamma}_p\left(\widehat{\theta}_{n, h_n}(p; x); x\right) = \left(\frac{p}{1-p}\right)^2 \widetilde{R}_+\left(\widehat{\theta}_{n, h_n}(p; x); x\right) + \widetilde{R}_-\left(\widehat{\theta}_{n, h_n}(p; x); x\right), \quad (5.19)$$

with the estimators of $R_-(\cdot; x)$ and $R_+(\cdot; x)$ are given, respectively, by

$$\widetilde{R}_-\left(\widehat{\theta}_{n, h_n}(p; x); x\right) = \frac{\sum_{i=1}^n \left(Y_i - \widehat{\theta}_{n, h_n}(p; x)\right)^2 K_i(x) \mathbf{1}_{\{Y_i - \widehat{\theta}_{n, h_n}(p; x) \leq 0\}}}{\sum_{i=1}^n K_i(x)}, \quad (5.20)$$

$$\widetilde{R}_+\left(\widehat{\theta}_{n, h_n}(p; x); x\right) = \frac{\sum_{i=1}^n \left(Y_i - \widehat{\theta}_{n, h_n}(p; x)\right)^2 K_i(x) \mathbf{1}_{\{Y_i - \widehat{\theta}_{n, h_n}(p; x) > 0\}}}{\sum_{i=1}^n K_i(x)}. \quad (5.21)$$

Consequently, the equations (5.18), (5.19), (5.20) and (5.21) allow to construct a consistent estimator of $\sigma_p^2(x)$. On the other hand, an application of Theorem 5.2 gives

$$\mathbb{P}\left[-U_{1-\frac{\alpha}{2}} \leq \left(\frac{n\phi_x(h_n)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\widehat{\theta}_{n, h_n}(p; x) - \theta(p; x)\right) \leq U_{1-\frac{\alpha}{2}}\right] = 1 - \alpha, \quad (5.22)$$

where $(1 - \alpha)$ is the confidence level and $U_{1-\frac{\alpha}{2}}$ the quantile of order $(1 - \frac{\alpha}{2})$ of the standard normal law. Thereby, with a confidence level $(1 - \alpha)$, the confidence interval is provided by

$$\widehat{\theta}_{n, h_n}(p; x) \pm U_{1-\frac{\alpha}{2}} \left(\frac{\sigma_p^2(x)}{n\phi_x(h_n)}\right)^{\frac{1}{2}}.$$

Of course in practice, we replace $\sigma_p^2(x)$ by its estimator $\widehat{\sigma}_p^2(x)$, defined by

$$\widehat{\sigma}_p^2(x) = n\phi_x(h_n) \frac{\sum_{i=1}^n K_i^2(x) \widehat{\gamma}_p\left(\widehat{\theta}_{n, h_n}(p; x); x\right)}{\left(\sum_{i=1}^n K_i(x)\right)^2 \widehat{\Lambda}_p^2\left(\widehat{\theta}_{n, h_n}(p; x); x\right)},$$

and we obtain the following confidence interval

$$\widehat{\theta}_{n, h_n}(p; x) \pm U_{1-\frac{\alpha}{2}} \left(\frac{\sum_{i=1}^n K_i^2(x) \widehat{\gamma}_p\left(\widehat{\theta}_{n, h_n}(p; x); x\right)}{\left(\sum_{i=1}^n K_i(x)\right)^2 \widehat{\Lambda}_p^2\left(\widehat{\theta}_{n, h_n}(p; x); x\right)}\right)^{\frac{1}{2}}. \quad (5.23)$$

5.4 On the potential impact of the functional expectile regression in NFDA

It is well documented that the expectile model (conditional or unconditional) has many theoretical and practical merits. From the theoretical point of view, the expectile regression can be interpreted as the *quantilization* of the conditional expectation. In the sense that the expectiles play the same role as the quantiles by using the tail expectations rather than the tail probabilities. From the empirical point of view, the use of the expectation criterion is more informative than the probability distribution (characterized by the frequency of the data), because the expectation is based on both (values of the data and its frequencies). Similarly to the conditional median the conditional expectation coincides with the expectile regression of order $p = 0.5$. In this context, the hypotheses and asymptotic results are identifiable to the results of [Ferraty and Vieu \(2006\)](#). On the other hand, the expectile regression functions are used to construct alternative estimators for both known risk measures such as Conditional Value at Risk (CVaR) or the Conditional Expected Shortfall (CES), refer to [Taylor \(2008\)](#). Thus we can say that the present contribution is of great importance in nonparametric functional statistics. It can be used as an alternative model for several statistical issues such as the discriminant analysis, the prediction problems or the risk analysis.

5.4.1 Supervised classification and prediction : expectile regression versus classical regression

The supervised curves classification is one of the fundamental problems in the functional data analysis. Recall that the classical regression is a basic tool to solve this issue. As suggested by [Ferraty and Vieu \(2006\)](#), the functional version of Nadarya-Watson estimator provides an efficient rule of the supervised classification problem. Precisely, from a n -sample $(X_i, Y_i)_i$ where $Y_i = \mathbf{1}_{C_i}$, C_i being a given class, we estimate $\mathbb{P}(C | X) = \mathbb{E}[Y | X]$ to decide on the affectation of X to the class C . Specifically, the new curve \mathcal{X} is affected to class C if $\mathbb{P}(C | X = \mathcal{X}) > \mathbb{P}(\bar{C} | X = \mathcal{X})$, where \bar{C} is the complementary set of C . Based on the same reasoning, the expectile regression provides an alternative classifier. Indeed, for a binary random variable $Y = \mathbf{1}_C$, the expectile regression $\theta(p; x)$ of order p is explicitly defined by

$$\theta(p; x) = \frac{p\mathbb{P}(Y = 1 | X = x)}{p\mathbb{P}(Y = 1 | X = x) + (1 - p)\mathbb{P}(Y = 0 | X = x)}. \quad (5.24)$$

Therefore, our new classification rule affects the new curve \mathcal{X} to the class C when

$$\widehat{\theta}_{n, h_n}(p; \mathcal{X}) \geq p.$$

Such procedure constitutes a generalization of the regression approach given by [Ferraty and Vieu \(2006\)](#). In particular, the mean regression can be obtained by taking $p = 0.5$. In addition, the main feature of our new classification procedure is the possibility to take into account the cost of the misclassification. More precisely, assume the cost of the misclassification of C is ϱ_1 against the misclassification of \bar{C} is costed by ϱ_2 , the expectile classifier associated to $p = \frac{\varrho_1}{\varrho_1 + \varrho_2}$ is more appropriate than the classical regression. On the other hand, as a generalisation of the regression operator, the expectile regression function plays also an important role in the prediction issue. Recall that the nonparametric functional prediction problem consists to determine a real characteristic Y for a given curve \mathcal{X} . For this purpose, we consider the functionals random variables $(X_i(t))_{t \in \mathbb{R}}$ $i = 1, \dots, n$. For each curve $(X_i(t))_{t \in \mathbb{R}}$, we know its required real characteristic Y_i . Thus, the random variable $\widehat{\theta}_{n, h_n}(p; x_{new})$, defined by (5.24), is the best approximation of Y , given x_{new} with respect to the loss function $\rho_p(\cdot, \cdot)$ given in (5.2).

5.4.2 Financial risk analysis : conditional expectile versus conditional quantile

The second attractive domain of the expectile function is the financial risk management. Such issue is crucial for bankers as well as the investors. At this stage the CVaR or the CES, refer to [Taylor \(2008\)](#), are the most common models for financial risk analysis. However, both risk measures tools have some drawbacks. For instance, the CVaR is not coherent. On the other hand, the CES measure is not elicitable models. In this context, the expectile regression is a good alternative which allows to overcome the limitations of the last cited risk measures. In particular, recall that the expectile regression is the unique risk model that is coherent and elicitable. Recall that the coherence regroup four properties such as invariant by translation, homogeneity, monotonicity and subadditivity. Each property is instrumental for the investment processing. For instance, the VaR model fails to satisfy the subadditivity. This property guaranties that the risk of a financial portfolio with two combined positions is smaller than the sum of the total risk associated with either position. Thus the VaR model cannot identify the risk measure of diversified financial portfolio. On the other hand, the elicitable property is also primordial of risk analysis. In particular, the elicibility means that the risk model is obtained with respect to some appropriate criterion. Therefore, the elicitable property allows to validate the estimation procedure of the risk model by using the backtesting procedure. So, the non elicibility of the expected shortfall exhibits its practical limitations. Because there is non credible rule to check the validity of this risk measure. Thus we can say there is an important gain by using the expectile regression compared to the conditional quantile. Moreover, the standards models CVaR or CES can be alternatively computed by using the expectile regression. In particular, the right-tail $CES_x(\alpha)$, is defined by

$$\begin{aligned} CES_x(\alpha) &= \Phi(\alpha, \theta(p; x)) \\ &= \left(\frac{1-p}{(1-2p)(1-\alpha)} \right) \theta(.5; x) - \left(\left(\frac{\alpha(1-2p)+p}{(1-2p)(1-\alpha)} \right) \theta(p; x) \right), \end{aligned} \quad (5.25)$$

where $p \in (0, 1)$ such that

$$\theta(p; x) = \text{VaR}_\alpha(x), \quad (5.26)$$

where $\text{VaR}_\alpha(x)$ is the conditional quantile of Y given $X = x$ of order $\alpha \in (0, 1)$. Noting that the one-to-one correspondence between the expectiles and the quantiles of a related distribution has been proved by [Jones \(1994\)](#). For a recent comparison between quantile and expectile regressions and references, see [Waltrup et al. \(2015\)](#).

Finally let's point out that the expectile regression can be improved as a risk measure by introducing its own expected shortfall. Similarly to the CES, it is defined as the conditional expectation of the financial position given that it exceeds the expectile. The statistical estimation of this model and the identification of its real impact in risk management are fundamental future tracks of the present study. In the same insight, the originality of this study is also motivated by the numerous open questions which can be studied in the future.

5.4.3 Simulation study

The main aim of this section is to show the easy implementation of the constructed estimator in practice. Of course, the practical use of this estimator is closely related to the flexibility of the selection method

of the different parameters involved in the estimator. As all kernel smoothing in nonparametric functional data analysis, the influence of kernel $K(\cdot)$ is less important compared to the others parameters and the choice of the metric $d(\cdot, \cdot)$ depends on the shape of the curves $(X_i)_i$ and its regularity assumption (order of drivability, continuity, discontinuity, ...). Thus the challenge in this kind of estimator is the choice of the bandwidth parameter h_n . Unfortunately, despite the pivotal role of this parameter in practice, its choice stays insufficiently explored from theoretical point of view. It is well known that only two popular rules are studied in functional statistics such as the mean quadratic cross-validation rule by [Rachdi and Vieu \(2007\)](#) and the Bayesian approach by [Shang \(2014\)](#). However, from the practical point of view, [Ferraty and Vieu \(2006\)](#) use three cross-validation rules based on the minimization of the mean square prediction error. They evaluate three predictor models that are the conditional expectation, the conditional median and the conditional mode. In this simulation paragraph, we explore the main feature of the expectile regression, that is its multi-connection with classical regression and conditional quantile, to examine the feasibility of the selector procedures, proposed by [Ferraty and Vieu \(2006\)](#), on the expectile regression estimator. More precisely, we conduct a simulation study to examine the efficiency of $\hat{\theta}_{n, h_n}(p; x)$ under the two following cross-validation rules

$$\left\{ \begin{array}{l} \arg \min_{h \in H_n} \sum_{i=1}^n (Y_i - \hat{\theta}_h(0.5; X_i))^2 \quad \text{RegCV rule,} \\ \arg \min_{h \in H_n} \sum_{i=1}^n (Y_i - \hat{q}_h(0.5; X_i))^2 \quad \text{MedCV rule,} \end{array} \right.$$

where $\hat{q}_h(0.5; \cdot)$ is the conditional median estimator and H_n given subset of positive real numbers. For this aim, we generate a functional input-variables using the following formula

$$X(t) = at^3 + bt^2 + ct + e, \quad \text{with } a, c \sim N(0, 0.2) \quad \text{and } b, e \sim N(0, 1).$$

To take an idea on the suitable metric $d(\cdot, \cdot)$ on this functional variables, we show the shape of the curves in the following [Figure 5.1](#).

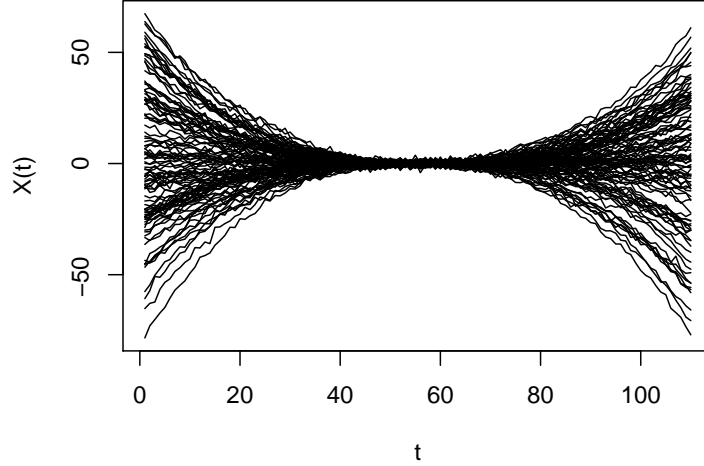


FIGURE 5.1 – The curves $X_{i=1,\dots,100}(t)$, $t_{j=1,\dots,110} \in [-3, 3]$

In view of the discontinuity of the curves, we can say that the PCA-metric (see [Ferraty and Vieu \(2006\)](#)) is adequate for this functional explanatory variables. Concerning the output variables Y , we generate $(Y_i)_i$ by the following regression function

$$Y = \sqrt{\int_{-3}^3 \log(X^2(t) + 1)} + \epsilon, \text{ where } \epsilon \text{ is white noise independent of } X.$$

Now, in order to check the behavior of $\hat{\theta}_{n,h_n}(p; x)$ in both cases (independent and dependent case), we simulate with two white noises. In the first situation, we suppose that $\epsilon \sim_{i.i.d.} N(0, 0.5)$, and in the second case, we generate it from the ARCH(1) model with parameters (0.01.0.9). For the computational task, we use the quadratic kernel $K(\cdot)$ on $(0, 1)$ and the PCA-metric associated to the third eigenvalue and we optimize the rules *MedCV* and (*RegCV*) using the local k -nearest neighborhood approach for which

$$H_n = \left\{ a \geq 0 \text{ such that } \sum_{i=1}^n \mathbb{1}_{B(x,a)}(X_i) = k \right\},$$

where $k \in \{5, 15, 25, \dots, 0.5n\}$. Finally, the efficiency of the expectile regression is evaluated by using the following backtesting criterion

$$Ex(p) = \frac{1}{n} \sum_{i=1}^n \rho_p(Y_i, \theta(p; X_i)),$$

where we recall that $\rho_p(\cdot, \cdot)$ is given in (5.2). The comparison results are reported in the Table 5.1 where we summarize the values of $Ex(p)$ for various sample size $n = 50, 150, 250$ and for three values of $p = 0.1, 0.05, 0.01$.

Cross-validation rule	White-noise type	n	$p = 0.1$	$p = 0.05$	$p = 0.01$
MedCV	IID	50	1.58	1.37	1.26
		150	1.22	1.14	1.02
		250	1.10	1.05	0.93
MedCV	ARCH	50	1.72	1.64	1.51
		150	1.56	1.33	1.12
		250	1.41	1.24	1.08
RegCV	IID	50	1.15	1.06	0.89
		150	0.82	0.74	0.67
		250	0.70	0.61	0.54
RegCV	ARCH	50	1.66	1.43	1.27
		150	1.30	1.16	1.03
		250	1.02	0.96	0.88

TABLE 5.1 – The $Ex(\cdot)$ values for different scenarios.

The obtained results confirm the compatibility of the rules *MedCV* and *RegCV* as bandwidth selector procedures on the expectile regression estimation with slight advantage to the *RegCV*-rule. Moreover, even if, without surprising, the simulation results are affected by the correlation of the data the $Ex(p)$ - values shows that $\hat{\theta}_{n,h_n}(p; x)$ has good performance in the dependence restriction. Finally, we can see that, the computational results incorporate the asymptotic theory about the consistency of the estimator $\hat{\theta}_{n,h_n}(p; x)$. In sense that the efficiency of $\hat{\theta}_{n,h_n}(p; x)$ increases with respect to the sample size n and it has satisfactory performance for moderate sample size when $n = 50$.

5.4.4 Real data example

In addition to the above discussions on the real impact of the expectile regression compared to its competitive models (quantile and regression), we will examine the easy implementation of this model in practice. In particular, we focus on the risk analysis area and we compare between the quantile and the expectile regressions. Although both models have the same role in the risk management, the expectile regression is more easier to compute and to interpret. Indeed, firstly the smoothing property of the score function associated with the expectile regression allows to increase the precision of its estimation and reduces its computational time cost. Secondly, in the financial risk area, the expectile function can be viewed as a threshold that measures a ratio between profit and loss which makes its interpretation as risk measure more easier than the quantile regression. So, in order to quantify this gain in practice, we conduct an empirical analysis based on the daily returns of $r(t)$ of 20 international financial stock indexes. The data is available in the website "<https://fred.stlouisfed.org/categories/32255/downloaddata>". It provides data during the period 01/ January/ 2010 to 31/December/2018. From this data, we consider 250 separate months independently chosen from different indexes. Of course, the analyzed functional data keeps all principal features of financial data that are volatility, excess kurtosis, skewness and non-normality (see, the Figure 5.2).

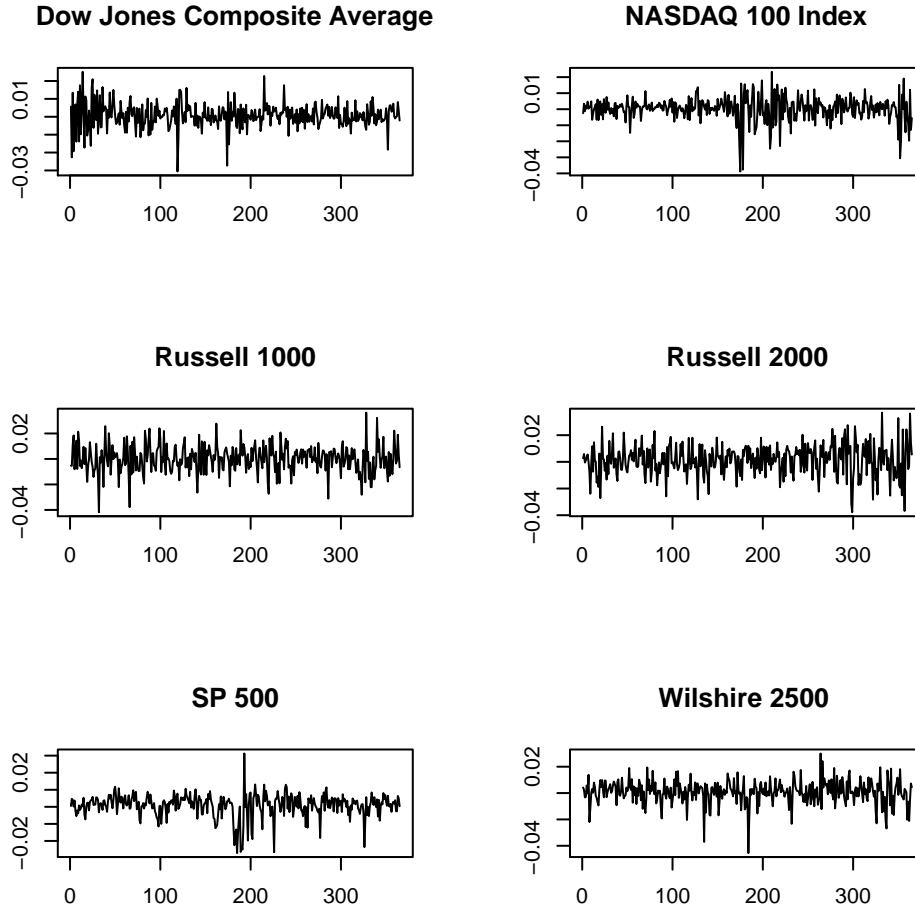


FIGURE 5.2 – The daily values of $Z(t) = -100 \log \left(\frac{r(t)}{r(t-1)} \right)$ for some stock index

Formally, in this real data study, we suppose that the curve is $X(\cdot)$ is the monthly curve of the time series $Z(t) = -100 \log \left(\frac{r(t)}{r(t-1)} \right)$ and the real response variable is $Y = Z(t_l + 1)$, where t_l is the last day of the month. As all statistical modeling, the efficiency of the two approaches is closely linked to the choice of the different parameters involved in the definition of the estimators. However, to conduct a fair comparison between both risk measures, we use the same rule on the choice of the fundamental parameters in the estimators. Precisely, concerning the conditional quantile we use the routine R-code "*funopare.quantile.lcv*" provided by [Ferraty and Vieu \(2006\)](#) which is available in the website "<http://www.lsp.ups-tlse.fr/staph/npfda>". More precisely, we point out that we have used this routine with a quadratic kernel on $(0, 1)$ and the PCA-metric. While the optimal bandwidth was chosen by splitting the data into subsets (Learning and testing samples $(200 + 50)$) and using the 200-observations of the learning to select the best smoothing parameter by the *RegCV*-rule. For sake of brevity, we have just considered the most popular method, that is, the cross-validated selected bandwidth. This may be extended to any other bandwidth selector such the bandwidth based on Bayesian ideas [Shang \(2014\)](#). As previously mentioned, the expectile regression estimator is computed by using the

same arguments. Finally, the efficiency of the two approaches is evaluated by using the following backtesting criterions

$$E(\boldsymbol{\theta}) = \frac{1}{50} \sum_{i=201}^{250} \rho_p(Y_i - \boldsymbol{\theta}(p; X_i)),$$

$$Q(\boldsymbol{\theta}) = \frac{1}{50} \sum_{i=201}^{250} L_p(Y_i - \boldsymbol{\theta}(p; X_i)),$$

where

$$L_p(t) = (2p - 1)(Y - t) + |Y - t|,$$

and $\boldsymbol{\theta}(p; X_i)$ means quantile or expectile regression. Such errors are evaluated for a fixed value of $p = 0.05$ and for both estimators conditional expectile or conditional quantile. The proposed criterions are performed over 100 replications, of each case, we interchange the observations between the learning and the testing parts. The obtained errors are displayed in the following box-plot figures.

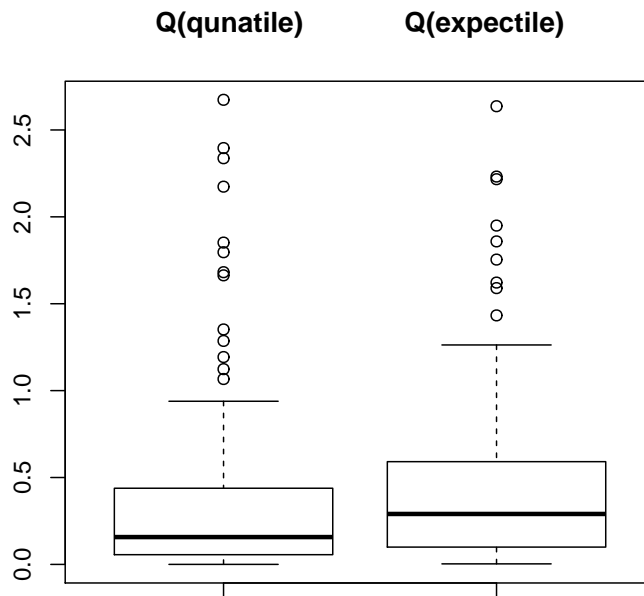


FIGURE 5.3 – Comparison between the quantile and the expectile using the Q -backtesting measure

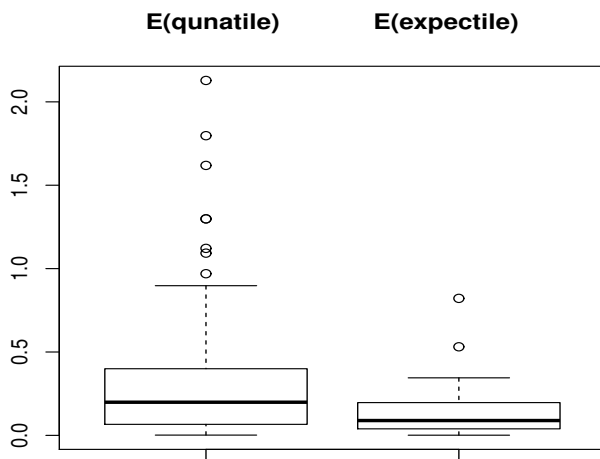


FIGURE 5.4 – Comparison between the quantile regression and the expectile regression using the E -backtesting measure

It appears clearly that the expectile regression has slight superiority in terms of precision compared to the quantile regression. Specifically, they are equivalent for the $Q(\cdot)$, in Figure 5.3, backtesting measure whereas the $E(\cdot)$, in Figure 5.4, measure gives a significative superiority of the expectile regression. Undoubtedly, this gain is due to the ease to determine an accurate estimator of the expectile regression compared to the last absolute error regression. However, we mention that both methods fit with satisfactory performance the volatility of the process.

5.5 Conclusion and perspectives

A natural perspective of this work is an extension to the multidimensional framework where $\mathbf{Y} \in \mathbb{R}^d$. On the basis of the multidimensional expectile in the paper of [Maume-Deschamps et al. \(2017\)](#), let us introduce the conditional multidimensional expectile : Let $\| \cdot \|$ be a norm on \mathbb{R}^d . We denote by $(\mathbf{Y}_1)_+$ the vector $(\mathbf{Y}_1)_+ = ((Y_1)_+, \dots, (Y_d)_+)^T$ and by $(\mathbf{Y})_-$ the vector $(\mathbf{Y})_- = (\mathbf{Y}_1)_-, \dots, (Y_d)_-)^T$. We define the following scoring function

$$s_\alpha(\mathbf{Y}, \mathbf{y}) = \alpha \| (\mathbf{Y} - \mathbf{y})_+ \|^2 + (1 - \alpha) \| (\mathbf{y} - \mathbf{Y})_+ \|^2,$$

for all $\mathbf{y} \in \mathbb{R}^d$. We call a *multivariate conditional expectile* any minimizer

$$\mathbf{y}^* \in \arg \min_{\mathbf{y} \in \mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{Y}, \mathbf{y}) \mid X = x].$$

It will be of interest to consider this extension in a future investigation. In addition to this issue of multivariate response, another challenging task would be to study the problem of the bandwidth selection in the expectile regression. Such question is primordial for a better use of this estimator in practice. Recall that, in our computational study the optimal smoothing parameter h_n was chosen by some cross-validation methods over a discrete set of bandwidth values defined by the k -Nearest Neighbors (k -NN). Even if these selector procedures give satisfactory results by employing the k -NN-approach, they lack a mathematical support to highlight the conditions of their optimality and to ensure their asymptotic outcomes. At this stage, the uniform consistency in number of neighborhoods, recently developed in functional statistics by [Kara et al. \(2017\)](#) or [Ling et al. \(2019\)](#) may be a best way to overcome this issue. Moreover, it should be noted that one of the advantages of the expectile regression is its possibility to behave with various statistical strategies (parametric, semiparametric or nonparametric modeling) in the same insight as in the regression and the conditional quantile. Thus, we can say that the linear or semiparametric version of the functional expectile is also a natural prospect of the present contribution. A particular attention will be paid to the semiparametric approach which is actually the trend in functional data analysis (see [Aneiros and Vieu \(2015\)](#) and [Novo et al. \(2019\)](#) for more discussions and motivations in this topic). Among the numerous open questions that can be concluded from the present contribution, we mention the construction of pointwise confidence intervals by combining the bootstrap ideas used by [Raña et al. \(2016\)](#) on the conditional expectation to those used by [Hahn \(1995\)](#) on the quantile regression. Finally, let's point out that it would be interesting to extend of our results to other functional time series cases (ergodic, long memory, associated process) which requires nontrivial mathematics, that goes well beyond the scope of the present paper.

5.6 The demonstration of asymptotic results

This section is devoted to the proofs of our results. The aforementioned notation is also used in what follows.

Proof of Theorem 5.1

For $t \in \mathbb{R}$ and by the fact that \widehat{G}_{n,h_n} is an increasing function, for any $\epsilon > 0$ small enough, we have

$$\begin{aligned} & \left\{ \left| \widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) - \boldsymbol{\theta}(p; x) \right| > \epsilon \right\} \\ & \Rightarrow \left\{ \widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) < \boldsymbol{\theta}(p; x) - \epsilon \right\} \cup \left\{ \widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) > \boldsymbol{\theta}(p; x) + \epsilon \right\} \\ & \Rightarrow \left\{ \widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x) \leq \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p; x) - \epsilon; x) \right\} \\ & \quad \cup \left\{ \widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x) \geq \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p; x) + \epsilon; x) \right\}. \end{aligned}$$

Since $G(\boldsymbol{\theta}(p; x); x) = \widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x)$, we obtain

$$\Rightarrow \begin{cases} G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) & \leq \widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \\ G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) + \epsilon; x) & \geq \widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p; x) + \epsilon; x). \end{cases}$$

Hence, we infer that

$$\begin{aligned} & \left\{ \left| \widehat{\boldsymbol{\theta}}_{n,h_n} - \boldsymbol{\theta} \right| > \epsilon \right\} \\ & \Rightarrow \left\{ \widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \geq G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \right\} \\ & \quad \cup \left\{ \widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p; x) + \epsilon; x) \leq G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) + \epsilon; x) \right\}. \end{aligned}$$

This implies that we have

$$\begin{aligned} & \left\{ \left| \widehat{\boldsymbol{\theta}}_{n,h_n} - \boldsymbol{\theta} \right| > \epsilon \right\} \\ & \Rightarrow \left\{ \left| \widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \right| \geq \left| G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \right| \right\} \\ & \quad \cup \left\{ \left| \widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p; x) + \epsilon; x) \right| \geq \left| G(\boldsymbol{\theta}(p; x) + \epsilon; x) - G(\boldsymbol{\theta}(p; x); x) \right| \right\}. \end{aligned}$$

From the last equation, we obtain that

$$\begin{aligned} & \sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) - \boldsymbol{\theta}(p; x) \right| > \epsilon \right) \\ & \leq \sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p; x) - \epsilon; x) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \right| \geq \left| G(\boldsymbol{\theta}(p; x); x) - G(\boldsymbol{\theta}(p; x) - \epsilon; x) \right| \right) \\ & \quad + \sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p; x) + \epsilon; x) - G(\boldsymbol{\theta}(p; x) + \epsilon; x) \right| \geq \left| G(\boldsymbol{\theta}(p; x) + \epsilon; x) - G(\boldsymbol{\theta}(p; x); x) \right| \right). \end{aligned}$$

Hence, by using the Proposition 5.1, for all $t \in \mathbb{R}$ and $\tau > 0$ we have,

$$\sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n}(t; x) - G(t; x) \right| > \tau \right) < \infty. \quad (5.27)$$

Then, the proof of Theorem 5.1 is a direct consequence of the Proposition 5.1. □

Proof of Proposition 5.1

We recall the definitions given in (5.6) and (5.9). We first decompose $\widehat{G}_{n,h_n}(t; x) - G(t; x)$ into the sum of two components, as follows :

$$\begin{aligned} \widehat{G}_{n,h_n}(t; x) - G(t; x) &= \frac{\widehat{G}_{n,h_n,-}(t)}{\widehat{G}_{n,h_n,+}(t)} - \frac{G^1(t; x)}{G^2(t; x)} \\ &= \frac{1}{\widehat{G}_{n,h_n,+}(t)} \left[\widehat{G}_{n,h_n,-}(t) - G^1(t; x) \right] \\ &\quad + \frac{G(t; x)}{\widehat{G}_{n,h_n,+}(t)} \left[G^2(t; x) - \widehat{G}_{n,h_n,+}(t) \right]. \end{aligned} \quad (5.28)$$

Lemmas 5.1–5.5 below are given in order to prove Proposition 5.1. To obtain the consistency results, one usually writes the difference between the estimator and the quantity to be estimated as the sum of a stochastic term and the so-called bias. The mathematical program is to determine the size of the stochastic term and to control the bias separately in the following Lemmas.

Lemme 5.1 *Assume that the hypotheses (A1)(for $i = 1$), (A2), (A3), (A4) and (A5) are satisfied. We have, as $n \rightarrow \infty$,*

$$\left| \widehat{G}_{n,h_n,-}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] \right| = O_{a.co.} \left(\sqrt{\frac{\log n}{n\phi_x(h_n)}} \right). \quad (5.29)$$

Lemme 5.2 *Assume that the hypotheses (A1)(for $i = 1$), (A3) and (A5) are fulfilled. We have as $n \rightarrow \infty$,*

$$\left| \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] - G^1(t; x) \right| = O \left(h_n^{k_1} \right). \quad (5.30)$$

Lemme 5.3 *Assume that the hypotheses (A1)(for $i = 1$), (A2), (A3), (A4) and (A5) are satisfied. We have, as $n \rightarrow \infty$,*

$$\left| \widehat{G}_{n,h_n,+}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n,+}(t; x) \right] \right| = O_{a.co.} \left(\sqrt{\frac{\log n}{n\phi_x(h_n)}} \right). \quad (5.31)$$

Lemme 5.4 *Assume that the hypotheses (A1)(for $i = 2$), (A3) and (A5) are satisfied. We have, as $n \rightarrow \infty$,*

$$\left| \mathbb{E} \left[\widehat{G}_{n,h_n,+}(t; x) \right] - G^2(t; x) \right| = O \left(h_n^{k_2} \right). \quad (5.32)$$

Lemme 5.5 *Assume that the hypotheses (A1)(for $i = 1$), (A2), (A3), (A4) and (A5) are satisfied. We have, as $n \rightarrow \infty$,*

$$\sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n,+}(t; x) \right| \leq \epsilon \right) < \infty. \quad (5.33)$$

In order to demonstrate the consistency of $\widehat{\theta}_{n,h_n}(p; x)$, it suffices to show that the estimator $\widehat{G}_{n,h_n,+}(t; x)$ converge almost completely, as $n \rightarrow \infty$.

Proof of Lemma 5.1

We will use the notation $Y_i^- = (Y_i - t)\mathbb{1}_{\{(Y_i - t) \leq 0\}}$. We have

$$\begin{aligned} & \left| \widehat{G}_{n,h_n;-}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n;-}(t; x) \right] \right| \\ &= \left| -\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)Y_i^- - \mathbb{E} \left[-\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x)Y_i^- \right] \right| \\ &= \frac{1}{n} \left\{ \left| \frac{1}{\mathbb{E}[K_1(x)]} \sum_{i=1}^n (K_i(x)Y_i^- - \mathbb{E}[K_i(x)Y_i^-]) \right| \right\} \\ &= \frac{1}{n} \left(\left| \sum_{i=1}^n \Delta_i \right| \right), \end{aligned}$$

where we have used the notation

$$\Delta_i = \frac{1}{\mathbb{E}[K_1(x)]} (K_i(x)Y_i^- - \mathbb{E}[K_i(x)Y_i^-]).$$

In order to control the stochastic term, we will make use of the Bernstein's exponential inequality for which the main ingredient is to evaluate asymptotically the m -th order moment of Δ_i :

$$\begin{aligned} & \mathbb{E} \left(\left| \frac{1}{\mathbb{E}[K_1(x)]} K_1 Y_1^- \right|^m \right) \\ &= \frac{1}{|\mathbb{E}[K_1(x)]|^m} \mathbb{E} \left[|K_1|^m |Y_1^-|^m \right] \\ &= \frac{1}{|\mathbb{E}[K_1(x)]|^m} \mathbb{E} \left[\mathbb{E} \left[|Y_1^-|^m \mid X = x \right] K_1^m \right] \\ &= \frac{1}{|\mathbb{E}[K_1(x)]|^m} \mathbb{E} \left[\varphi_m(Y_1^-) K_1^m \right] = \mathbb{E} \left[\varphi_m(Y_1^-) (\widehat{K}_1)^m \right], \end{aligned}$$

where

$$\widehat{K}_i(\cdot) = \frac{K_i(\cdot)}{\mathbb{E}[K_i(\cdot)]} \text{ and recall that } \varphi_m(Y) = \mathbb{E} \left[|Y|^m \mid X = x \right].$$

By using the assumptions **(A2)**, **(A3)** and **(A5)**, we get

$$\mathbb{E} \left| \left(\widehat{K}_1 Y_1^- \right)^m \right| = \mathbb{E} \left| \left(\frac{1}{\mathbb{E}[K_1(x)]} K_1 Y_1^- \right)^m \right| \leq C_3 (\phi_x(h_n))^{-m+1}.$$

This readily implies that

$$\mathbb{E} \left| \left(\widehat{K}_1 Y_1^- \right)^m \right| = O \left((\phi_x(h_n))^{-m+1} \right). \tag{5.34}$$

By the use of the Newton's binomial formula, we obtain that

$$\begin{aligned}
 & \mathbb{E} \left| \left(Y_1^- \widehat{K}_1 - \mathbb{E} \left[Y_1^- \widehat{K}_1 \right] \right) \right|^m \\
 &= \mathbb{E} \left| \sum_{l=0}^m C_l^m \left(Y_1^- \widehat{K}_1 \right)^l \left(\mathbb{E} \left[Y_1^- \widehat{K}_1 \right] \right)^{m-l} (-1)^{m-l} \right| \\
 &\leq \sum_{l=0}^m C_l^m \mathbb{E} \left| Y_1^- \widehat{K}_1 \right|^l \left| \mathbb{E} \left[Y_1^- \widehat{K}_1 \right] \right|^{m-l} \\
 &\leq \sum_{l=0}^m C_l^m \mathbb{E} \left| Y_1^- \widehat{K}_1 \right|^l \left| \mathbb{E} \left[\mathbb{E} \left[Y_1^- \widehat{K}_1 \right] \mid X_1 \right] \right|^{m-l} \\
 &\leq C \sum_{l=0}^m C_l^m \left(\phi_x(h_n) \right)^{-l+1} \left| G_1^{X_1}(t) \right|^{m-l} \\
 &\leq C \max_{0 \leq l \leq m} \left(\phi_x(h_n) \right)^{-l+1} = C \left(\phi_x(h_n) \right)^{-m+1}.
 \end{aligned}$$

We therefore obtain that

$$\mathbb{E} \left| \left(Y_1^- \widehat{K}_1 - \mathbb{E} \left(Y_1^- \widehat{K}_1 \right) \right) \right|^m = O \left(\left(\phi_x(h_n) \right)^{-m+1} \right). \quad (5.35)$$

Thus, to achieve this proof, it suffices to use the classical Bernstein's inequality (see Corollary A8 in [Ferraty and Vieu \(2006\)](#)) with

$$a_n^2 = \left(\phi_x(h_n) \right)^{-1},$$

and by using [\(A4\)](#) we obtain, for all $\tau > 0$, that

$$\begin{aligned}
 \mathbb{P} \left(\left| \widehat{G}_{n,h_n,-}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] \right| > \tau \sqrt{\frac{\log n}{n \phi_x(h_n)}} \right) &= \mathbb{P} \left(\left| \sum_{i=1}^n \Delta_i \right| > n \tau \sqrt{\frac{\log n}{n \phi_x(h_n)}} \right) \\
 &\leq 2 \exp \left(-\frac{1}{2} \vartheta_n \right),
 \end{aligned}$$

where

$$\vartheta_n = \frac{\tau^2 \log n}{1 + \tau \sqrt{\frac{\log n}{n \phi_x(h_n)}}},$$

and

$$\sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n,-}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] \right| > \tau \sqrt{\frac{\log n}{n \phi_x(h_n)}} \right) \leq 2 \sum_{n \geq 1} \exp \left(-C \tau^2 \log n \right).$$

Thereby, for a strictly positive constant C , we choose $\tau > \frac{1}{\sqrt{C}}$ to obtain that

$$\sum_{n \geq 1} \exp \left(-C \tau^2 \log n \right) < \infty.$$

This, implies that

$$\left| \widehat{G}_{n,h_n,-}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] \right| = O_{a.co.} \left(\sqrt{\frac{\log n}{n \phi_x(h_n)}} \right). \quad (5.36)$$

Hence the proof is complete. □

Proof of Lemma 5.2

Notice that we have

$$\begin{aligned}
 & \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] - G^1(t; x) \\
 &= \mathbb{E} \left[-\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x) Y_i^- \right] - (-\mathbb{E}[Y^- | X = x]) \\
 &= \mathbb{E}[Y^- | X = x] - \frac{1}{\mathbb{E}[K_1(x)]} \mathbb{E}[K_1(x) Y_1^-] \\
 &= \frac{1}{\mathbb{E}[K_1(x)]} \{ \mathbb{E}[K_1 \mathbb{E}[Y^- | X = x]] - \mathbb{E}[\mathbb{E}[K_1 Y_1^- | X_1]] \} \\
 &= \frac{1}{\mathbb{E}[K_1(x)]} \{ \mathbb{E}[K_1(x) \mathbb{E}[Y^- | X = x]] - \mathbb{E}[Y_1^- | X_1] \} \\
 &= \frac{1}{\mathbb{E}[K_1(x)]} \left\{ \mathbb{E} \left[K_1(x) \left(G_1^{X_1}(t) - G^1(t; x) \right) \right] \right\},
 \end{aligned}$$

where

$$Y_i^- = (Y_i - t) \mathbb{1}_{\{(Y_i - t) \leq 0\}}.$$

We have

$$\begin{aligned}
 & \mathbb{E} \left[K_1(x) \left(G^1(t; X_1) - G^1(t; x) \right) \right] \\
 &= \mathbb{E} \left(K \left(\frac{d(x, X_1)}{h_n} \right) \left(G^1(t; X_1) - G^1(t; x) \right) \right) \\
 &= \int K \left(\frac{d(x, z)}{h_n} \right) \left(G^1(t; z) - G^1(t; x) \right) d\mathbb{P}^{X_1} \\
 &= \int_{d(x,z) \leq h_n} K \left(\frac{d(x, z)}{h_n} \right) \left(G^1(t; z) - G^1(t; x) \right) d\mathbb{P}^{X_1} \\
 &= \int_{B(x, h_n)} K \left(\frac{d(x, z)}{h_n} \right) \left(G^1(t; z) - G^1(t; x) \right) d\mathbb{P}^{X_1} \\
 &= \mathbb{E} \left[K_1(x) \left(G^1(t; X_1) - G^1(t; x) \right) \mathbb{1}_{B(x, h_n)} \right].
 \end{aligned}$$

By using the condition **(A1)** (for $i = 1$), we obtain that

$$\begin{aligned}
 & \left| \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] - G^1(t; x) \right| \\
 &= \frac{1}{\mathbb{E}[K_1(x)]} \left| \mathbb{E} \left[K_1(x) \left(G^1(t; X_1) - G^1(t; x) \right) \mathbb{1}_{B(x, h_n)} \right] \right| \\
 &\leq C_1 \mathbb{1}_{B(x, h_n)} d^{k_1}(x, X_1) \\
 &\leq C_1 h_n^{k_1}.
 \end{aligned}$$

This readily implies that

$$\left| \mathbb{E} \left[\widehat{G}_{n,h_n,-}(t; x) \right] - G^1(t; x) \right| = O \left(h_n^{k_1} \right). \quad (5.37)$$

By combining the results (5.37) and (5.36), we establish the almost complete convergence of $\widehat{G}_{n,h_n,-}(t; x)$ to $G^1(t; x)$. Thus, for all $\epsilon > 0$, we have

$$\sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n;-}(t; x) - G^1(t; x) \right| > \epsilon \right) < \infty. \quad (5.38)$$

Finally, by Lemma 5.2 and 5.1, we get

$$\left| \widehat{G}_{n,h_n;-}(t; x) - G^1(t; x) \right| = O \left(h_n^{k_1} \right) + O_{a.co.} \left(\sqrt{\frac{\log n}{n \phi_x(h_n)}} \right). \quad (5.39)$$

Hence the proof is complete. \square

Proof of Lemma 5.3

By using the same steps as those used for proving Lemma 5.1, replacing $Y_i^- = (Y_i - t) \mathbb{1}_{\{(Y_i - t) \leq 0\}}$ with $Y_i^+ = (Y_i - t) \mathbb{1}_{\{(Y_i - t) > 0\}}$ and by applying the Bernstein's inequality we derive that

$$\left| \widehat{G}_{n,h_n;+}(t; x) - \mathbb{E} \left[\widehat{G}_{n,h_n;+}(t; x) \right] \right| = O_{a.co.} \left(\sqrt{\frac{\log n}{n \phi_x(h_n)}} \right). \quad (5.40)$$

Hence the proof is complete. \square

Proof of Lemma 5.4

We observe that we have

$$\begin{aligned} & \mathbb{E} \left[\widehat{G}_{n,h_n;+}(t; x) \right] - G^2(t; x) \\ &= \mathbb{E} \left[\frac{1}{n \mathbb{E} [K_1(x)]} \sum_{i=1}^n K_i(x) Y_i^+ \right] - \mathbb{E} [Y^+ | X = x] \\ &= \frac{1}{\mathbb{E} [K_1(x)]} \left\{ \mathbb{E} \left[K_1(x) \left(G_2^{X_1}(t) - G^2(t; x) \right) \right] \right\}, \end{aligned}$$

where

$$Y_i^+ = (Y_i - t) \mathbb{1}_{\{(Y_i - t) > 0\}}.$$

Then, under assumption (A1)(ii), we get :

$$\left| \mathbb{E} \left[\widehat{G}_{n,h_n;+}(t; x) \right] - G^2(t; x) \right| = O \left(h_n^{k_2} \right). \quad (5.41)$$

By combining the results (5.40) and (5.41), we obtain the almost complete convergence of $\widehat{G}_{n,h_n;+}(t; x)$ to $G^2(t; x)$. Thus for all $\epsilon > 0$, we have

$$\sum_{n \geq 1} \mathbb{P} \left(\left| \widehat{G}_{n,h_n;+}(t; x) - G^2(t; x) \right| > \epsilon \right) < \infty. \quad (5.42)$$

Finally, the application of Lemmas 5.4 and 5.3 yields to

$$\left| \widehat{G}_{n,h_n,+}(t;x) - G^2(t;x) \right| = O\left(h_n^{k_2}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \quad (5.43)$$

Hence the proof is complete. □

Proof of Lemma 5.5

By taking into account the fact that $G^2(t;x) > 0$, for all $t \in \mathbb{R}$, we readily infer that

$$\left| \widehat{G}_{n,h_n,+}(t;x) \right| \leq \frac{1}{2}G^2(t;x) \implies \left| \widehat{G}_{n,h_n,+}(t;x) - G^2(t;x) \right| > \frac{1}{2}G^2(t;x).$$

This, implies that

$$\mathbb{P}\left(\left|\widehat{G}_{n,h_n,+}(t;x)\right| \leq \frac{1}{2}G^2(t;x)\right) \leq \mathbb{P}\left(\left|\widehat{G}_{n,h_n,+}(t;x) - G^2(t;x)\right| > \frac{1}{2}G^2(t;x)\right).$$

An application of the Lemma 5.3, with the particular choice that $\delta = \frac{1}{2}G^2(t;x)$, gives

$$\sum_{n \geq 1} \mathbb{P}\left(\left|\widehat{G}_{n,h_n,+}(t;x)\right| \leq \frac{1}{2}G^2(t;x)\right) < \infty. \quad (5.44)$$

Hence the proof is complete. □

Proof of Theorem 5.2

Let us introduce

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\widehat{\theta}_{n,h_n}(p;x) - \theta(p;x)\right).$$

For $z \in \mathbb{R}$, we set

$$\delta_p(z,x) = \theta(p;x) + z(n\phi_x(h))^{-\frac{1}{2}}\sigma_p(x).$$

We have the following decomposition

$$\begin{aligned} \mathbb{P}(Z_n \leq z) &= \mathbb{P}\left(\widehat{\theta}_{n,h_n}(p;x) \leq \delta_p(z,x)\right) \\ &= \mathbb{P}\left(\left\{\widehat{\theta}_{n,h_n}(p;x) \leq \delta_p(z,x)\right\} \cap \left\{\widehat{G}_{n,h_n,+}(\delta_p(z,x)) = 0\right\}\right) \\ &\quad + \mathbb{P}\left(\left\{\widehat{\theta}_{n,h_n}(p;x) \leq \delta_p(z,x)\right\} \cap \left\{\widehat{G}_{n,h_n,+}(\delta_p(z,x)) \neq 0\right\}\right) \\ &= I_1(x) + I_2(x). \end{aligned} \quad (5.45)$$

The term $I_2(x)$ can be written as follows

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) \leq \delta_p(z, x) &\implies \frac{p}{1-p} \leq \widehat{G}_{n,h_n}(\delta_p(z, x); x) \\ &\implies \frac{p}{1-p} \leq \frac{\widehat{G}_{n,h_n;-}(\delta_p(z, x); x)}{\widehat{G}_{n,h_n;+}(\delta_p(z, x); x)}.\end{aligned}$$

Using the fact that the function $\widehat{G}_{n,h_n}(\cdot; x)$ is an increasing function and $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$ is the unique solution of the equation $\widehat{G}_{n,h_n}(\boldsymbol{\theta}; x) = \frac{p}{1-p}$, we readily obtain

$$\begin{aligned}I_2(x) &\leq \mathbb{P}\left(\frac{p}{1-p} \leq \frac{\widehat{G}_{n,h_n;-}(\delta_p(z, x); x)}{\widehat{G}_{n,h_n;+}(\delta_p(z, x); x)}\right) \\ &\leq \mathbb{P}\left(\left(\frac{p}{1-p}\right) \widehat{G}_{n,h_n;+}(\delta_p(z, x); x) - \widehat{G}_{n,h_n;-}(\delta_p(z, x); x) \leq 0\right).\end{aligned}$$

We have

$$I_2(x) \leq \mathbb{P}\left(\widehat{L}_p(z, x) - \mathbb{E}\left[\widehat{L}_p(z, x)\right] \leq \mathbb{E}\left[-\widehat{L}_p(z, x)\right]\right), \quad (5.46)$$

where

$$\widehat{L}_p(z, x) = \left(\frac{p}{1-p}\right) \widehat{G}_{n,h_n;+}(\delta_p(z, x)) - \widehat{G}_{n,h_n;-}(\delta_p(z, x)).$$

In order to complete the proof of Theorem 5.2, we need first to prove the following Lemmas. Lemma 5.6 shows that $I_1(x)$ converges to zero, and the Lemmas 5.7, 5.8 show the convergence in distribution of $I_2(x)$ to a standard normal random variable, as n tends to infinity.

Lemme 5.6 *Assume that the hypotheses (A3), (A5) and (B3) are satisfied. We have, as $n \rightarrow \infty$,*

$$I_1(x) = o(1), \quad (5.47)$$

where $I_1(x)$ is defined in the equation (5.45).

Lemme 5.7 *Assume that the hypotheses of Theorem 5.2 are satisfied. We have, as $n \rightarrow \infty$,*

$$\left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x))^{-1} \mathbb{E}\left[-\widehat{L}_p(z, x)\right] = z + o(1). \quad (5.48)$$

Lemme 5.8 *Suppose that the hypotheses of Theorem 5.2 hold true. We have, as $n \rightarrow \infty$,*

$$(n\phi_x(h)) \text{Var}\left(\sum_{i=1}^n \mu_i(x)\right) \longrightarrow \frac{\alpha_2(x)\gamma_p(\boldsymbol{\theta}(p; x); x)}{\alpha_1^2(x)}, \quad (5.49)$$

and

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x))^{-1} \left(\widehat{L}_p(z, x) - \mathbb{E}\left[\widehat{L}_p(z, x)\right]\right) \xrightarrow{\mathcal{D}} N(0, 1), \quad (5.50)$$

where

$$\mu_i(x) = \frac{1}{n\mathbb{E}[K_1(x)]} K_i(x) (Y_i - \delta_p(z, x)) \left(\frac{p}{1-p} \mathbf{1}_{A_i} + \mathbf{1}_{A_i^c}\right),$$

with

$$A_i = \{Y_i > \delta_p(z, x)\} \quad \text{and} \quad A_i^c = \{Y_i \leq \delta_p(z, x)\}.$$

Proof of Lemma 5.6

Since the variables X_i are independent, we get, under **(A3)**, **(A5)** and **(B3)**, that, as $n \rightarrow \infty$

$$\begin{aligned} I_1(x) &= \leq \mathbb{P} \left[\left(\widehat{G}_{n,h_n,+}(\delta_p(z,x)) = 0 \right) \right] \\ &= \mathbb{P} \left(\bigcap_{i=1}^n \{ \omega \in \Omega \text{ such that } X_i(\omega) \notin B(x, h_n) \} \right) \\ &= (1 - \phi_x(h))^n \\ &\leq \exp(-n\phi_x(h)) \rightarrow 0. \end{aligned}$$

Thus the proof is complete. □

Proof of Lemma 5.7

Keep in mind the definition

$$\widehat{L}_p(z,x) = \left(\frac{p}{1-p} \right) \widehat{G}_{n,h_n,+}(\delta_p(z,x)) - \widehat{G}_{n,h_n,-}(\delta_p(z,x)),$$

and

$$G(\boldsymbol{\theta}(p;x);x) = \frac{G^1(\boldsymbol{\theta}(p;x);x)}{G^2(\boldsymbol{\theta}(p;x);x)}.$$

By the fact that $\boldsymbol{\theta}(p;x)$ is the conditional expectile of $G(\cdot;x)$ of order p , i.e.,

$$G(\boldsymbol{\theta}(p;x);x) = \frac{p}{1-p}.$$

We have

$$\begin{aligned} &\mathbb{E} \left[-\widehat{L}_p(z,x) \right] \\ &= \mathbb{E} \left[\widehat{G}_{n,h_n,-}(\delta_p(z,x);x) \right] - \frac{p}{1-p} \mathbb{E} \left[\widehat{G}_{n,h_n,+}(\delta_p(z,x);x) \right] \\ &= \left(\mathbb{E} \left[\widehat{G}_{n,h_n,-}(\delta_p(z,x);x) \right] - G^1(\delta_p(z,x);x) \right) \\ &\quad + \frac{p}{1-p} \left(G^2(\delta_p(z,x);x) - \mathbb{E} \left[\widehat{G}_{n,h_n,+}(\delta_p(z,x);x) \right] \right) \\ &\quad + \left(G^1(\delta_p(z,x);x) - G^1(\boldsymbol{\theta}(p;x);x) \right) \\ &\quad - \frac{p}{1-p} \left(G^2(\delta_p(z,x);x) - G^2(\boldsymbol{\theta}(p;x);x) \right) \\ &\quad + \left(G^1(\boldsymbol{\theta}(p;x)) - \frac{p}{1-p} G^2(\boldsymbol{\theta}(p;x);x) \right) \\ &= J_1(x) + J_2(x) + J_3(x) + J_4(x) + J_5(x). \end{aligned} \tag{5.51}$$

By the fact that $\boldsymbol{\theta}(p;x)$ is the unique solution of the equation (5.6), we infer that the term $J_5(x)$, in the last equation, is equal to zero. We next evaluate the term $J_1(x)$ in the right side of (5.51). Taking into

consideration the conditions **(A1)**(for $i = 1$), **(A5)** and **(B3)**, we have the following bound

$$\begin{aligned} |J_1(x)| &= \left| \mathbb{E} \left[\widehat{G}_{n,h_n;-}(\delta_p(z, x)) \right] - G^1(\delta_p(z, x); x) \right| \\ &\leq C_1 h_n^{k_1}. \end{aligned}$$

Following similar steps for $J_2(x)$ and using the hypotheses **(A1)**(for $i = 2$), **(A5)** and **(B3)**, we show that

$$\begin{aligned} |J_2(x)| &= \frac{p}{1-p} \left| \mathbb{E} \left[\widehat{G}_{n,h_n;+}(\delta_p(z, x)) \right] - G^2(\delta_p(z, x); x) \right| \\ &\leq C_2 h_n^{k_2}, \end{aligned}$$

which implies, for $i = 1, 2$,

$$|J_i(x)| = O(h_n^{k_i}).$$

In order to evaluate the terms $J_3(x)$ and $J_4(x)$ in the right side of (5.51), we start with the Taylor's expansion of the function $G^1(\cdot; x)$ to obtain

$$\begin{aligned} J_3(x) &= G^1(\delta_p(z, x); x) - G^1(\boldsymbol{\theta}(p; x); x) \\ &= z\sigma_p(x) (n\phi_x(h))^{-\frac{1}{2}} G^{1'}(\boldsymbol{\theta}(p; x); x) \\ &\quad + o\left((n\phi_x(h))^{-\frac{1}{2}}\right), \end{aligned}$$

and the similar arguments for $J_4(x)$ yield to

$$\begin{aligned} J_4(x) &= \left(-\frac{p}{1-p}\right) G^2(\delta_p(z, x); x) - G^2(\boldsymbol{\theta}(p; x); x) \\ &= z\sigma_p(x) (n\phi_x(h))^{-\frac{1}{2}} \left(\left(-\frac{p}{1-p}\right) G^{2'}(\boldsymbol{\theta}(p; x); x) \right) \\ &\quad + o\left((n\phi_x(h))^{-\frac{1}{2}}\right) \end{aligned}$$

Then, by setting

$$\Lambda_p(\boldsymbol{\theta}(p; x); x) = G^{1'}(\boldsymbol{\theta}(p; x); x) - \left(\frac{p}{1-p}\right) G^{2'}(\boldsymbol{\theta}(p; x); x),$$

one can see that

$$\left| \sum_{i=1}^5 J_i(x) \right| = O(h_n^\eta) + z\sigma_p(x) (n\phi_x(h))^{-\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x)) + o\left((n\phi_x(h))^{-\frac{1}{2}}\right).$$

Under the assumption **(B3)**, for $\eta > \min(k_1, k_2)$, we get the following result

$$\begin{aligned} \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x))^{-1} \left| \mathbb{E} \left(-\widehat{L}_p(z, x) \right) \right| &= O\left(nh_n^{2\eta}\phi_x(h)\right)^{\frac{1}{2}} + z + o(1) \\ &= z + o(1). \end{aligned}$$

Thus the proof is complete. □

Proof of Lemma 5.8

In what follows, we set

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)} \right)^{\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x))^{-1} \left(\widehat{L}_p(z, x) - \mathbb{E} \left[\widehat{L}_p(z, x) \right] \right).$$

We can apply the Lyapunov CLT to conclude that, as $n \rightarrow \infty$,

$$Z_n = \frac{1}{S_n} \sum_{i=1}^n (\mu_i(x) - \mathbb{E}[\mu_i(x)]) \xrightarrow{\mathcal{D}} N(0, 1), \quad (5.52)$$

where

$$S_n = \left(\sum_{i=1}^n \text{Var}(\mu_i(x)) \right)^{\frac{1}{2}}.$$

In order to demonstrate that equation (5.52) holds, we start by computing the asymptotic variance of $\widehat{L}_p(z, x)$. We keep in mind the following equations

$$\begin{aligned} \widehat{L}_p(z, x) &= \widehat{L}_p(z, x) = \left(\frac{p}{1-p} \right) \widehat{G}_{n, h_n; +}(\delta_p(z, x); x) - \widehat{G}_{n, h_n; -}(\delta_p(z, x); x), \\ \delta_p(z, x) &= \delta_p(z, x) = \boldsymbol{\theta}(p; x) + z(n\phi_x(h))^{-\frac{1}{2}} \sigma_p(x). \end{aligned}$$

We first write $\widehat{L}_p(z, x)$ in the following form

$$\begin{aligned} \widehat{L}_p(z, x) &= \frac{p}{1-p} \left[\frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x) (Y_i - \delta_p(z, x)) \mathbf{1}_{(Y_i - \delta_p(z, x)) > 0} \right] \\ &\quad - \frac{1}{n\mathbb{E}[K_1(x)]} \left[\sum_{i=1}^n K_i(x) (Y_i - \delta_p(z, x)) \mathbf{1}_{(Y_i - \delta_p(z, x)) \leq 0} \right] \\ &= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K_i(x) (Y_i - \delta_p(z, x)) \left[\frac{p}{1-p} \mathbf{1}_{A_i} + \mathbf{1}_{A_i^c} \right] \\ &= \sum_{i=1}^n \mu_i(x), \end{aligned}$$

where we recall that $\mu_i(x)$, A_i and A_i^c are defined in Lemma 5.8. Thus we can simplify the problem of obtaining the asymptotic variance of $\widehat{L}_p(z, x)$ by computing instated the variance of the sum of $\mu_i(x)$. We

can therefore write the following equalities

$$\begin{aligned}
 & (n\phi_x(h)) \text{Var} \left(\widehat{L}_p(z, x) \right) \\
 &= (n\phi_x(h)) \text{Var} \left(\sum_{i=1}^n \mu_i(x) \right) \\
 &= \frac{\phi_x(h)}{\mathbb{E}^2 [K_1(x)]} \text{Var} \left(K_1(x) (Y_1 - \delta_p(z, x)) \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right) \right) \\
 &= \phi_x(h) \frac{\mathbb{E} [K_1^2(x)]}{\mathbb{E}^2 [K_1(x)]} \mathbb{E} \left[\frac{K_1^2(x) (Y_1 - \delta_p(z, x))^2 \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right)^2}{\mathbb{E} [K_1^2(x)]} \right] \\
 &- \phi_x(h) \mathbb{E}^2 \left[\frac{K_1(x) (Y_1 - \delta_p(z, x)) \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right)}{\mathbb{E} [K_1(x)]} \right] \\
 &= T_0(x)T_1(x) + T_2(x). \tag{5.53}
 \end{aligned}$$

By conditioning with respect to X_1 , we readily obtain

$$\begin{aligned}
 & \mathbb{E} \left[K_1(x) (Y_1 - \delta_p(z, x)) \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right) \right] \\
 &= \mathbb{E} \left[K_1(x) \left\{ \left(\frac{p}{1-p} \right) G^2(\delta_p(z, x), X_1) - G^1(\delta_p(z, x), X_1) \right\} \right] \\
 &= \mathbb{E} \left[K_1(x) \left\{ \left(\frac{p}{1-p} \right) M_2(x) - M_1(x) \right\} \right],
 \end{aligned}$$

where for $i = 1, 2$

$$M_i(x) = G^i(\delta_p(z, x), X_1) - G^i(\boldsymbol{\theta}(p; x); x).$$

Under the assumption **(A1)**, it follows immediately from the triangle inequality that, for $i = 1, 2$,

$$\mathbb{1}_{d(x, X_1) \leq h} |M_i(x)| \leq C_i h^{k_i} + z (n\phi_x(h))^{-\frac{1}{2}} \sigma_p(x) G^{i'}(\boldsymbol{\theta}(p; x); x).$$

From which we deduce that

$$\mathbb{E} \left[\frac{K_1(x) (Y_1 - \delta_p(z, x)) \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right)}{\mathbb{E} [K_1(x)]} \right] = o(1).$$

This, implies that

$$T_2(x) = o(1). \tag{5.54}$$

We next evaluate the term $T_1(x)$ in the right side of (5.53). We set

$$\lambda_p(\delta_p; x) = \left(\frac{p}{1-p} \right)^2 R_+(\delta_p(z, x); x) + R_-(\delta_p(z, x); x),$$

and

$$\gamma_p(\boldsymbol{\theta}(p; x); x) = \left(\frac{p}{1-p} \right)^2 R_+(\boldsymbol{\theta}(p; x); x) + R_-(\boldsymbol{\theta}(p; x); x).$$

By the same previous arguments used to treat the $T_2(x)$, we have

$$\begin{aligned} T_1(x) &= \mathbb{E} \left[\frac{K_1^2(x) (Y_1 - \delta_p(z, x))^2 \left(\frac{p}{1-p} \mathbb{1}_{A_1} + \mathbb{1}_{A_1^c} \right)^2}{\mathbb{E} [K_1^2(x)]} \right] \\ &= \mathbb{E} \left[\frac{K_1^2(x) \lambda_p(\delta_p, X_1)}{\mathbb{E} [K_1^2(x)]} \right] \\ &= \mathbb{E} \left[\frac{K_1^2(x) \left\{ \left(\frac{p}{1-p} \right)^2 N_2(x) + N_1(x) \right\}}{\mathbb{E} [K_1^2(x)]} \right] + \gamma_p(\boldsymbol{\theta}(p; x); x), \end{aligned}$$

where

$$N_1(x) = R_-(\delta_p(z, x); X_1) - R_-(\boldsymbol{\theta}(p; x); x),$$

and

$$N_2(x) = R_+(\delta_p(z, x); X_1) - R_+(\boldsymbol{\theta}(p; x); x).$$

An application of the triangle inequality implies readily that

$$\begin{aligned} \mathbb{1}_{d(x, X_1) \leq h} |N_1(x)| &\leq C_1 h^{r_1} + z (n \phi_x(h))^{-\frac{1}{2}} \sigma_p(x) R'_-(\boldsymbol{\theta}(p; x); x), \\ \mathbb{1}_{d(x, X_1) \leq h} |N_2(x)| &\leq C_2 h^{r_2} + z (n \phi_x(h))^{-\frac{1}{2}} \sigma_p(x) R'_+(\boldsymbol{\theta}(p; x); x). \end{aligned}$$

By combining the assumptions **(A3)** and **(B1)**, **(B3)**, we infer that

$$T_1(x) = \gamma_p(\boldsymbol{\theta}(p; x); x) + o(1). \quad (5.55)$$

We evaluate the term $T_0(x)$ in the right side of (5.53). Using similar arguments as in Ferraty and Vieu (2006), we can show that

$$\mathbb{E} [K_1^j(x)] = K^j(1) \phi_x(h) - \int_0^1 ((K^j))'(s) \phi_x(sh) ds + o(\phi_x(h)) \text{ for } j = 1, 2. \quad (5.56)$$

This when combined with (5.15) gives that, as $n \rightarrow \infty$,

$$\frac{\phi_x(h) \mathbb{E} [K_1^2(x)]}{\mathbb{E}^2 [K_1(x)]} \rightarrow \frac{\alpha_2(x)}{\alpha_1^2(x)}. \quad (5.57)$$

By combining (5.54), (5.55) and (5.57) we obtain the statement (5.49). To derive the asymptotic normality, we will show the Lyapunov conditions for the independent variables $\mu_i(x)$. Since the observations are i.i.d. and making use of the C_r -inequality (see (Loève, 1963, p.155)), this implies, for some $\eta > 0$, that

$$\begin{aligned} &(n \phi_x(h))^{\frac{2+\eta}{2}} \sum_{i=1}^n \mathbb{E} \left[\left| \mu_i(x) - \mathbb{E} [\mu_i(x)] \right|^{2+\eta} \right] \\ &\leq 2^{\eta+1} n (n \phi_x(h))^{\frac{2+\eta}{2}} \mathbb{E} \left[\left| \mu_1(x) \right|^{2+\eta} \right] \\ &\quad + 2^{\eta+1} n (n \phi_x(h))^{\frac{2+\eta}{2}} \left| \mathbb{E} [\mu_1(x)] \right|^{2+\eta} \\ &= A_1(x) + A_2(x). \end{aligned}$$

Through conditioning with respect to X_1 and using the conditions **(B2)** and **(B3)**, we obtain

$$\begin{aligned}
 A_1(x) &= 2^{\eta+1} (n\phi_x(h))^{-\frac{\eta}{2}} \left(\frac{\phi_x(h)}{\mathbb{E}[K_1(x)]} \right)^{2+\eta} \\
 &\quad \times \mathbb{E} \left[\frac{K_1^{2+\eta}(x)}{\phi_x(h)} \left| Y_1 - \delta_p(z, x) \right|^{2+\eta} \left| \frac{p}{1-p} \mathbf{1}_A + \mathbf{1}_{A^c} \right|^{2+\eta} \right] \\
 &\leq \left(\max \left(1, \frac{p}{1-p} \right) \right)^{2+\eta} 2^{\eta+1} (n\phi_x(h))^{-\frac{\eta}{2}} \left(\frac{\phi_x(h)}{\mathbb{E}[K_1(x)]} \right)^{2+\eta} \\
 &\quad \times \mathbb{E} \left[\frac{K_1^{2+\eta}(x)}{\phi_x(h)} \mathbb{E} \left(\left| Y_1 - \delta_p(z, x) \right|^{2+\eta} \mid X_1 \right) \right] \\
 &\leq C (n\phi_x(h))^{-\frac{\eta}{2}}.
 \end{aligned}$$

This, implies that

$$A_1(x) \longrightarrow 0 \text{ as } n \longrightarrow \infty. \quad (5.58)$$

In order to evaluate $A_2(x)$, we use the same arguments as those involved in the proof of $T_2(x)$. We get

$$\begin{aligned}
 A_2(x) &= 2^{\eta+1} n (n\phi_x(h))^{-\frac{2+\eta}{2}} \left| \mathbb{E}[\mu_1(x)] \right|^{2+\eta} \\
 &= 2^{\eta+1} n^{-\frac{\eta}{2}} (\phi_x(h))^{-\frac{2+\eta}{2}} \left| \mathbb{E} \left[\frac{K_1(x)(Y_1 - \delta_p(z, x))(\frac{p}{1-p} \mathbf{1}_A + \mathbf{1}_{A^c})}{\mathbb{E}[K_1(x)]} \right] \right|^{2+\eta} \\
 &= 2^{\eta+1} n^{-\frac{\eta}{2}} (\phi_x(h))^{-\frac{2+\eta}{2}} \left| T_2(x) \right|^{2+\eta}.
 \end{aligned}$$

By combing the condition (5.54) with the assumption **(A3)**, we obtain

$$A_2(x) = o(1). \quad (5.59)$$

Consequently, from (5.58), (5.59) and (5.49), for some $\eta > 0$, we infer that, as $n \longrightarrow \infty$,

$$\frac{1}{(n\phi_x(h))^{-\frac{2+\eta}{2}} S_n^{2+\eta}} (n\phi_x(h))^{-\frac{2+\eta}{2}} \sum_{i=1}^n \mathbb{E} \left[\left| \mu_i(x) - \mathbb{E}[\mu_i(x)] \right|^{2+\eta} \right] \longrightarrow 0. \quad (5.60)$$

We conclude from (5.60) that the Lyapunov condition is verified which, implies the convergence in distribution of Z_n to a standard normal $N(0, 1)$. We therefore obtain, as $n \longrightarrow \infty$,

$$(n\phi_x(h)) \text{Var} \left(\widehat{L}_p(z, x) \right) \longrightarrow \sigma_p^2(x) \Lambda_p^2(\boldsymbol{\theta}(p; x); x). \quad (5.61)$$

Combining (5.49) and (5.61), we infer

$$\sigma_p^2(x) \Lambda_p^2(\boldsymbol{\theta}(p; x); x) = \frac{\alpha_2(x) \gamma_p(\boldsymbol{\theta}(p; x); x)}{\alpha_1^2(x)}.$$

Thus by (5.45) and (5.47) with (5.46), we get

$$\mathbb{P}(Z_n \leq z) \leq o(1) + \mathbb{P} \left(Z_n \leq U_n \cdot \mathbb{E} \left[-\widehat{L}_p(z, x) \right] \right), \quad (5.62)$$

where

$$U_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)} \right)^{\frac{1}{2}} (\Lambda_p(\boldsymbol{\theta}(p; x); x))^{-1}.$$

Finally, in order to prove Theorem 5.2, we combine (5.62) with the equations (5.48), (5.49) and (5.50). \square

Proof of Theorem 5.3

The proof is based on the same arguments as those used in the proof of Theorem 5.1. However the bias term is not affected by the dependence assumption. So, it remains to evaluate the asymptotic behaviors of the variance terms of $\widehat{G}_{n,h_n;-}(t; x)$ and $\widehat{G}_{n,h_n;+}(t; x)$ to prove the convergence in probability of the estimator $\widehat{\theta}_{n,h_n}(p; x)$ stated in Theorem 5.3.

Lemma 5.9 *Under the hypotheses of the Theorem 5.3, we have, as $n \rightarrow \infty$,*

$$\text{Var} \left[\widehat{G}_{n,h_n;-}(t; x) \right] = O \left(\frac{n}{\phi_x(h_n)} \right), \quad (5.63)$$

and

$$\text{Var} \left[\widehat{G}_{n,h_n;+}(t; x) \right] = O \left(\frac{n}{\phi_x(h_n)} \right). \quad (5.64)$$

Proof of Lemma 5.9

The proof of this lemma is omitted, it follows the same lines of the proof of Lemma 11.5 in [Ferraty and Vieu \(2006\)](#). □

5.7 Appendix

5.7.1 Demonstration of Remark 1 : the function $G(\cdot; x)$ is increasing.

Recall that

$$G(t; x) = \frac{G^1(t; x)}{G^2(t; x)}, \quad (5.65)$$

where

$$\begin{aligned} G^1(t; x) &= -\mathbb{E} \left[(Y - t) \mathbf{1}_{\{(Y-t) \leq 0\}} \mid X = x \right], \\ G^2(t; x) &= \mathbb{E} \left[(Y - t) \mathbf{1}_{\{(Y-t) > 0\}} \mid X = x \right]. \end{aligned}$$

If $t_1 > t_2$, we get, $-(y - t_1) \geq -(y - t_2)$ and

$$\int_{-\infty}^{t_2} -(y - t_2) F_Y^x(y) \leq \int_{-\infty}^{t_2} -(y - t_1) dF_Y^x(y),$$

where $F_Y^x(\cdot)$ is the distribution function of Y conditional on $\{X = x\}$: then

$$\int_{-\infty}^{t_2} -(y - t_2) dF_Y^x(y) \leq \int_{-\infty}^{t_1} -(y - t_1) dF_Y^x(y),$$

which implies that

$$\mathbb{E} \left(-(Y - t_2) \mathbb{1}_{(Y-t_2) \leq 0} \mid X = x \right) \leq \mathbb{E} \left(-(Y - t_1) \mathbb{1}_{(Y-t_1) \leq 0} \mid X = x \right).$$

Then, $G^1(\cdot; x)$ is an increasing function. Similarly, if $t_1 > t_2$ for $G^2(\cdot; x)$

$$\begin{aligned} \int_{t_2}^{+\infty} (y - t_2) dF_Y^x(y) &\geq \int_{t_2}^{+\infty} (y - t_1) dF_Y^x(y) \\ &\geq \int_{t_1}^{+\infty} (y - t_1) dF_Y^x(y), \end{aligned}$$

This implies that

$$\mathbb{E} \left((Y - t_2) \mathbb{1}_{(Y-t_2) > 0} \mid X = x \right) \geq \mathbb{E} \left((Y - t_1) \mathbb{1}_{(Y-t_1) > 0} \mid X = x \right).$$

We readily infer that $G^2(\cdot; x)$ is a decreasing function. Therefore we obtain that $G(\cdot; x)$ is an increasing function. \square

5.7.2 Bernstein's inequality.

Let Z_1, \dots, Z_n be independent real zero mean random variables. If for all $m \geq 2$, there exists a constant $C_m > 0$, such that

$$\mathbb{E}|Z_1^m| \leq C_m a^{2(m-1)},$$

we have for $\epsilon > 0$,

$$\mathbb{P} \left(\left| \sum_{i=1}^n Z_i \right| > \epsilon n \right) \leq 2 \exp \left(\frac{-\epsilon^2 n}{2a^2(1 + \epsilon)} \right). \quad (5.66)$$

5.7.3 Lyapunov CLT.

In this variant of the central limit theorem, the random variables X_i have to be independent, but not necessarily identically distributed. The theorem also requires that random variables have moments of order $(2 + \delta)$, $\delta > 0$.

Theorem 5.4 *Suppose $\{X_1, X_2, \dots\}$ is a sequence of independent random variables, each with a finite expected value μ_i and variance σ_i^2 . Define*

$$s_n^2 = \sum_{i=1}^n \sigma_i^2.$$

If for some $\delta > 0$, the Lyapunov's condition

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E} \left[|X_i - \mu_i|^{2+\delta} \right] = 0,$$

is satisfied, then we have, as $n \rightarrow \infty$,

$$\frac{1}{s_n} \sum_{i=1}^n (X_i - \mu_i) \xrightarrow{\mathcal{D}} N(0, 1).$$

Bibliographie

- Abdous, B. and Rémillard, B. (1995). Relating quantiles and expectiles under weighted-symmetry. *Ann. Inst. Statist. Math.*, **47**(2), 371–384.
- Abdous, B., Ghoudi, K., and Rémillard, B. (2003). Nonparametric weighted symmetry tests. *Canad. J. Statist.*, **31**(4), 357–381.
- Aigner, D. J., Amemiya, T., and Poirier, D. J. (1976). On the estimation of production frontiers : maximum likelihood estimation of the parameters of a discontinuous density function. *Internat. Econom. Rev.*, **17**(2), 377–396.
- Al-Awadhi, F., Kaid, Z., Laksaci, A., Ouassou, I., and Rachdi, M. (2019). Functional data analysis : local linear estimation of the l_1 -conditional quantiles. *Statistical Methods and Applications*, **28**, 217–240.
- Aneiros, G. and Vieu, P. (2015). Partial linear modelling with multi-functional covariates. *Comput. Statist.*, **30**(3), 647–671.
- Aneiros, G., Cao, R., and Vieu, P. (2019a). Editorial on the special issue on functional data analysis and related topics. *Computational Statistics*, **34**(2), 447–450.
- Aneiros, G., Cao, R., Fraiman, R., Genest, C., and Vieu, P. (2019b). Recent advances in functional data analysis and high-dimensional statistics. *Journal of Multivariate Analysis*, **170**, 3 – 9. Special Issue on Functional Data Analysis and Related Topics.
- Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, **9**(3), 203–228.
- Bellini, F. and Bernardino, E. D. (2017). Risk management with expectiles. *The European Journal of Finance*, **23**(6), 487–506.
- Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quant. Finance*, **15**(5), 725–733.
- Bellini, F., Klar, B., Müller, A., and Rosazza Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.*, **54**, 41–48.
- Bellini, F., Bignozzi, V., and Puccetti, G. (2018). Conditional expectiles, time consistency and mixture convexity properties. *Insurance Math. Econom.*, **82**, 117–123.
- Crambes, C., Delsol, L., and Laksaci, A. (2008). Robust nonparametric estimation for functional data. *J. Nonparametr. Stat.*, **20**(7), 573–598.

- Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Comm. Statist. Theory Methods*, **41**(7), 1254–1268.
- Daouia, A., Girard, S., and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc., Ser. B, Stat. Methodol.*, **80**(2), 263–292.
- Daouia, A., Gijbels, I., and Stupfler, G. (2019). Extremiles : A new perspective on asymmetric least squares. *Journal of the American Statistical Association*, **114**(527), 1366–1381.
- Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statist. Sinica*, **1**(1), 93–125.
- Ehm, W., Gneiting, T., Jordan, A., and Krüger, F. (2016). Of quantiles and expectiles : consistent scoring functions, choquet representations and forecast rankings. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, **78**(3), 505–562.
- Emmer, S., Kratz, M., and Tasche, D. (2015). What is the best risk measure in practice? a comparison of standard measures. *Journal of Risk*, **18** :2, 31–60.
- Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Comm. Statist. Theory Methods*, **37**(16-17), 2735–2759.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Sankhyā*, **67**(2), 378–398.
- Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data : inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**(3), 267–286.
- Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference*, **140**(2), 335–352.
- Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.*, **106**(494), 746–762.
- Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics [editorial]. *Journal of Multivariate Analysis*, **146**, 1 – 6. Special Issue on Statistical Models and Methods for High or Infinite Dimensional Spaces.
- Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory*, **11**(1), 105–121.
- Holzmann, H. and Klar, B. (2016). Expectile asymptotics. *Electron. J. Stat.*, **10**(2), 2355–2371.
- Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer Series in Statistics. Springer, New York.
- Hsing, T. and Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester.
- Jones, M. C. (1994). Expectiles and M -quantiles are quantiles. *Statist. Probab. Lett.*, **20**(2), 149–153.

- Kara, L.-Z., Laksaci, A., Rachdi, M., and Vieu, P. (2017). Data-driven knn estimation in nonparametric functional data analysis. *Journal of Multivariate Analysis*, **153**, 176 – 188.
- Koenker, R. and Bassett, Jr., G. (1978). Regression quantiles. *Econometrica*, **46**(1), 33–50.
- Koenker, R. and Bassett, Jr., G. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, **50**(1), 43–61.
- Kokoszka, P. and Reimherr, M. (2017). *Introduction to functional data analysis*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL.
- Krätschmer, V. and Zähle, H. (2017). Statistical inference for expectile-based risk measures. *Scand. J. Stat.*, **44**(2), 425–454.
- Kuan, C.-M., Yeh, J.-H., and Hsu, Y.-C. (2009). Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*, **150**(2), 261 – 270. Recent Development in Financial Econometrics.
- Laksaci, A., Lemdani, M., and Ould-Saïd, E. (2009). A generalized L^1 -approach for a kernel estimator of conditional quantile with functional regressors : consistency and asymptotic normality. *Statist. Probab. Lett.*, **79**(8), 1065–1073.
- Laksaci, A., Lemdani, M., and Ould-Saïd, E. (2011). Asymptotic results for an L^1 -norm kernel estimator of the conditional quantile for functional dependent data with application to climatology. *Sankhya A*, **73**(1), 125–141.
- Ling, N. and Vieu, P. (2018). Nonparametric modelling for functional data : selected survey and tracks for future. *Statistics*, **52**(4), 934–949.
- Ling, N., Meng, S., and Vieu, P. (2019). Uniform consistency rate of knn regression estimation for functional time series data. *Journal of Nonparametric Statistics*, **31**(2), 451–468.
- Loève, M. (1963). *Probability theory*. Third edition. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London.
- Maume-Deschamps, V., Rullière, D., and Said, K. (2017). Multivariate extensions of expectiles risk measures. *Depend. Model.*, **5**(1), 20–44.
- Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, **55**(4), 819–847.
- Novo, S., Aneiros, G., and Vieu, P. (2019). Automatic and location-adaptive estimation in functional single-index regression. *Journal of Nonparametric Statistics*, **31**(2), 364–392.
- Raña, P., Aneiros, G., Vilar, J., and Vieu, P. (2016). Bootstrap confidence intervals in functional nonparametric regression under dependence. *Electron. J. Stat.*, **10**(2), 1973–1999.
- Rachdi, M. and Vieu, P. (2007). Nonparametric regression for functional data : Automatic smoothing parameter selection. *Journal of Statistical Planning and Inference*, **137**(9), 2784 – 2801.

- Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Second Edition. Springer, New York, NY.
- Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.*, **42**, 43–47.
- Shang, H. L. (2014). Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density. *J. Nonparametr. Stat.*, **26**(3), 599–615.
- Taylor, J. W. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, **6**(2), 231–252.
- Volkonskiĭ, V. A. and Rozanov, Y. A. (1959). Some limit theorems for random functions. I. *Teor. Veroyatnost. i Primenen*, **4**, 186–207.
- Waltrup, L., Otto-Sobotka, F., Kneib, T., and Kauermann, G. (2015). Expectile and quantile regression—david and goliath? *Statistical Modelling*, **15**, 433–456.
- Yao, Q. and Tong, H. (1996). Asymmetric least squares regression estimation : a nonparametric approach. *J. Nonparametr. Statist.*, **6**(2-3), 273–292.
- Ziegel, J. F. (2016). Coherence and elicibility. *Math. Finance*, **26**(4), 901–918.

Chapitre 6

Conclusion et Perspectives.

6.1 Conclusion

Dans cette thèse, les résultats que nous énonçons sont liés aux propriétés asymptotiques du modèle de régression dont lequel la variable aléatoire réponse Y est réelle tandis que la variable explicative est à valeurs dans un espace de dimension infinie. La méthode du noyau a été utilisée pour le modèle fonctionnel à indice simple (MFIS ou FSIM) et pour la régression expectile.

Dans la première partie de cette thèse, le MFIS a été étudié via la méthode du noyau de Nadaraya-Watson (NW) (cf. [Nadaraya \(1964\)](#) et [Watson \(1964\)](#)) pour estimer la fonction de régression dans le cas où les variables aléatoires sont fonctionnelles et quasi-associées. Dans cette dernière, nous avons établi la consistance et la normalité asymptotique de notre estimateur, et la construction de l'intervalle de confiance a été obtenue comme une conséquence directe de ce résultat.

Nous avons considéré ensuite une autre approche pour l'estimation de la fonction de régression, approche proposée par [Burba *et al.* \(2009\)](#) pour résoudre des préoccupations les plus importantes dans les problèmes de grande dimension est celui de la structure locale des données, et il apparaît qu'un outil bien adapté est les k plus proches voisins (k -NN). Cette méthode a été introduite comme une méthode alternative à la méthode du noyau de NW. Nous avons considéré le même modèle MFIS et nous avons étudié la consistance de la fonction de régression via la méthode k -NN.

Dans la deuxième partie de cette thèse, nous avons remarqué qu'à travers la littérature que l'utilisation de la régression classique dans plusieurs domaines est assez importante, néanmoins cette méthode a des points faibles pour capturer l'effet des régresseurs sur les queues de la distribution de la variable réponse, à moins que l'effet des régresseurs soit uniforme sur toute la distribution de la variable réponse. C'est dans ce contexte que d'autres méthodes de régression ont vu le jour, et qui sont capables d'estimer plus que la moyenne conditionnelle. Parmi ces méthodes, il y a la régression expectile (RE). Cette partie de thèse prolonge les avancées récentes de l'analyse des données fonctionnelles non-paramétriques, et nous pensons que ce sont les premiers résultats disponibles dans la littérature portant sur l'estimation non-paramétrique de l'expectile conditionnel via la méthode du noyau. Pour un état de l'art sur l'analyse des données fonctionnelles non paramétriques, on peut se référer à [Ling and Vieu \(2018\)](#).

Plus précisément, nous avons étudié l'estimation non-paramétrique de l'expectile conditionnel dans le cas où le régresseur X est fonctionnel et la variable réponse Y est réelle, pour des données complètes et indépendantes. Cette dernière partie donne les propriétés asymptotiques de l'estimateur et propose un estimateur efficace de la variance afin de fournir un intervalle de confiance pour l'expectile conditionnel. Le résultat du Théorème 5.1 de la consistance peut être considéré comme une généralisation du Théorème 6.1 dans Ferraty and Vieu (2006) sur l'opérateur de régression, puisque la régression expectile d'ordre $p = 0.5$ se réduit à la régression classique $\mathbb{E}[Y | X = x]$.

De plus, la normalité asymptotique dans le Théorème 5.2 généralise la normalité asymptotique de l'opérateur de régression obtenue dans le Théorème 2 de Ferraty *et al.* (2007). Il est important de noter que la normalité asymptotique présente un intérêt particulier pour certaines études asymptotiques diverses. En particulier, il peut être utilisé comme un résultat préliminaire pour prouver la consistance en moyenne quadratique ou pour prouver l'intégrabilité uniforme de l'estimateur, on peut se référer à Crambes *et al.* (2008) pour plus de détails et une discussion approfondie. En fin, nous avons examiné la mise en œuvre de ce modèle dans la pratique. En particulier, nous avons concentré sur le domaine de l'analyse des risques pour comparer entre la régression quantile et expectile.

Ces résultats ouvrent la porte à de nouvelles perspectives, d'ordre théorique (uniformité des résultats, allègement de l'hypothèse de dépendance au cadre α -mélangeant,...) ou bien d'ordre appliqué (choix du paramètre de lissage...). Dans la section suivante, nous listons quelques exemples de ces perspectives.

6.2 Perspectives

Il sera intéressant, pour conclure les travaux de cette thèse, d'envisager quelques développements futurs possibles en vue d'améliorer et d'étendre nos résultats

Sur le paramètre de lissage :

- Étudions le problème de la sélection du paramètre de lissage dans la régression des expectiles, un tel problème est primordial pour une meilleure utilisation de cet estimateur en pratique,
- Le choix du paramètre k dans le modèle fonctionnel à indice simple pour la méthode des k plus proches voisins (k -NN).

Sur les données fonctionnelles :

- Étendre nos résultats de l'expectile conditionnel à d'autres cas de séries temporelles fonctionnelles (ergodique, mémoire longue, β -mélangeantes, processus associés, et dans le cas des données censurées ...).

Sur la normalité asymptotique :

- Nous étudions, dans un futur proche, la normalité asymptotique de la méthode des k -NN dans le modèle MIFS pour les variables aléatoires associées ou processus β -mélangeant. Pour cela, les travaux récemment développés dans les statistiques fonctionnelles par Kara *et al.* (2017) et Ling *et al.* (2019) peut être un meilleur moyen de surmonter ce problème.
- Nous étudions aussi, dans un futur proche, la normalité asymptotique de l'expectile conditionnel dans le cas de mélange.

Sur l'expectile conditionnel multivarié :

- Montrons la consistance de l'expectile conditionnel uni et multivariés par la méthode k -NN,

- L'un des avantages de la régression expectile est sa possibilité de se comporter avec différentes stratégies statistiques (modélisation paramétrique, semi-paramétrique ou non-paramétrique) dans la même perspective que dans la régression et le quantile conditionnel. Ainsi, nous pouvons dire que la version linéaire ou semi-paramétrique de l'expectile fonctionnel est également une perspective naturelle de la contribution actuelle. Donc, Une attention particulière sera accordée à l'approche semi-paramétrique qui est en fait la tendance de l'analyse des données fonctionnelles (voir [Aneiros and Vieu \(2015\)](#) et [Novo et al. \(2019\)](#) pour plus de discussions et de motivations sur ce sujet),
- Généralisons l'expectile au cadre multidimensionnel où $\mathbf{Y} \in \mathbb{R}^d$ dans les deux cas i.i.d. et mélange. Pour cela, en se basant sur l'expectile multidimensionnel du papier de [Maume-Deschamps et al. \(2017\)](#), nous introduisons l'expectile conditionnel multidimensionnel : Soit la norme $\| \cdot \|$ dans \mathbb{R}^d . Nous désignons par $(\mathbf{Y})_+$ le vecteur $(\mathbf{Y})_+ = ((Y_1)_+, \dots, (Y_d)_+)^T$ et par $(\mathbf{Y})_-$ le vecteur $(\mathbf{Y})_- = (Y_1)_-, \dots, (Y_d)_-)^T$. Nous définissons la fonction de risque suivante

$$s_\alpha(\mathbf{Y}, \mathbf{y}) = \alpha \| (\mathbf{Y} - \mathbf{y})_+ \|^2 + (1 - \alpha) \| (\mathbf{y} - \mathbf{Y})_+ \|^2,$$

pour tout $\mathbf{y} \in \mathbb{R}^d$. Sous certaines conditions sur la norme pour garantir la stricte convexité de la fonction s_α , nous définissons *l'expectile conditionnel multivarié* à seuil commun par

$$\mathbf{y}^* \in \arg \min_{\mathbf{y} \in \mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{Y}, \mathbf{y}) \mid X = x].$$

À long terme, cette extension sera intéressante dans une future enquête.

Sur le quantile conditionnel :

- La construction des intervalles de confiance ponctuels en combinant les idées de bootstrap utilisées par [Raña et al. \(2016\)](#) sur l'espérance conditionnelle à celles utilisées par [Hahn \(1995\)](#) sur la régression quantile,
- Généralisons le quantile conditionnel au cadre multidimensionnel, et nous traitons dans ce cas les propriétés asymptotiques de l'estimateur construit à partir de la méthode du noyau de NW et par la méthodes des k -NN pour des données i.i.d. et dépendantes.

Bibliographie

- Aneiros, G. and Vieu, P. (2015). Partial linear modelling with multi-functional covariates. *Comput. Statist.*, **30**(3), 647–671.
- Burba, F., Ferraty, F., and Vieu, P. (2009). k-nearest neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics*, **21**(4), 453–469.
- Crambes, C., Delsol, L., and Laksaci, A. (2008). Robust nonparametric estimation for functional data. *J. Nonparametr. Stat.*, **20**(7), 573–598.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer Series in Statistics. Springer, New York. Theory and practice.
- Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric regression on functional data : inference and practical aspects. *Aust. N. Z. J. Stat.*, **49**(3), 267–286.
- Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory*, **11**(1), 105–121.
- Kara, L.-Z., Laksaci, A., Rachdi, M., and Vieu, P. (2017). Data-driven knn estimation in nonparametric functional data analysis. *Journal of Multivariate Analysis*, **153**, 176 – 188.
- Ling, N. and Vieu, P. (2018). Nonparametric modelling for functional data : selected survey and tracks for future. *Statistics*, **52**(4), 934–949.
- Ling, N., Meng, S., and Vieu, P. (2019). Uniform consistency rate of knn regression estimation for functional time series data. *Journal of Nonparametric Statistics*, **31**(2), 451–468.
- Maume-Deschamps, V., Rullière, D., and Said, K. (2017). Multivariate extensions of expectiles risk measures. *Depend. Model.*, **5**(1), 20–44.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, **9**(1), 141–142.
- Novo, S., Aneiros, G., and Vieu, P. (2019). Automatic and location-adaptive estimation in functional single-index regression. *Journal of Nonparametric Statistics*, **31**(2), 364–392.
- Raña, P., Aneiros, G., Vilar, J., and Vieu, P. (2016). Bootstrap confidence intervals in functional nonparametric regression under dependence. *Electron. J. Stat.*, **10**(2), 1973–1999.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhyā : The Indian Journal of Statistics, Series A (1961-2002)*, **26**(4), 359–372.