

N° d'ordre :

**Université de Saida– Dr. Moulay Tahar  
Faculté des Sciences**

## **Thèse**

Présentée pour obtenir le diplôme de

## **Doctorat de 3<sup>ème</sup> Cycle**

**Spécialité : Analyse Stochastique et Statistique des Processus**

**Filière : Mathématiques**

Par :

**Bouanani Oussama**

Thème :

**Contribution aux modèles conditionnels  
par la méthode des polynômes locaux**



Thèse soutenue le 23/06/2020 devant le jury composé de :

<b>N°</b>	<b>Nom et prénom</b>	<b>Grade</b>	<b>Etablissement</b>	<b>Qualité</b>
01	Tebboune Fethallah Ouahbi	Prof.	Université de Saida – Dr. Moulay Tahar	Président
02	Rahmani Saâdia	MCA.	Université de Saida – Dr. Moulay Tahar	Rapporteur
03	Laksaci Ali	Prof.	Université Djillali Liabès de Sidi Bel Abbès	Co-rapporteur
04	Bouchentouf Amina Angelika	Prof.	Université Djillali Liabès de Sidi Bel Abbès	Examineur
05	Kandouci Abdeldjebbar	Prof.	Université de Saida – Dr. Moulay Tahar	Examineur
06	Rachdi Mustapha	Prof.	Université Grenoble-Alpes- France	Invité

---

## *Dédicaces*



*Je dédie ce travail qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents, **Berrezoug, Chérifi khedidja** qui ne cessent de me donner avec amour le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux vous protège et que la réussite soit toujours à ma portée pour que je puisse vous combler de bonheur.*

*Je dédie aussi ce modeste travail à :*

- ▷ *Mes grands-parent et en particulier à mon défunt grand-père.*
- ▷ *Ma soeur Imen et sa famille et ma petite soeur Khaoither.*
- ▷ *Ma femme Khansaa et sa famille.*
- ▷ *Tous mes ami(e)s, mes collègues et tous ceux qui m'estiment.*
- ▷ *Et en dernier à mon neveu le petit prince Younès.*

*Bouanani Oussama*

---

## *Remerciements*

En tout premier lieu, je remercie le bon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

Qu'il me soit permis d'exprimer ma gratitude envers ma directrice de thèse Mademoiselle *Saâdia Rahmani*, pour ses multiples conseils et pour toutes les heures qu'elle a consacrées à diriger cette recherche.

J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et son respect. Me permettre de travailler sous sa direction fut le plus grand honneur que l'on m'ait fait.

Mes remerciements vont également à mon co-directeur de thèse Monsieur le professeur *Ali Laksaci* qui, par ses écrits et ses conseils a guidé mes réflexions et m'a accompagné pas à pas dans mon travail. Son énergie et sa confiance ont été des éléments moteurs pour moi.

J'aimerais également adresser mes remerciements les plus profonds à Monsieur le professeur *Mustapha Rachdi*, pour son soutien constant durant ces trois années de travail, ainsi que pour ses conseils avisés et son écoute qui ont été prépondérants pour la bonne réussite de cette thèse.

Je remercie les membres du Laboratoire de Modèles stochastiques, Statistique et Applications de l'Université de Saida, ainsi qu'à Monsieur le professeur *Abdeldjebbar Kandouci* le directeur du laboratoire. Je le remercie pour son accueil chaleureux à chaque fois que j'ai sollicité son aide, ainsi que pour ses multiples encouragements, notamment lors de mon premier exposé au laboratoire. Je le remercie également d'avoir accepté d'examiner mon modeste travail.

Je remercie chaleureusement le Président du jury, *Tebboune Fethallah Ouahbi*, qui a eu la délicate mission d'accorder les différentes voix présentes dans le jury.

De même, je suis grandement reconnaissant à Professeur *Angelika Amina Bouchentouf* de l'intérêt et du temps qu'elle a bien voulu accorder à l'expertise de cette thèse, en acceptant d'en être examinatrice.

Enfin, je tiens à remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de cette thèse.

# Table des figures

1.1	Courbes spectrométriques. . . . .	10
1.2	Consommation d'électricité mensuelle au USA entre janvier 1973 et février 2001. . . . .	10
1.3	Courbes correspondantes au phénomène El Niño. . . . .	11
1.4	Courbes $X_i$ , pour $i = 1, \dots, 250$ . . . . .	12
1.5	Courbes $X_i$ , pour $i = 1, \dots, 200$ . . . . .	13
2.1	A sample of 100 curves . . . . .	48
2.2	The asymptotic distribution : conditional distribution (left plot) and the conditional density (right plot) . . . . .	50
2.3	Near-infrared spectra curves . . . . .	51
2.4	Predicted values ( $y$ -axis) versus Test values ( $x$ -axis) results : local linear method (left plot) and the kernel smoothing approach (right plot) . . . . .	52

## Notations

Nous adopterons les notations ci-dessous dans les différents chapitres de ce manuscrit.

### *Abréviations et Symboles*

$\mathfrak{F}$	Espace semi-métrique.
$d(., .)$	Semi-métrique sur $\mathfrak{F}$ .
$X$	Variable aléatoire fonctionnelle, (v.a.f).
$X_i, i = 1, \dots, n$	Échantillon de v.a.f.
$x$	Observations de v.a.f.
$Y$	Variable aléatoire réelle, (v.a.r).
$\mathbb{E}(X)$	Espérance mathématique de la variable aléatoire $X$ .
$Var(X)$	Variance de la variable aléatoire $X$ .
$:=$	Définition d'une quantité.
$\mathbb{E}(Y X)$	Espérance conditionnelle de $Y$ sachant $X$ .
$Var(Y X)$	Variance conditionnelle de $Y$ sachant $X$ .
$Cov(X, Y)$	Covariance entre les variables aléatoires $X$ et $Y$ .
$B(x, h)$	Boule de centre $x$ et de rayon $h$ dans l'espace $(\mathfrak{F}, d)$ .
$\phi_x(h)$	Mesure de probabilité ou la probabilité de concentration.
$F^x(y)$	Distribution conditionnelle de v.a.r. $Y$ sachant v.a.f. $X_i = x$ .
$\xrightarrow{D}$	Convergence en distribution.
$\xrightarrow{p}$	Convergence en probabilité.
$\xrightarrow{p.s.}$	Convergence presque sûre.
$\mathbb{1}_{\{\}}$	Fonction indicatrice.
$C$ où $C'$	Constantes positives réelles.
$u_n \simeq v_n$ où $u_n \cong v_n$	Existence de deux constantes $C$ et $C'$ telles que : $0 < C v_n < u_n < C' v_n$ .
$u_n = o(v_n)$	$\frac{u_n}{v_n} \xrightarrow{n \rightarrow +\infty} 0$ .
$u_n = O(v_n)$	Existence d'une constante $C > 0$ telle que : $u_n \leq C v_n$ .
$u_n = o_p(v_n)$	$\frac{u_n}{v_n} \xrightarrow{p, n \rightarrow +\infty} 0$ , avec $u_n$ et $v_n$ deux suites aléatoires.
$[x]$	Partie entière de $x$ .

# Table des matières

<b>1</b>	<b>Introduction et présentation</b>	<b>6</b>
1.1	Statistique non paramétrique pour des données fonctionnelles . . . . .	6
1.1.1	Quelques exemples sur les données fonctionnelles . . . . .	9
1.1.2	Génération des données fonctionnelles . . . . .	11
1.1.3	Présentation sur la proximité des objets fonctionnels . . . . .	13
1.2	Originalité et modernisation de la méthode des polynômes locaux . . . . .	16
1.2.1	Cas vectoriel . . . . .	16
1.2.2	Cas fonctionnel . . . . .	17
1.3	Organisation du manuscrit . . . . .	19
1.4	Brève présentation des résultats . . . . .	20
1.4.1	Présentation des modèles et ses estimateurs . . . . .	20
1.4.2	Résultats : Cas i.i.d. . . . .	22
1.4.3	Résultats : Cas $\alpha$ -mélangeant . . . . .	26
	<b>Bibliographie</b>	<b>29</b>
<b>2</b>	<b>Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics</b>	<b>37</b>
2.1	Introduction and motivations . . . . .	38
2.2	Conditional cumulative distribution estimation . . . . .	40
2.2.1	Assumptions and notations . . . . .	41
2.2.2	Comments on the assumptions . . . . .	42
2.2.3	Main results . . . . .	43
2.3	Conditional density and its derivatives . . . . .	44
2.3.1	Assumptions and main result . . . . .	44

2.4	Conditional mode estimation . . . . .	46
2.5	Computational studies . . . . .	47
2.5.1	On simulated data . . . . .	48
2.5.2	Real data application . . . . .	50
2.6	Proofs . . . . .	54
2.7	Appendix . . . . .	65
<b>Bibliographie</b>		<b>71</b>
<b>3</b>	<b>Local linear conditional cumulative distribution function with mixing data</b>	<b>74</b>
3.1	Introduction and motivations . . . . .	75
3.2	The model . . . . .	77
3.2.1	The estimate . . . . .	77
3.2.2	Assumptions and notations . . . . .	78
3.3	Main Results . . . . .	80
3.3.1	Proof of Theorem 11 . . . . .	81
3.4	Confidence Intervals . . . . .	82
3.5	Appendix . . . . .	83
3.5.1	Proof of Lemma 3.3.1. . . . .	83
3.5.2	Proof of Lemma 3.3.2 . . . . .	96
<b>Bibliographie</b>		<b>99</b>
<b>4</b>	<b>Asymptotic normality of conditional mode estimator for functional dependent data</b>	<b>103</b>
4.1	Introduction and motivations . . . . .	104
4.2	Local linear estimator construction . . . . .	106
4.3	Main result . . . . .	107
4.3.1	Assumptions and notations . . . . .	107
4.3.2	Comments on the assumptions . . . . .	109
4.3.3	Proof of Theorem 14. . . . .	109
4.4	Appendix . . . . .	110
<b>Bibliographie</b>		<b>120</b>

<b>Annexe</b>	<b>123</b>
4.5 Notions de convergence . . . . .	123
4.6 Quelques inégalités utiles . . . . .	124
4.7 Théorème central limite . . . . .	126
<b>Conclusion et perspectives</b>	<b>127</b>
<b>Bibliographie générale</b>	<b>129</b>



## ملخص

المشكلة التي نتناولها في هذه الأطروحة هي الحالة الطبيعية المتقاربة لنماذج شرطية ذات متغير عشوائي دالي و متغير الاستجابة حقيقياً، باستعمال الطريقة الخطية المحلية. من الجانب التطبيقي لهذه النتيجة المتعلقة بالتقارب الطبيعي هي بناء فترات التنبؤ أو التحقق من الاختبارات الإحصائية. في البداية نعتبر متتالية من العينات المستقلة و تتبع نفس القانون. وضعنا الهدف المتمثل في إنشاء الحالة الطبيعية للمقدر محلي خطي لدالة التوزيع الشرطي والكثافة الشرطية ومشتقاتها وكذلك الوضع الشرطي. نقدم أيضاً تطبيق النتائج التي تم الحصول عليها على المحاكاة البيانات و البيانات الحقيقية. في المرحلة الثانية نعمم نتائجنا في حالة سلسلة من الملاحظات المرتبطة بنوع المعامل  $\alpha$ . نقدر وظيفة التوزيع الشرطي بالطريقة المحلية الخطية. بعد ذلك ندرس الحالة الطبيعية المتقاربة لهذا المقدر في ظل افتراضات تركيز القانون المشترك للملاحظات والفرضيات على معاملات المعامل  $\alpha$ . يمكن استخدام هذه النتيجة لمشكلة التنبؤ بالسلاسل الزمنية. في الجزء الثالث من هذه الأطروحة ، نعتبر نفس النوع من الملاحظات كما في الحالة السابقة ، و نهم على تقدير الوضع الشرطي بالطريقة المحلية الخطية و قدمنا في ظل ظروف معينة ، خاصية إحصائية تتعلق بالحالة الطبيعية المتقاربة لهذا المقدر.

## كلمات مفتاحية:

النماذج الشرطية ، التركيب متعدد الحدود المحلي ، تحليل البيانات الوظيفية (FDA) ، التقدير اللامعبي ، احتمال الكرة الصغيرة.

# Résumé

La problématique abordée dans cette thèse concerne la loi limite des estimateurs localement linéaires de certains modèles conditionnels, lorsque la variable explicative est de type fonctionnel et la variable réponse est réelle. Notons que la conséquence pratique de cette propriété asymptotique est la construction des intervalles de confiance où la réalisation des tests statistiques.

Nous supposons pour commencer que l'échantillon que nous étudions est constitué de variables indépendantes et identiquement distribuées. Nous avons fixé comme objectif l'établissement de la normalité asymptotique des estimateurs localement linéaires de la fonction de répartition conditionnelle, la densité conditionnelle et ses dérivées ainsi que le mode conditionnel. Les résultats obtenus sont également appliqués sur des données simulées puis réelles.

Dans la deuxième partie de ce manuscrit, nous généralisons nos résultats dans le cas où les observations sont fortement mélangeantes. Nous estimons la fonction de répartition conditionnelle par la méthode locale linéaire. Ensuite, nous étudions la normalité asymptotique de cet estimateur sous des hypothèses de concentration de la loi conjointe des observations et des hypothèses sur les coefficients de mélange. Ce résultat peut être utilisé pour le problème de la prévision en série chronologique.

Dans la troisième partie de cette thèse, nous considérons le même type de dépendance des observations que le cas précédent, en s'intéressant à l'estimation du mode conditionnel par la méthode locale linéaire, pour lequel nous avons établi sous certaines conditions, une propriété statistique qui concerne la loi limite de cet estimateur.

**Mots clés :** Modèles conditionnels, polynômes locaux, analyse de données fonctionnelles (FDA), estimation non paramétrique, probabilité de petite boules.

# Summary

The problem addressed in this thesis concerns the limit law of locally linear estimators of some conditional models, when the explanatory variable is of functional nature and the response variable is real. Note that the practical consequence of this asymptotic property is the construction of confidence intervals or statistical tests.

First, we consider a sequence of independent and identically distributed observations. We establish the asymptotic normality of the conditional distribution function estimators and the conditional density and its derivatives as well as the conditional mode. The established results were applied on both simulated and real data.

In the second part of this manuscript, we generalize our results in cases where the observations are strongly mixed. We estimate the conditional distribution function by the linear local method. Next, we study the asymptotic normality of this estimator under concentration assumptions of the joint law of observations  $(X_i, X_j)$  and assumptions about the mixing coefficients. This result can be used for the time series prediction problem.

In the third and last part of this thesis, we consider the same type of dependence of the observations as the previous case, by focusing on the estimation of the conditional mode by the linear local method, for which we have established under certain conditions, the limit law of this estimator.

**Key words :** Conditional models, local polynomial fitting, functional data analysis (FDA), nonparametric estimation, small ball probability.

# Liste des travaux

## I Publications dans des journaux internationaux

1. Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*, **46**, 199-233.
2. Bouanani, O., Rahmani, S., Hennani, L.A (2020). Local linear conditional cumulative distribution function with mixing data. *Arabian Journal of Mathematics*, **9** :289-307.
3. Bouanani, O., Rahmani, S., Laksaci, A., and Rachdi, M. (2020). Asymptotic normality of conditional mode estimator for functional dependent data. *India journal of pure and applied mathematics*. **51** (2) : 465-481.

## II Communications internationales

1. Participation à Onzième Rencontre d'Analyse Mathématique et Applications(RAMA11). " Asymptotic results of a local linear modelization of the conditional distribution for functional data" à Sidi Bel Abbès, Algérie, 21-24 novembre 2019.
2. Participation à "International Workshop on Perspectives on High-dimensional Data Analysis", Marrakech, 09-13 avril 2018.

## III Communications orales nationales

Participation aux séminaire hebdomadaire de LMSSA :

1. Introduction à la méthode des polynômes locaux dans le cas fonctionnel, 25 octobre 2017.
2. Propriétés asymptotiques des modèles conditionnels par la méthode locale linéaire, 30 septembre 2018.

#### **IV Actes de conférences**

1. Participation à l'école du printemps : "Les équations différentielles stochastiques rétrogrades et contrôle stochastique" à Saïda, Algérie, 22-26 avril 2018.
2. Participation à l'école d'automne : "L'inférence statistique et théorie de l'information pour des processus Markoviens et semi-Markoviens" à Saïda, Algérie, 23-28 octobre 2018.
3. Participation à l'école : "La méthodologie d'élaboration d'une thèse et d'une publication scientifique" à Bouira, Algérie, 9-10 décembre 2018.
4. Participation à l'école de Recherche CIMPA "Analyse stochastique et Applications" à Saïda, Algérie, 01-09 Mars 2019.

#### **V Projet de Recherche**

Membre du Projet de Recherche PRFU : C00L03UN200120180002 "Statistique fonctionnelle et étude des modèles de files d'attente". Agréé à partir de 01/01/2018.

# Introduction et présentation

## 1.1 Statistique non paramétrique pour des données fonctionnelles

L'analyse des données fonctionnelles est devenue ces derniers temps un outil important dans la recherche en statistique. Notons que la statistique fonctionnelle est étroitement liée à l'étude d'ensembles de données dans lesquelles les observations sont des courbes ou des surfaces. Il s'agit de données de dimension infinie, c'est à dire rentrant dans le champ de "grande dimension". Elles apparaissent de plus en plus fréquemment dans de nombreux domaines scientifiques grâce aux progrès techniques, à l'augmentation des capacités de stockage de l'information, à l'amélioration de l'outil informatique et de ses capacités de traitement, ainsi qu'à la multiplication des systèmes de monitoring et le perfectionnement des capteurs.

Ces nouvelles technologies récentes facilitent la collecte des données sur des grilles de plus en plus fines. Pratiquement les données sont représentées sous la forme d'un vecteur  $(X_i(t_{i1}), \dots, X_i(t_{ip}))_{1 \leq i \leq n}$ , où  $(t_{i1}, \dots, t_{ip})$  est une suite ordonnée et  $n$  la taille de l'échantillon. Intuitivement, on peut considérer ce type de données comme réalisations d'un processus stochastique  $\{X_t\}_{t \in T}$ , où  $T$  est un ensemble continu pouvant représenter par exemple un intervalle de temps ou un spectre de longueurs d'ondes. On parle alors d'analyse statistique de courbes ou d'analyse de données fonctionnelles. La littérature consacrée à l'étude des données fonctionnelles s'est considérablement développée. Historiquement les premiers développements statistiques sur les variables aléatoires fonctionnelles reviennent à la contribution de Rao [73] et celle

de Tucker [45] ; ces auteurs se sont intéressés à l'analyse en composantes principales. Frank et Friedman [47] ont comparé différentes méthodes de régression paramétrique pour variables fonctionnelles afin de développer la régression ridge (dont le principe a été posé par Hoerl et Kennard [51]). Helland [50] s'est intéressé à la régression des moindres carrés partiels, cet article fut suivi d'un commentaire de Hastie et Mallows [48] qui ont remarqué que ces méthodes, qui ne prennent pas en considération la nature fonctionnelle des données, ne sont pas adaptées lorsque les courbes ne sont pas lisses ou mesurées avec des erreurs lorsque la discrétisation des courbes n'est pas réalisée aux mêmes points, ce qui est souvent le cas en pratique. Par la suite, plusieurs auteurs ont proposé des modèles qui tiennent compte de la nature fonctionnelle des variables (voir Ramsay et Silverman [72], Cardot *et al.* [13], Bosq [12]).

Le traitement non paramétrique des données fonctionnelles est beaucoup plus récent que l'analyse paramétrique. Dans cette direction les premiers ouvrages dans ce sujet ont été introduits par Ferraty et Vieu [32]. Ces auteurs ont étudié certaines propriétés asymptotiques de l'estimateur à noyau de l'opérateur de régression dans le cas de données à dimension fractale. Ces résultats ont été étendus par Ferraty *et al.* [34] aux problèmes non standards de la régression telle que la prévision dans le contexte des séries chronologiques. Une année plus tard, Ferraty et Vieu [72] ont également fourni une nouvelle approche en modèle de régression pour une variable explicative à valeurs dans un espace semi-métrique, ils ont établi des résultats asymptotiques en utilisant de nouvelles méthodes sur l'estimateur de Nadaraya-Watson dans le cas dépendant. En suite, Dabo-Niang et Rhomari [17] se sont focalisés sur la convergence en norme  $L^p$  de l'estimateur à noyau de la régression non paramétrique. La convergence presque complète pour le cas fortement mélangeant a été étudié par Ferraty *et al.* [36]. Plus généralement, il faut noter que la contribution de Ferraty et Vieu [38] peut être considérée comme déterminante dans le cadre non paramétrique fonctionnel. Ces auteurs ont montré la convergence presque complète des estimateurs à noyau de la plus part des modèles conditionnels en précisant leurs vitesses de convergence dans le cas où les observations sont i.i.d. En 2007, Dabo-Niang et Laksaci [18] ont ajouté des résultats sur la convergence en norme  $L^p$  de l'estimateur à noyau du mode conditionnel dans le cas i.i.d. Laksaci et Meref [54] ont traité la convergence presque complète de l'estimateur à noyau de la fonction de répartition conditionnelle et les quantiles conditionnels, dans le cas de variables fonctionnelles spatialement dépendantes. En 2011 Dabo-Niang *et al.* [21] ont étudié la convergence

presque sûre d'un estimateur à noyau pour la fonction de régression. Une année plus tard, Laksaci et *al.* [56] ont établi la convergence presque complète de l'estimateur local linéaire de la fonction de répartition conditionnelle et des dérivées successives de la densité conditionnelle. Tandis que la fonction de hasard conditionnelle a été étudiée par Ferraty et *al.* [39], en établissant la convergence presque complète d'un estimateur à noyau de ce modèle non paramétrique.

Récemment, et en utilisant des conditions sur l'uniformité de l'intégral de l'entropie, Kara-Zaitri et *al.* [53] ont étudié la convergence uniforme par rapport au paramètre de lissage en estimation non paramétrique de plusieurs modèles statistiques telles que la régression, la densité conditionnelle, la fonction de hasard conditionnelle et la fonction de répartition conditionnelle.

En ce qui concerne la normalité asymptotique des estimateurs à noyau du mode conditionnel et des quantiles conditionnels, plusieurs auteurs se sont intéressés à cette propriété asymptotique, nous nous référons à Ezzahrioui et Ould-saïd [25, 26, 27] qui ont traité les deux cas i.i.d et dépendant. Un estimateur alternatif des quantiles conditionnels a été proposée par Lemdani et *al.* [59], ces auteurs ont traité les quantiles conditionnels comme un modèle robuste appartenant à la classe des M-estimateurs. Les résultats asymptotiques de cet article sont la convergence presque complète et la normalité asymptotique dans le cas i.i.d. La fonction de hasard conditionnelle a été étudiée par Quintela-Del-Río [71] en considérant des observations  $\alpha$ -mélangeantes, cet auteur a établi la convergence presque complète et la normalité asymptotique et a illustré ces résultats asymptotiques par une application sur des données sismiques. Cette propriété asymptotique pour le cas de la régression robuste a été étudiée par Attouch et *al.* [3]. Dabo-Niang en collaboration avec Laksaci [19] ont généralisé leur résultats sur la convergence en norme  $L^p$  de l'estimateur à noyau du mode conditionnel pour des données fortement mélangeantes. Notons que parmi les contributions les plus récentes sur ce sujet, nous pouvons citer Ardjoun et *al.* [2] qui ont établi la normalité asymptotique de l'estimateur à noyau récursive du mode conditionnel dans le cas ergodique.

La littérature sur le cas d'une variable réponse fonctionnelle est très restreinte en statistique fonctionnelle. On citera, dans ce contexte l'article de Van Keilegom en collaboration avec Ferraty et Vieu [45] pour la normalité asymptotique de l'estimateur de la fonction de régression, lorsque les observations sont indépendantes et identiquement distribuées. Le cas dépendant a été récemment considéré par Ferraty et *al.*



[43]. Dans ce travail, les auteurs ont établi la convergence presque complète de l'estimateur à noyau de l'opérateur de régression pour des observations  $\beta$ -mélangeantes. D'un point de vue pratique, les données fonctionnelles se retrouvent dans de nombreuses applications. Parmi les plus fameuses, nous pouvons citer : Müller et al. [68] sur des données biologiques, Chiou et al. [14] sur des données démographiques, l'analyse des mouvements en biomécanique par sørensen et al. [76], l'étude des profils de croissance par sauder et al. [75], ainsi que l'étude de l'évolution des taux d'intérêt au cours de temps par laurini [57].

### 1.1.1 Quelques exemples sur les données fonctionnelles

Le premier exemple concerne des données spectrométriques, (voir Ferraty et al. [33]). La spectroscopie dans le proche infrarouge est une technique analytique basée sur le principe d'absorption des rayonnements par la matière organique en utilisant un spectromètre. Cette méthode d'absorption est utilisée pour estimer la composition chimique des aliments (fourrages, viandes, ...). On s'intéresse à un échantillon de 215 morceaux de viandes finement hachées contenant chacun un certain taux de matière grasse (lipide). Pour chaque morceau, on observe le spectre dans le proche infrarouge pour 100 longueurs d'onde  $(\lambda_j)_{j=1,\dots,100}$  réparties entre 850 et 1050 nanomètres avec un pas constant (Figure 1.1). On observe donc pour chaque morceau  $i$ , une famille " discrète "  $X_i^d = \{X_i(\lambda_j)\}_{j=1,\dots,100}$ , que l'on peut considérer comme une version discrétisée de la variable fonctionnelle  $X_i = \{X_i(\lambda)\}_{\lambda \in [850;1050]}$ . En outre, pour chaque unité  $i$ , nous avons la matière grasse de  $y_i$  qui est obtenue par un traitement chimique analytique. Le fichier <https://npfda-spectrometric.dat> contient le couple  $(x_i, y_i)_{i=1,\dots,215}$ . Ces données réelles ont été utilisées dans le cas où les variables sont indépendantes.

Un deuxième exemple est tiré des relevés de consommation d'électricité mensuelle des secteurs résidentiels et commerciaux des États-Unis, entre janvier 1973 et février 2001. Chacune des 28 courbes (Figure 1.2) est alors composée de 12 points représentant les relevés de chaque année. Ces données réelles ont été utilisées dans le cas où les variables sont dépendantes, et leurs descriptions sont disponibles sur le site du centre de prévision du climat américain : <http://www.cpc.ncep.noaa.gov/data/indices/>.

Le troisième exemple concerne l'étude du phénomène El Niño : El Niño désigne

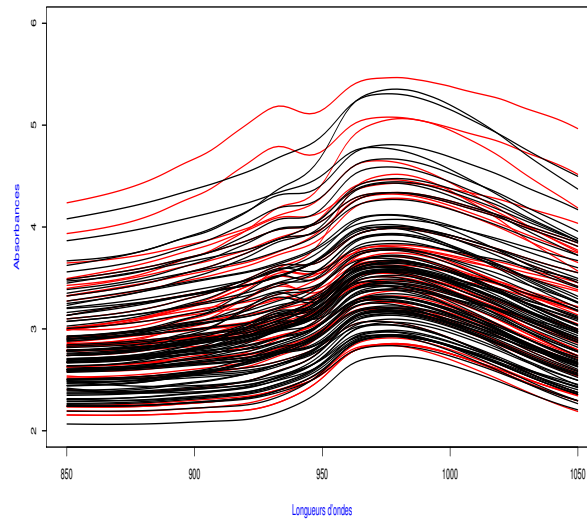


FIGURE 1.1 – Courbes spectrométriques.

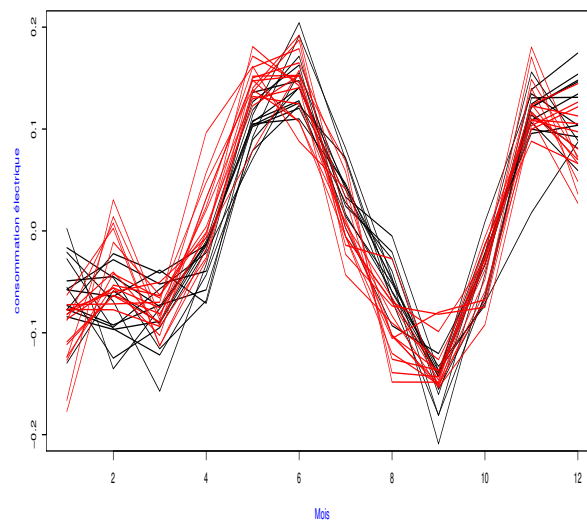


FIGURE 1.2 – Consommation d'électrique mensuelle au USA entre janvier 1973 et février 2001.

à l'origine un courant côtier saisonnier chaud au large du Pérou et de l'Équateur mettant fin à la saison de pêche. Le terme désigne maintenant par extension un phénomène climatique particulier, différent du climat usuel, il se caractérise par des

températures anormalement élevées de l'eau dans la partie Est de l'océan Pacifique Sud, représentant une extension vers le Sud du courant chaud péruvien. Il a été relié à un cycle de variation de la pression atmosphérique globale entre l'est et l'ouest du Pacifique, nommé l'oscillation australe, et les deux phénomènes sont réunis sous le titre de ENSO (El Niño / Southern Oscillation).

Ce jeu de données dont nous disposons 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 (Figure 1.3). Ces données et leurs descriptions sont disponibles sur le site du centre de prévision du climat américain : <http://www.cpc.ncep.noaa.gov/data/indices/>.

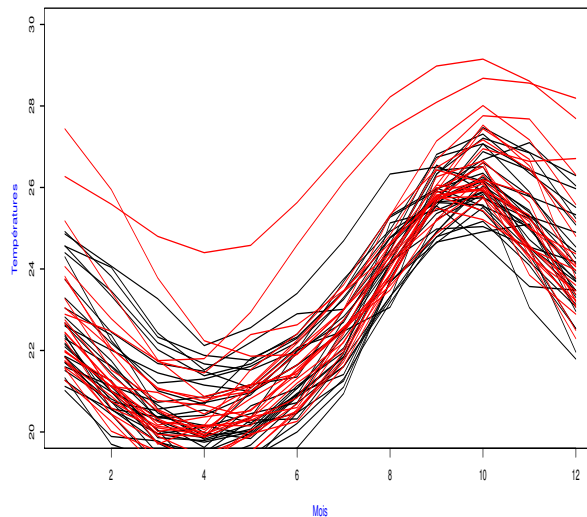


FIGURE 1.3 – Courbes correspondantes au phénomène El Niño.

### 1.1.2 Génération des données fonctionnelles

Deux exemples de simulation sont étudiés, le premier correspond au cas i.i.d. En pratique l'expression fonctionnelle des courbes observées est inconnue et nous avons uniquement accès aux observations discrètes mesurées à des temps précis. Nous simulons 250 variables  $a_i$  (respectivement,  $w_i$ ) indépendantes, uniformément

distribuées sur l'intervalle  $[1.5, 2.5]$ , (respectivement,  $[0.2, 2]$ ). Nous supposons que ces courbes sont observées sur une grille formée d'une discrétisation en 100 points de l'intervalle  $[0, \pi]$ . Considérons les courbes explicatives suivantes :

$$X_i(t) = a_i * \cos(w_i * t_j) + \sum_{k=1}^j W_{ik}, \quad i = 1, \dots, 250 \text{ et } j = 1, \dots, 100,$$

où les variables  $W_{ik}$  sont indépendantes, et distribuées selon la loi normal  $\mathcal{N}(0, 0.01)$ . La simulation de ces courbes est illustrée à la Figure 1.4

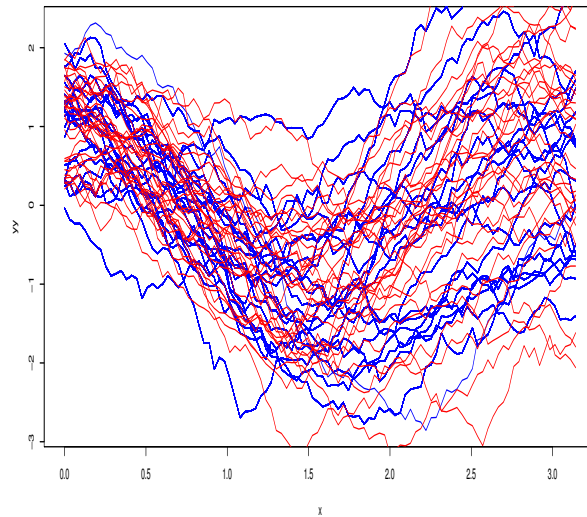


FIGURE 1.4 – Courbes  $X_i$ , pour  $i = 1, \dots, 250$ .

L'autre exemple considéré correspond au cas dépendant. Nous récoltons des observations  $\alpha$ -mélangeantes générées à l'aide du processus suivant :

$$b_i = \frac{1}{\sqrt{8}} * b_{i-1} + Z_i, \quad i = 1, \dots, 200,$$

où les variables  $b_i$  (respectivement,  $Z_i$ ) sont distribuées selon la loi normal  $\mathcal{N}(1, 0.4)$ , indépendamment de  $Z_i$  (respectivement, normal  $\mathcal{N}(0, 1)$ ). Nous simulons 200 variables  $a_i$  uniformément distribuées sur  $[0, 4\pi/3]$ . Ensuite, nous créons pour chaque  $a_i$  et  $b_i$  un vecteur  $(X_i(t_j))_{j=1, \dots, 200}$ , généré par :

$$t \rightarrow \exp \left( a_i * t^2 + \sin \left( b_i * \left( t - \frac{\pi}{3} \right) \right) \right), \quad i = 1, \dots, 200, \quad t \in [0, \pi/3].$$

Les courbes  $X_i$  sont représentées régulièrement sur une grille de discrétisation de 200 points dans l'intervalle  $[0, \pi/3]$ . La simulation de ces courbes est illustrée à la Figure 1.5

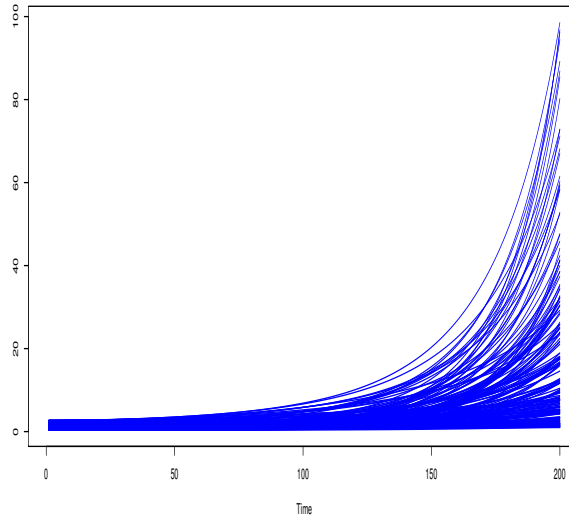


FIGURE 1.5 – Courbes  $X_i$ , pour  $i = 1, \dots, 200$ .

### 1.1.3 Présentation sur la proximité des objets fonctionnels

L'une des difficultés de l'approche statistique fonctionnelle vient de la nécessité de définir une mesure de proximité (i.e., distance) entre deux objets fonctionnels.

En effet, dans un espace de dimension finie toutes les distances sont équivalentes, cette propriété n'est plus valable dans un espace fonctionnel (i.e., dans un espace de dimension infinie). Ainsi, l'utilisation d'une distance standard peut provoquer l'apparition du "fléau de la dimension". Dans ce qui suit, nous rappelons quelques semi-métriques souvent utilisées dans le contexte non paramétrique fonctionnel. En considérant un espace fonctionnel  $\mathfrak{F}$ , une semi-métrique  $d$  est une application définie sur  $\mathfrak{F} \times \mathfrak{F}$  et à valeurs dans  $\mathbb{R}^+$  telle que :

- $\forall x \in \mathfrak{F}, d(x, x) = 0,$
- $\forall (x, \hat{x}) \in \mathfrak{F} \times \mathfrak{F}, d(x, \hat{x}) = d(\hat{x}, x),$
- $\forall (x, \hat{x}, x'') \in \mathfrak{F} \times \mathfrak{F} \times \mathfrak{F}, d(x, \hat{x}) \leq d(x, x'') + d(x'', \hat{x}).$

Ainsi, une distance peut-être vue comme une semi-métrique qui vérifie également l'axiome :

$$\forall (x, \hat{x}) \in \mathfrak{F} \times \mathfrak{F}, d(x, \hat{x}) = 0 \implies x = \hat{x}.$$

Par exemple, dans le cas d'un espace  $\mathfrak{F}_m = C^m[0, 1]$ , i.e., l'ensemble des fonctions  $m$  fois continûment différentiables sur  $[0, 1]$ ,  $m \geq 0$ ; et la semi-métrique  $d(., .)$  est définie comme :

$$d_m(x_1, x_2) = \sqrt{\int_0^1 (x_1^{(m)}(t) - x_2^{(m)}(t))^2 dt}, \quad x_1, x_2 \in \mathfrak{F}_m,$$

où  $x_1^m$  et  $x_2^m$  désignent les dérivés d'ordre  $m$  de  $x_1$  et  $x_2$ , respectivement. Dans le cas où  $m = 3$ , on déduit facilement que si  $x_1(t) = t$ ,  $x_2(t) = t^2$ ,  $t \in [0, 1]$  alors  $d_3(x_1, x_2) = 0$ . Donc  $(\mathfrak{F}_3, d_3)$  est un espace semi-métrique.

L'utilisation d'une semi-métrique au lieu d'une distance revient alors à réduire la dimension de l'espace fonctionnel, ce qui nous permet de limiter l'impact du fléau de la dimension .

Les trois principales familles de semi-métriques les plus utilisées dans la littérature statistique fonctionnelle sont respectivement basées sur la dérivation, l'analyse en composantes principales (voir Jolliffe, I [52]) et la semi-métrique de modèle d'indice fonctionnel. À titre d'exemple motivant l'introduction de cette première famille de pseudo-métriques, de nombreuses études présentant des applications sur des données spectrométriques ont constaté que de bien meilleurs résultats pouvaient être

obtenus en considérant comme variable explicative fonctionnelle la dérivée seconde des courbes brutes (voir Zhou, L et *al*). [85]).

La famille de semi-métriques  $d^{deriv}$  basée sur la dérivation se définit ainsi :

$$d_m^{deriv}(x_i, x_j) = \sqrt{\int_0^1 (x_i^{(m)}(t) - x_j^{(m)}(t))^2 dt},$$

où  $x_i^m$  désigne la dérivée d'ordre  $m$  de la fonction  $x_i$ . Ainsi, cette famille de semi-métriques est indexée par un entier  $m$ . Nous pouvons remarquer que la distance  $L^2$  est un cas particulier de cette famille lorsque  $m = 0$ . Le calcul répété des dérivées successives d'une fonction étant numériquement très sensible, il est préférable d'opérer une décomposition préalable des fonctions  $x_i$  et  $x_j$  dans une base de fonctions dont on connaît les dérivées exactes, comme par exemple la base B-spline.

La deuxième famille de semi-métriques  $d^{f pca}$  basée sur l'analyse en composantes principales fonctionnelles et indexée par un entier  $q$  indiquant les  $q$  premiers vecteurs propres de l'opérateur de variance-covariance empirique, associés aux  $q$  plus grandes valeurs propres.

$$d_q^{f pca}(x_i, x_j) = \sqrt{\sum_{k=1}^q \left( \int_0^1 (x_i(t) - x_j(t)) v_k(t) dt \right)^2}.$$

où  $v_1, \dots, v_q$  sont déduits de l'analyse en composantes principales de la variable fonctionnelle dont les  $x_i$  sont les réalisations. Cette méthode consiste à décomposer la variable fonctionnelle  $x$  selon des directions orthonormées  $v_1, v_2, \dots, v_q$  telles que :  $Var \left( \int v_q(t) x(t) dt \right)$  est maximale pour  $j = 1, \dots, q - 1$ ,  $\int v_j(t) v_q(t) = 0$ . Plus précisément cette technique maximise la variance expliquée par le sous-espace vectoriel engendré par  $v_1, \dots, v_q$ .

Cette famille de semi-métriques peut être utilisée seulement si les données sont équilibrées (les courbes sont observées aux mêmes points). Ceci pourrait apparaître comme un inconvénient pour l'usage d'un tel genre de cette famille mais, son avantage principal est d'être utilisée même si les courbes son irrégulières.

La troisième famille c'est la semi-métrique de modèle d'indice fonctionnel : pour une variable fonctionnelle  $x$  dans un espace de Hilbert  $\mathbb{H}$  muni du produit scalaire

$\langle \cdot, \cdot \rangle$ , on définit la semi-métrique de modèle d'indice fonctionnel par :

$$\forall x, \theta, y \in \mathbb{H} \quad , d_{\theta}(x, y) = | \langle x - y, \theta \rangle |.$$

Par des techniques de type validation croisée on pourrait estimer  $\theta$  qui donne la "meilleure" prédiction.

## 1.2 Originalité et modernisation de la méthode des polynômes locaux

### 1.2.1 Cas vectoriel

La modélisation locale linéaire est une approche statistique alternative à la méthode de Nadaraya-Watson [83, 69], elle possède de nombreux avantages par rapport à cette dernière, notamment la plupart des estimateurs qui dépendent d'un paramètre de lissage, ce qui aboutit à une vitesse de convergence sur les points aux bords plus faible que ceux de l'intérieur. Dans la littérature, ce problème est désigné par l'effet de bord. Il est donc utile de faire quelques modifications pour traiter ce problème. La régression polynomiale locale est encore attractive dans ce cas là, Tibshirani et Hastie [79] ont remarqué empiriquement que la régression locale linéaire corrige automatiquement les effets de bord. Ce fait a été montré théoriquement par Fan et Gijbels [29]. Ruppert et Wand [74] qui ont étendu ce résultat au cas des estimateurs polynomiaux locaux.

Historiquement, les premières applications de l'estimation polynomiale locale reviennent à Stone [77], cette contribution est déterminante aussi bien sur la modélisation locale linéaire que sur l'analyse non paramétrique d'une manière plus générale. En effet, Stone [77] a déterminé la vitesse de convergence minimax d'une famille d'estimateurs construits à partir d'une fonction de poids, puis il a donné les conditions nécessaires et suffisantes permettant d'obtenir la consistance de ces estimateurs. Lejeune [58] a étudié le comportement de l'erreur moyenne quadratique et il a montré une équivalence entre la régression polynomiale locale et la régression par noyaux optimaux, qui toutes les deux éliminent le biais jusqu'à un ordre donné et sont de variance minimales. Tsybakov [80], a étudié l'optimisation de la vitesse de convergence de la méthode des polynômes locaux dans le cas de l'estimation robuste. Müller [67] quant à lui, a étudié l'efficacité de l'estimation par la



méthode des polynômes locaux. Fan [28] a étudié l'estimateur local linéaire de la régression en établissant la vitesse de convergence de l'erreur quadratique moyenne conditionnelle et celle de l'erreur quadratique moyenne conditionnelle intégrée. Dans ce même article, Fan a montré que le lissage local linéaire est asymptotiquement plus efficace que tout autre lissage, que ce soit par l'estimation à noyau, par projection orthogonale ou par des splines. Par ailleurs, concernant les propriétés minimax de l'estimateur local linéaire de la fonction de régression, nous nous référons à Fan [30]. La généralisation de ces derniers résultats ont été établis par Ruppert et Wand [74] au cas où la variable explicative (covariable) est multidimensionnelle. Deux ans plus tard Fan et Gijbels [31] ont montré que la régression polynomiale locale présente plusieurs avantages sur les autres méthodes de régression non paramétriques. Dans cette monographie, ces auteurs ont remarqué que la capacité de la méthode des polynômes locaux à contourner le principal inconvénient des méthodes à noyau classique en corrigeant automatiquement les effets de bords tout en conservant les propriétés d'optimalité théoriques. Dans la même année, Masry [62], a repris ce travail en étudiant la convergence forte et la normalité asymptotique de l'estimateur de la fonction de régression lorsque les observations sont fortement mélangeantes. Pour l'estimation locale linéaire des dérivées de la fonction de régression, nous référons aussi à Masry et Mielniczuk[63].

En résumé les vingt dernières années ont été marquées par un développement intéressant de l'estimation polynomiale locale, tant au point de vue théorique qu'appliqué.

### 1.2.2 Cas fonctionnel

La modélisation des polynômes locaux dans le cadre fonctionnel a été abordée très récemment. Les premiers ouvrages sur ce sujet reviennent à Baïllo et Grané [4]. Ils ont introduit une nouvelle technique de régression non paramétrique. Cette approche consiste à proposer un estimateur local linéaire de l'opérateur de régression lorsque la variable explicative prend ses valeurs dans un espace de Hilbert, ils ont étudié son comportement asymptotique en déterminant son biais ainsi que sa variance. Dans le même cas hilbertien, Berlinet et *al.* [7] ont construit un autre estimateur local linéaire de la fonction de régression basé sur l'inverse de l'opérateur de covariance. En outre il ont donné une majoration de l'erreur quadratique moyenne de cet estimateur. Dans le cas Banachique, Barrientos-Marin et *al.* [5] ont proposé une version alternative facile à utiliser même quand la covariable appartient à un

espace semi-métrique, et ils ont prouvé la convergence presque complète de l'estimateur construit, ils ont également obtenu sa vitesse de convergence. Depuis ce temps l'estimation polynomiale locale de la fonction de régression a connu un grand succès en dimension infinie. Cette technique a été appliquée pour l'estimation des autres modèles conditionnelle. Citons, Demongeot et *al.* [20] qui ont obtenu des estimateurs locaux linéaires de la densité conditionnelle quand la réponse est scalaire et la variable explicative est à valeurs dans un espace semi-métrique. Puis sous certaines conditions, ils ont montré la convergence presque complète ponctuelle et uniforme de l'estimateur construit et ils ont précisé la vitesse de convergence pour chaque mode. De plus, ils ont utilisé les résultats obtenus pour donner quelques propriétés asymptotiques de l'estimateur local linéaire du mode conditionnel. La convergence en moyenne quadratique de l'estimateur local linéaire de la fonction de répartition conditionnelle a été établie par Demongeot et *al.* [22]. Tous les résultats précités ont été obtenus dans le cas où les observations sont i.i.d. Dans le cas où les données sont spatialement dépendantes, Chouaf et Laksaci [15] quant à eux, ont généralisé les résultats dus à Barrientos et *al.*[5]. Laksaci et *al.* [55], ont montré également la convergence presque complète des estimateurs locaux linéaire de plusieurs modèles non paramétriques conditionnels, tels que la fonction de répartition conditionnelle, la densité conditionnelle, ainsi que le mode conditionnel. En 2015, Messaci et *al.* [66] ont étudié l'estimation locale linéaire des quantiles conditionnels, en établissant la convergence presque complète de l'estimateur construit. Récemment Demongeot et *al.* [23] ont étudié l'estimation non paramétrique de la fonction de régression, lorsque la variable réponse et la covariable sont fonctionnelles. Ces auteurs ont construit un estimateur local linéaire de l'opérateur de régression, ils ont évalué son erreur d'estimation ensuite ils ont montré la convergence presque complète et uniforme de l'estimateur construit puis, ils ont précisé sa vitesse de convergence. Notons que la vitesse de convergence de cet estimateur dépend des paramètres liés à la structure topologique de l'espace fonctionnel des variables et/ou à la méthode de construction de cet estimateur. Ces résultats confirment la supériorité de la méthode locale linéaire sur la méthode à noyau. En effet, la méthode à noyau fournit un biais d'ordre  $O(h)$ . En revanche, avec la méthode locale linéaire, on obtient un biais d'ordre  $O(h^2)$ . Par ailleurs, sur la partie variance on a la même constante que dans le cas multidimensionnel, mais dans le cadre fonctionnelle, le terme de variance est exprimé en fonction de la concentration  $\phi(h)$ . Dans [24] les mêmes auteurs ont étudié l'estimation non paramétrique du même estimateur, lorsque la variable explicative est dans

un espace semi-métrique et la variable réponse est dans un espace de Hilbert. Il ont montré la convergence presque complète et uniforme dans le cas où les observations sont i.i.d. Plus précisément, la vitesse de convergence est obtenu sous des conditions usuelles et elle est exprimée en fonction de la régularité du modèle, de la concentration de la variable fonctionnelle, ainsi que de l'entropie de l'espace fonctionnel. Dans la même année Leulmi et *al.* [60] ont étudié l'estimation locale linéaire de la fonction de régression généralisée dans le cas d'une variable réponse scalaire et d'une variable explicative à valeurs dans un espace semi-métrique et ils ont montré la convergence presque complète ponctuelle puis uniforme de l'estimateur construit.

En ce qui concerne les travaux de la normalité asymptotique, nous nous référons à Zhou and Linal [86]. Ces auteurs ont établi sous des conditions de régularité, la normalité asymptotique de l'estimateur local linéaire de l'opérateur de régression. Belarbi et *al.* [6] ont étudié l'estimation robuste du modèle de régression linéaire local fonctionnel et ils ont établi la convergence presque complète ainsi que la normalité asymptotique de l'estimateur construit. Tous les résultats précités ont été obtenus dans le cas où les observations sont i.i.d. En 2017 Xianzhu et *al.* [84], ont établi la normalité asymptotique d'un estimateur local linéaire de la densité conditionnelle dans le cas où les observations sont  $\alpha$ -mélangeantes.

### 1.3 Organisation du manuscrit

Cette thèse se décompose en quatre chapitres.

Nous avons commencé par un chapitre introductif, en s'intéressant d'abord aux notions élémentaires sur la statistique non paramétrique pour des données fonctionnelles. Cette partie a été suivie par une présentation de quelques exemples d'applications pour donner une idée sur la simulation des données réelles et des données générées par une fonction. Ensuite nous avons exposé un bref descriptif historique sur la méthode des polynômes locaux dans les deux cas vectoriel et fonctionnel. Enfin, cette introduction s'achève par une dernière partie incluant une synthèse méthodologique et une synthèse des résultats originaux obtenus dans cette thèse. Nous y offrons de nombreuses références bibliographiques sur le sujet.

Dans le deuxième chapitre nous traitons le cas où les observations sont indépendantes et identiquement distribuées. Comme résultats asymptotiques nous établissons la normalité asymptotique des estimateurs localement linéaires de la fonction de

répartition conditionnelle, la densité conditionnelle et ses dérivées ainsi que le mode conditionnel. Notons que la loi limite de ces estimateurs est donnée en fonction de la propriété de concentration de la variable fonctionnelle étudiée sur des petites boules. Nous avons illustré notre méthode par une application sur des données simulées et sur des données réelles, l'objectif est l'étude comparative entre la méthode locale constante et la méthode locale linéaire, et à travers la normalité asymptotique de ces estimateurs nous obtenons les intervalles de confiance pour nos valeurs prédites. Ce travail a fait l'objet d'une publication parue dans la revue internationale *Behaviormetrika*.

Dans le troisième chapitre nous étudions la même propriété asymptotique pour la fonction de répartition conditionnelle dans le cas des observations fortement mélangées. Plus précisément nous construisons un estimateur local linéaire adapté à cette situation pour lequel nous avons établie la normalité asymptotique. Ensuite nous déterminons une forme explicite de l'intervalle de confiance construit. Ce travail a fait l'objet d'une publication parue dans la revue internationale *Arabian Journal of Mathematics*.

Le quatrième chapitre est consacré au problème de l'estimation locale linéaire du mode conditionnel. On suppose qu'on dispose des observations fortement mélangées et on établit la loi limite de notre estimateur. Les résultats de ce chapitre ont été publiés dans la revue internationale *Indian Journal of Pure and Applied Mathematics*.

L'annexe est consacré aux outils mathématiques et techniques utilisées tout au long de cette thèse. Nous achevons ce manuscrit par une conclusion générale ainsi que quelques perspectives de recherche permettant d'étendre et parfois de généraliser les résultats de cette thèse.

## 1.4 Brève présentation des résultats

Nous donnons dans cette section une brève présentation des résultats obtenus dans cette thèse.

### 1.4.1 Présentation des modèles et ses estimateurs

Soit  $(X, Y)$  un couple de variables aléatoires à valeurs dans  $\mathfrak{F} \times \mathbb{R}$ , où  $\mathfrak{F}$  est un espace semi-métrique. On note  $d$  la semi métrique sur  $\mathfrak{F}$ . On désigne par  $F^x$  la

fonction de répartition conditionnelle de  $Y$  sachant  $X = x$  défini par

$$\forall x \in \mathfrak{F}, \forall y \in \mathbb{R}, \quad F^x(y) = \mathbb{P}(Y \leq y | X = x).$$

On suppose que cette distribution conditionnelle de  $Y$  sachant  $X$  est absolument continue par rapport à la mesure de Lebesgue dont on note  $f^x$  sa densité. On désigne par  $f^{x(j)}$  la dérivée d'ordre  $j$  de cette densité conditionnelle.

Soit  $(X_i, Y_i)_{i \in \mathbb{N}^*}$  des couples aléatoires ayant la même loi que  $(X, Y)$ , On construit un estimateur de la fonction de répartition conditionnelle notée  $\widehat{F}^x = \widehat{a}$  solution du problème des moindres carrés pondérés suivant :

$$\sum_{i=1}^n (H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K(h_K^{-1}\delta(x, X_i)), \quad (1.1)$$

où  $\beta(\cdot, \cdot)$  et  $\delta(\cdot, \cdot)$  sont des opérateurs connus de  $\mathfrak{F} \times \mathfrak{F}$  dans  $\mathbb{R}$ , tels que  $\forall z \in \mathcal{F}$ ,  $\beta(z, z) = 0$  et  $|\delta(\cdot, \cdot)| = d(\cdot, \cdot)$ . La fonction  $K$  est un noyau,  $H$  est une fonction de répartition et  $h_K = h_{K,n}$  (resp.  $h_H = h_{H,n}$ ) est une suite de nombres réels positifs. Formellement, l'estimateur fonctionnel local linéaire est défini par

$$\widehat{F}^x(y) = \frac{\sum_{i,j=1}^n w_{ij} H(h_H^{-1}(y - Y_j))}{\sum_{i,j=1}^n w_{ij}} = \frac{\sum_{j=1}^n \Delta_j K_j H_j}{\sum_{j=1}^n \Delta_j K_j}, \quad (1.2)$$

où  $w_{ij} = \beta_i(\beta_i - \beta_j) K_i K_j$  et  $\Delta_j = K_j^{-1} \left( \sum_{i=1}^n w_{ij} \right) = \sum_{i=1}^n \beta_i^2 K_i - \left( \sum_{i=1}^n \beta_i K_i \right) \beta_j$ , avec  $\beta_i = \beta(X_i, x)$ ,  $K_i = K(h_K^{-1}(\delta(x, X_i)))$  et  $H_j = H(h_H^{-1}(y - Y_j))$ .

De même, nous proposons de définir l'estimateur local linéaire de la densité conditionnelle et ses dérivées  $\widehat{f}^{x(j)}$  de  $f^{x(j)}$  par  $\widehat{f}^{x(j)} = \widehat{a}$  obtenu par la minimisation de :

$$\sum_{i=1}^n (h_H^{-j-1} H^{(j+1)}(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K(h_K^{-1}\delta(x, X_i)).$$

Cet estimateur peut être écrit sous la forme :

$$\widehat{f}^{x(j)}(y) = \frac{\sum_{i \neq k}^n W_{ik}(x) H^{(j+1)}(h_H^{-1}(y - Y_k))}{h_H^{j+1} \sum_{i \neq k}^n W_{ik}(x)} = \frac{\sum_{k=1}^n \Delta_k K_k H_k^{(j+1)}}{h_H^{j+1} \sum_{k=1}^n \Delta_k K_k}. \quad (1.3)$$

En ce qui concerne l'estimation du mode conditionnel on suppose qu'il existe un compact  $S$  où le mode est unique. On note par  $\theta$  ce mode. On estime ce dernier par l'approche locale linéaire par la variable aléatoire  $\hat{\theta}$  solution du problème d'optimisation suivant :

$$\hat{f}^x(\hat{\theta}(x)) = \sup_{y \in S} \hat{f}^x(y). \quad (1.4)$$

Dans la section suivante, nous expliciterons les résultats de la normalité asymptotiques de ces paramètres conditionnels.

### 1.4.2 Résultats : Cas i.i.d.

Nous supposons que les variables aléatoires fonctionnelles sont i.i.d. Pour une meilleure lisibilité, nous commençons par citer l'ensemble des notations et des hypothèses nécessaires pour établir la normalité asymptotique des estimateurs (1.2), (1.3) et (1.4).

#### Notations générales et hypothèses :

Dans ce qui suit, nous fixons  $x$  (resp.  $y$ ) dans  $\mathfrak{F}$  (resp.  $\mathbb{R}$ ), et nous désignons par  $\mathcal{V}_x$  (resp.  $\mathcal{V}_y$ ) un voisinage de  $x$  (resp.  $y$ ) et  $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(x, X) \leq r_1)$ .

Avant d'établir nos principaux résultats, nous introduisons les quantités  $M_j$  et  $N(a, b)$ , qui apparaîtront dans les termes dominants du biais et de la variance :

$$M_j = K^j(1) - \int_{-1}^1 (K^j(u))' \chi_x(u) du \quad \text{où } j = 1, 2,$$

$$\text{et pour tous } a > 0 \text{ et } b = 2, 4, \quad N(a, b) = K^a(1) - \int_{-1}^1 (u^b K^a(u))' \chi_x(u) du.$$

De plus, nous notons

$$\hat{F}^x(y) = \frac{\hat{F}_N^x(y)}{\hat{F}_D^x} \quad \text{et} \quad \hat{f}^x(y) = \frac{\hat{f}_N^x(y)}{\hat{F}_D^x},$$

$$\text{où } \hat{F}_N^x(y) = \frac{1}{n \mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_k K_j H_j, \quad \hat{f}_N^{x(j)}(y) = \frac{1}{n h_H^{j+1} \mathbb{E}(\Delta_1 K_1)} \sum_{k=1}^n \Delta_j K_k H_k^{(j+1)},$$

$$\hat{F}_D^x = \frac{1}{n \mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_j K_j,$$

et

$$\varphi_l(\cdot, y) = \frac{\partial^l f(\cdot)(y)}{\partial y^l}, \quad \psi_l(s) = \mathbb{E}[\varphi_l(X, y) - \varphi_l(x, y) | \beta(X, x) = s], \quad \text{avec } l \in \{j, j+2\}.$$

Pour énoncer le résultat concernant la loi limite de l'estimateur (1.2), nous avons utilisé les hypothèses suivantes :

(H1) Pour tout  $r > 0$ ,  $\phi_x(r) := \phi_x(-r, r) > 0$ , et il existe une fonction  $\chi_x(\cdot)$  telle

$$\text{que, pour tout } t \in [-1, 1], \quad \lim_{h_K \rightarrow 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \chi_x(t).$$

(H2) La fonction de répartition conditionnelle  $F^x$  est telle que : il existe

$$b_1 > 0, b_2 > 0, \quad \forall (y_1, y_2) \in \mathcal{V}_y \times \mathcal{V}_y \text{ et } \forall (x_1, x_2) \in \mathcal{V}_x \times \mathcal{V}_x$$

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C \left( |\delta(x_1, x_2)|^{b_1} + |y_1 - y_2|^{b_2} \right).$$

où  $C$  est une constante positive dépendante de  $x$ .

(H3) Les fonctions  $\delta(\cdot, \cdot)$  et  $\beta(\cdot, \cdot)$  sont telles que :

$$(i) \quad \forall z \in \mathfrak{F}, \quad |\delta(x, z)| = d(x, z) \text{ et } C_1 |\delta(x, z)| \leq |\beta(x, z)| \leq C_2 |\delta(x, z)|$$

où  $C_1 > 0, C_2 > 0$ .

$$(ii) \quad \sup_{u \in B(x, r)} |\beta(u, x) - \delta(x, u)| = o(r).$$

(H4) Le noyau  $K$  est une fonction positive, différentiable de support  $[-1, 1]$  et pour lequel la dérivée première  $K'$  satisfait

$$K^2(1) - \int_{-1}^1 (K^2(u))' \chi_x(u) du > 0.$$

(H5) Le noyau  $H$  est une fonction différentiable, telle que :

$$\int |t|^{b_2} H'(t) dt < \infty.$$

(H6) Les paramètres de lissage satisfaites : il existe un entier positif  $n_0$ , tel que,

$\forall n > n_0$  :

$$-\frac{1}{\phi_x(h_K)} \int_{-1}^1 \phi_x(zh_K, h_K) \frac{d}{dz} (z^2 K(z)) dz > C_3 > 0,$$

$$\lim_{n \rightarrow \infty} h_K = 0, \quad \lim_{n \rightarrow \infty} h_H = 0, \quad \lim_{n \rightarrow \infty} \frac{\log n}{n\phi_x(h_K)} = 0,$$

et

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

où  $dP(x)$  est la distribution cumulative de  $X$ .

Nous avons donc établi le résultat suivant.

**Théorème 1.1.** [9] *Sous les hypothèses (H1)-(H6), on a :*

$$\sqrt{n\phi_x(h_K)}(\widehat{F}^x(y) - F^x(y) - B_n(x, y)) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, y)),$$

où

$$V_{HK}(x, y) = \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)),$$

et

$$B_n(x, y) = \frac{\mathbb{E}(\widehat{F}_N^x(y))}{\mathbb{E}(\widehat{F}_D^x)} - F^x(y).$$

Pour établir la loi limite de  $\widehat{f}^{x(j)}(y)$  nous aurons besoin des hypothèses du Théorème (1.1) et des hypothèses supplémentaires suivantes :

(H1')  $\forall l \in \{j, j+2\}$ , la quantité  $\psi_l^{(2)}(0)$  existe, où  $\psi_l^{(2)}$  désigne la seconde dérivée de  $\psi_l$ .

(H6')  $\lim_{n \rightarrow +\infty} \frac{\log n}{nh_H^{2j+1}\phi_x(h_K)} = 0$ .

(H7) (i) Il existe une constante positive  $C$  telle que, pour tout  $(y_1, y_2) \in \mathbb{R}^2$ ,

$$|H^{(j+1)}(y_1) - H^{(j+1)}(y_2)| \leq C|y_1 - y_2|.$$

(ii) Il existe une constante positive  $\kappa$  tel que, pour tout  $j' \leq j$ ,

$$\lim_{y \rightarrow \infty} |y|^{1+\kappa} |H^{(j'+1)}(y)| = 0.$$

(iii)  $H^{(j+1)}$  est une fonction bornée, et  $\int (H^{(1)}(t))^\nu dt < \infty$  où  $\nu > 1$ .

On établit le résultat suivant.



**Théorème 1.2.** [9] Si les hypothèses (H1'), (H2)-(H5), (H6') et (H7), sont vérifiées, on a

$$\sqrt{nh_H^{2j+1}\phi_x(h_K)}(\widehat{f}^{x(j)}(y) - f^{x(j)}(y) - B_n(x, y)) \xrightarrow{D} \mathcal{N}(0, V_{H,K}(x, y)),$$

où

$$V_{H,K}(x, y) = \frac{M_2}{M_1^2} f^x(y) \int (H^{(j+1)}(t))^2 dt,$$

et

$$B_n(x, y) = \frac{\mathbb{E}(\widehat{f}_N^{x(j)}(y))}{\mathbb{E}(\widehat{F}_D^x)} - f^{x(j)}(y).$$

Afin d'établir la normalité asymptotique de  $\widehat{\theta}(x)$ , nous avons introduit les hypothèses supplémentaires suivantes :

(H9) Il existe un entier  $j > 0$ , tel que,  $f^x$  soit  $j$ -fois continûment différentiable par rapport à  $y$  sur  $[\theta - \xi, \theta + \xi]$  pour  $\xi > 0$ .

$$(H10) \begin{cases} f^{x(l)}(\theta) = 0 & \text{si } 1 \leq l < j, \\ |f^{x(j)}(\theta)| > 0. \end{cases}$$

(H11) Le paramètre de lissage  $h_H$  satisfait  $\lim_{n \rightarrow +\infty} nh_H^3 \phi_x(h_K) = 0$ .

**Théorème 1.3.** [9] Sous les hypothèses du Théorème 1.2 et si les hypothèses (H9)-(H11) sont vérifiées, on obtient

$$\left( \frac{nh_H^3 \phi_x(h_K) (f^{x(2)}(\theta(x)))^2}{V_{H,K}(x, \theta(x))} \right)^{1/2} (\widehat{\theta}(x) - \theta(x) - B_n(x, \theta(x))) \xrightarrow{D} \mathcal{N}(0, 1),$$

où

$$B(x, \theta(x)) = B_H(x, \theta(x))h_H^2 + B_K(x, \theta(x))h_K^2 + o(h_H^2) + o(h_K^2),$$

avec

$$B_H(x, \theta(x)) = \frac{1}{2} f^{x(3)}(\theta(x)) \int t^2 H^{(1)}(t) dt,$$

$$B_K(x, \theta(x)) = \frac{1}{2} \psi_1^{(2)}(0) \frac{N(1, 2)}{M_1},$$

et

$$V_{H,K}(x, \theta(x)) = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt.$$

Une discussion sur les hypothèses et les détails de la démonstration des Théorèmes (1.1), (1.2) et (1.3) sont exposés au Chapitre 2.

### 1.4.3 Résultats : Cas $\alpha$ -mélangeant

Afin de généraliser les résultats obtenus dans le cas i.i.d. à des observations dépendantes, nous renforçons les hypothèses précédentes, en ajoutant des hypothèses sur la concentration de loi conjointe du couple  $(X_i, X_j)$  et sur le coefficient de mélange.

**Notations générales et hypothèses :**

On note Pour tout  $l \in \{0, 2, 3\}$ , par :  $g_l(x, y) = \frac{\partial^l F^x(y)}{\partial y^l}$ ,  $\lambda_l(x, y) = \frac{\partial^l f^x(y)}{\partial y^l}$  et

$G_l(s) = \mathbb{E}[g_l(X, y) - g_l(x, y) | \beta(x, X) = s]$ ,  $\Lambda_l(s) = \mathbb{E}[\lambda_l(X, y) - \lambda_l(x, y) | \beta(x, X) = s]$ .

En suite, nous avons imposé les conditions suivantes afin d'établir la normalité asymptotique de l'estimateur local linéaire de la fonction de répartition conditionnelle (1.2).

(F1) Pour tout  $l \in \{0, 2\}$  la quantité  $G_l^{(2)}(0)$  existe.

(F2)  $\sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) \leq \psi_x(h_K)$ , où  $\psi_x(h_K) = O(\phi_x^2(h_K))$ .

(F3) Les coefficients de  $\alpha$ -mélangeant de la suite  $(X_i, Y_i)_{i \in \mathcal{N}^*}$  vérifient les conditions :

(i)  $\exists s > 1, \exists c > 0 : \forall n \in \mathbb{N}; \quad \alpha(n) \leq cn^{-s}$ ,

(ii)  $\sum_{k=1}^{+\infty} k^\delta (\alpha(k))^{\frac{1}{p}} < \infty$ , avec  $p > 0$  et  $\delta > \frac{1}{p}$ .

(F4) Le paramètre de lissage  $h_K$  vérifie :

(i)  $\lim_{n \rightarrow +\infty} n \phi_x^m(h_K) = +\infty$ , pour  $m = 1, 2$ ,

(ii)  $\lim_{n \rightarrow +\infty} n^\varsigma \phi_x(h_K) = 0$ , pour  $\varsigma = 2, 3$ ,

(iii)  $\exists C_1 > 0, \quad C_2 > 0, \quad C_2 n^{1-p} \leq \phi_x(h_K) \leq C_1 n^{\frac{1}{1-p}}$  pour  $p \geq 2$ .

(F5) Il existe deux suites positives  $(v_n)$  et  $(r_n)$  tendant vers l'infini avec  $n$  et satisfaisant :

$$(i) \quad v_n = o\left((n\phi_x(h_K))^{\frac{1}{2}}\right) \text{ et } \lim_{n \rightarrow +\infty} \left(\frac{n}{\phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) = 0,$$

$$(ii) \quad q_n v_n = o\left((n\phi_x(h_K))^{\frac{1}{2}}\right) \text{ et } \lim_{n \rightarrow +\infty} q_n \left(\frac{n}{\phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) = 0, \text{ où } q_n \text{ est un entier tel que } q_n(r_n + v_n) = O(n).$$

Nous avons donc établi le résultat suivant :

**Théorème 1.4.** [10] *Sous les hypothèses du Théorème 1.1 et si les hypothèses (F1)-F(5) sont vérifiées, on obtient :*

$$\sqrt{n\phi_x(h_K)}(\widehat{F}^x(y) - F^x(y) - B_n(x, y)) \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)),$$

où

$$V_{HK}(x, y) = \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)),$$

et

$$B_n(x, y) = \mathbb{E}(\widehat{F}^x(y)) - F^x(y) = B_H(x, y)h_H^2 + B_K(x, y)h_K^2 + o(h_K^2) + o(h_K^2),$$

où

$$\begin{aligned} B_H(x, y) &= \frac{1}{2} \frac{\partial^2 F^x(y)}{\partial y^2} \int t^2 H'(t) dt, \\ B_K(x, y) &= \frac{1}{2} G_0^{(2)}(0) \frac{N(1, 2)}{M_1}. \end{aligned}$$

la discussion sur les hypothèses et les détails de la démonstration du Théorème 1.4 seront donnés au Chapitre 3 de cette thèse.

Pour obtenir la normalité asymptotique de l'estimateurs (1.4) dans le cas dépendant, on garde les hypothèses (H9), H(10) et (F3(i)) et on remplace (F1), (F2) et (F4) par les hypothèses suivantes :

(F1') Pour tout  $l \in \{0, 2, 3\}$  la quantité  $\Lambda_l^{(2)}(0)$  existe.

(F2')  $\sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) \leq \psi_x(h_K)$ , où  $\psi_x(h_K)$  est telle que

$$\exists \epsilon \in ]0, \frac{a+1}{a-1}], \quad 0 < \psi_x(h_K) = O(\phi_x^{1+\epsilon}(h_K)).$$

(F4') Le paramètre de lissage  $h_K$  satisfait :

$$\exists C_1 > 0, \quad C_2 > 0, \quad C_2 n^{1-a} \leq \phi_x(h_K) \leq C_1 n^{\frac{1}{1-a}} \text{ avec } a > 1$$

$$\text{et } \lim_{n \rightarrow +\infty} \frac{\log(n)}{nh_H^3 \phi_x^2(h_K)} = 0.$$

On établit le résultat suivant :

**Théorème 1.5.** [11] *Sous les hypothèses (H9), H(10), (F2'), (F3(i)), (F3') et (F4') nous avons :*

$$\left( \frac{nh_H^3 \phi_x(h_K) (f^{x(2)}(\theta(x)))^2}{V_{HK}(x, \theta(x))} \right)^{1/2} (\hat{\theta}(x) - \theta(x)) \xrightarrow{d} \mathcal{N}(0, 1),$$

où

$$V_{HK}(x, \theta(x)) = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt.$$

La preuve de ce résultat et le détail des conditions imposées seront donnés dans le Chapitre 4.

# Bibliographie

- [1] Albert, W. M., Ingram, O. and Barry C, A. (2011). *Inequalities : Theory of Majorization and Its Applications.*, Springer Science+Business Media, LLC, 2nd edn.
- [2] Ardjoun, F.Z., Hennani, L.A. and Laksaci, A. (2016). A recursive kernel estimate of the functional modal regression under ergodic dependence condition. *J. Stat. Theory Pract.*, **10**, 475-496.
- [3] Attouch, M. Laksaci, A. and Ould Saïd, E. (2009). Asymptotic distribution of robust estimator for functional nonparametric models. *Comm. Statist. Theory and Methods.*, **38**, 1317-1335.
- [4] Baillo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102-111.
- [5] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *J. of Nonparametric Stat.*, **22**, 617-632.
- [6] Belarbi, F., Chemikh, S. and Laksaci, A. (2017). Local linear estimate of the nonparametric robust regression in functional data. *Statistics et Probability Letters.*, **134**, 128-133.
- [7] Berlinet, A., Elamine, A. and Mas, A. (2011). Local linear regression for functional data. *Inst. Statist. Math.*, **63**, 1047-1075.
- [8] Bienaymé, I.J. (1853). Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés. *C.R. Acad. Sci.*,

- 
- Paris* **37**, 309-324. Also as (1867). Liouville's *Journal de Mathématiques Pures et Appliquées*, **12**, 158-176.
- [9] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*, **46**, 199-233.
- [10] Bouanani, O., Rahmani, S., and Hennani, L.A (2020). Local linear conditional cumulative distribution function with mixing data. *Arabian Journal of Mathematics*, **9** :289-307.
- [11] Bouanani, O., Rahmani, S., Laksaci, A., and Rachdi, M. (2020). Asymptotic normality of conditional mode estimator for functional dependent data. *India journal of pure and applied mathematics*..**51**, (2), 465-481.
- [12] Bosq, D. (2000). *Linear processes in function spaces : theory and applications*.. New York : Springer.
- [13] Cardot, H., Ferraty, F. and Sarda, P. (2003). Spline estimators for the functional linear model. *Statist. Sinica.*, **13**, 571-591.
- [14] Chiou, J.-M., and Müller, H.-G. (2009). Modeling hazard rates as functional data for the analysis of cohort lifetables and mortality forecasting. *Journal of the American Statistical Association.*, **104**, 572-585.
- [15] Chouaf, A. and Laksaci, A. (2012). On the functional local linear estimate for spatial regression. *Stat. Risk. Model.*, **29**, 189-214.
- [16] Chung, KL. (2001). *A Course In Probability Theory* . Academic Press, 3rd edn.
- [17] Dabo-Niang, S. and Rhomari, N. (2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R. Acad. Sci. Paris*, **336**, 75-80.
- [18] Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique de mode conditionnel pour variable explicative fonctionnelle. *C. R., Math. Acad. Sci. Paris*, **344**, 49-52.

- [19] Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data. *Statistica.*, **70**, 83-94.
- [20] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics.*, **47**, 26-44.
- [21] Dabo-Niang, S., Rachdi, M. and Yao, A-F. (2011). Spatial kernel regression estimation and prediction for functional random fields. *Far. East. J. Statist.*, **37**, 77-113.
- [22] Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the local linear modelization of the conditional distribution for functional data. *Sankhya A.*, **76**, 328-355.
- [23] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2016). Estimation locale linéaire de la fonction de régression pour des variables hilbertiennes. *C.R. Math. Acad. Sci. Paris.*, **354**, 847-850.
- [24] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2017). Local linear regression modelization when all variables are curves. *Statist. Probab. Lett.*, **121**, 37-44.
- [25] Ezzahrioui, M. and Ould-Saïd, E. (2008a). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Stat.*, **20**, 3-18.
- [26] Ezzahrioui, M. and Ould-Saïd, E. (2008b). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East J. Theor. Stat.*, **25**, 15-38.
- [27] Ezzahrioui, M. and Ould-Saïd, E. (2008c). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Comm. Statist. Theory Methods.*, **37**, 2735-2759.
- [28] Fan, J. (1992). Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998-1004.

- [29] Fan, J. and Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *Ann. of Statist.*, **20**, 2008-2036.
- [30] Fan, J. (1993). Local linear regression smoothers and their minimax efficiency. *Ann. of Statist.*, **20**, 196-216.
- [31] Fan, J., Gijbels, I.(1996). *Local Polynomial Modelling and its Applications*. London, Chapman and Hall.
- [32] Ferraty, F. and Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Math. Acad. Sci. Paris.*, **330**, 403-406.
- [33] Ferraty, F. et Vieu, P. (2002) The functional nonparametric model and application to spectrometric data. *Comput. Statist.*, **17**, 545-564.
- [34] Ferraty, F., Goia, A. and Vieu, P. (2002). Functional nonparametric model for time series : a fractal approach for dimension reduction. *Test*, **11**, 317-344.
- [35] Ferraty, F. and Vieu, P. (2003). Curves discrimination a nonparametric functional approach. *Comput. Statist and Data Anal.*, **44**, 161-173.
- [36] Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. *J. Nonparametric. Statist .*, **16**, 111-127.
- [37] Ferraty, F., Laksaci, A. and Vieu, P. (2005). Functional time series prediction via conditional mode estimation. *C. R. Math. Acad. Sci. Paris .*, **340** 389-392.
- [38] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Theory and Practice. Springer Verlag.
- [39] Ferraty, F., Rabhi, A. et Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Roumaine Math. Pures Appl.*, **53**, 1-18.
- [40] 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**, 335-352.



- [41] Ferraty, F., Van Keilegom, I. and Vieu, P. (2010). On the validity of the bootstrap in non-parametric functional regression. *Scand. J. Stat.*, **37**, 286-306.
- [42] Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2011). Kernel regression with functional response. *J. Electron. Statist.*, **5**, 159-171.
- [43] Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2012). Estimation de la fonction de régression pour variable explicative et réponse fonctionnelles dépendantes. *C.R. Math. Acad. Sci. Paris.*, **350**, 717-720.
- [44] Ferraty, F. and Vieu, P. (2011). *Kernel regression estimation for functional data*. The Oxford handbook of functional data analysis, 72-129, Oxford Univ. Press, Oxford.
- [45] Ferraty, F., Van Keilegom, I. and Vieu, P. (2012). Regression when both response and predictor are functions. *J. Multivariate Anal.*, **109**, 10-28.
- [46] Foata, D. and Fuchs, A. (1996). *Calcul des probabilités*, Masson.
- [47] Frank, I. et Friedman, J. (1993). A statistical view of some chemometric regression tools. *Technometrics.*, **35**, 109-135.
- [48] Hastie, T. et Mallows, C. (1993). Discussion of a statistical view of some chemometrics regression tools. by frank, i.e. and friedman, j. h. *Technometrics.*, **35**, 140-143.
- [49] Helal, N. and Ould-Saïd, E. (2016). Kernel conditional quantile estimator under left truncation for functional regressors. *Opuscula Mathematica.*, **36**, 25-48.
- [50] Helland, I. S. (1990). Partial least squares regression and statistical models. *Scand.J. Statist.*, **17**, 97-114.
- [51] Hoerl, A. E. et Kennard, R. W. (1970). Ridge regression : Biased estimation for nonorthogonal problems. *Technometrics.*, **12**, 55-67.
- [52] Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.
- [53] Kara-Zaitri, L., Laksaci, A., Rachdi, M. and Vieu, P.(2016). Data-driven kNN estimation in nonparametric functional data analysis, *J. Multivariate Analysis.*, **153**, 176-188.

- 
- [54] Laksaci, A. et Maref, F. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *C. R. Math. Acad. Sci. Paris.*, **347**, 1075-1080.
- [55] Laksaci, A., Madani, F. and Rachdi, M. (2013). Kernel conditional density estimation when the regressor is valued in a semi metric space. *Commun. in Statit.-Theory and Meth.*, **42**, 3544-3570.
- [56] Laksaci, A., Rachdi, M. and Rahmani, S., (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1-23.
- [57] Laurini, M. P. (2014). Dynamic functional data analysis with non-parametric state space models. *Journal of Applied Statistics.*, **41**, 142-163.
- [58] Lejeune, M. (1985). Estimation non-parametrique par noyaux : regression polynomiale mobile. *Revue de Statist. Appliq.*, **33**, 43-68.
- [59] Lemdani, M. ., Ould-Saïd, E. ., Poulin, N. (2009). Asymptotic properties of a conditional quantile estimator with randomly truncated data. *J. Multivariate Anal.*, **100**, 546-559.
- [60] Leulmi, S. and Messaci, F. (2017). Local linear estimation of a generalized regression function with functional dependent data. *Communications in Statistics - Theory and Methods.*, **47**, 5795-5811.
- [61] Liebscher E. (2001). Central limit theorems for  $\alpha$ -mixing triangular arrays with applications to nonparametric statistics. *Math. Meth. of Statist.*, **10**, 194-214.
- [62] Masry, E. (1996). Multivariate local polynomial regression for time series : uniform strong consistency and rates. *J. Time Ser. Anal.*, **6**, 571-599.
- [63] Masry, E. and Mielniczuk, J. (1999). Local linear regression estimation for time series with long-range dependence. *Stochastic Process. Appl.*, **82**, 173-193.
- [64] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stochastic Process. Appl.*, **115**, 155-177.
- [65] Massy, W. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association.*, **60**, 234-256.

- [66] Messaci, F., Nemouchi, N., Ouassou, I. and Rachdi, M. (2015). Local polynomial modelling of the conditional quantile for functional data. *Stat. Methods. Appl.*, **24**, 597-622.
- [67] Müller, H.G. (1987). Weighted local regression and kernel methods for nonparametric curve fitting. *J. Amer. Statist. Assoc.*, **82**, 231-238.
- [68] Müller, H.-G. and Stadtmüller, U. (2005). Generalized functional linear models. *The Annals of Statistics.*, **33**, 774-805.
- [69] Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications.*, **9**, 141-142.
- [70] Nonnenmacher, D. J. F., Zagst, R. (1995). A new form of Jensen's inequality and its application to statistical experiments. *ANZIAM Journal.*, **36**, 389-398.
- [71] Quintela-del-Río, A. (2008). Hazard function given a functional variable : Non-parametric estimation under strong mixing conditions. *J. Nonparametr. Stat.*, **20**, 413-430.
- [72] Ramsay, J. O. et Silverman, B. W. (1997). *Functional data analysis*. Springer Series in Statistics. Springer, New York.
- [73] Rao, C. (1958). Some statistical methods for comparison of growth curves. *Biometrics.*, **14**, 1-17.
- [74] Ruppert, D. Wand, M.P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, **22**, 1346-1370.
- [75] Sauder, C., Cardot, H., Disenhaus, C., Le Cozler, Y. (2013). Non-parametric approaches to the impact of Holstein heifer growth from birth to insemination on their dairy performance at lactation one. *Journal of Agricultural Science.*, **151**, 578 - 589.
- [76] Sørensen, H., Tolver, A., Thomsen, M. H. et Andersen, P. H. (2012). Quantification of symmetry for functional data with application to equine lameness classification. *J. App. Statist.*, **39**, 317-360.
- [77] Stone, C.J. (1977). Consistent nonparametric regression. *Ann. of Statist.*, **5**, 595-620.

- [78] Thoma, M. (2004). Electrical energy usage over the business cycle. *Energy Economics*, **26**, 463-485.
- [79] Tibshirani, R. et T. Hastie. (1987). Local likelihood estimation. *J. Amer. Statist. Assoc.*, **82**, 59-567.
- [80] Tsybakov, A.B. (1986). Robust reconstruction of functions by the local approximation method. *Problems of Information Transmission.*, **22**, 133-146.
- [81] Tucker, L. (1958). Determination of parameters of a functional relation by factor analysis. *psychometrika.*, **23**, 19-23.
- [82] Van der Vaart.A.W. (2012). *Asymptotic Statistics*, Cambridge series in statistical and probabilistic mathematics.
- [83] Watson, G.S. (1964). Smooth regression analysis. *Sankhya A*, **26**, 359-372.
- [84] Xianzhu, X., Peiqin, L.and Chen, A. (2017). Asymptotic normality of the local linear estimation of the conditional density for functional time series data. *Communications in Statistics - Theory and Methods.*, **47**, 3418-3440.
- [85] Zhou, L, H. Wu, J. Li, Z. Wang, and L. Zhang. (2012) Determination of fatty acids in broiler breast meat by near-infrared reflectance spectroscopy. *Meat science.*, **90**, 658-664.
- [86] Zhou, Z. Y., Lin, Z. Y. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *Journal of Nonparametric Statistics.*, **28**, 1-16.

# Chapitre 2

## Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics

Ce chapitre fait l'objet d'une publication parue dans la revue internationale Behaviormetrika.



**Asymptotic normality of some conditional nonparametric functional  
parameters in high- dimensional statistics**

Oussama Bouanani <sup>1</sup>, Ali Laksaci <sup>2</sup>, Mustapha Rachdi <sup>3</sup> and Saâdia Rahmani <sup>1</sup>

<sup>1</sup> Laboratoire de Modèles Stochastiques, Statistique et Applications, Université Dr moulay Tahar, Saïda 20000, Algeria. E-mail : oussamaproba@gmail.com. E-mail : saadia.rahmani@gmail.com.

<sup>2</sup> Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia. E-mail : alilak@yahoo.fr.

<sup>3</sup> Université Grenoble Alpes (France), Laboratoire AGEIS, EA 7407, AGIM Team, UFR SHS, BP. 47, F38040 Grenoble Cedex 09, France. E-mails : mustapha.rachdi@univ-grenoble-alpes.fr

**Abstract :** This paper deals with the convergence in distribution of estimators of some conditional parameters in the Functional Data Analysis framework. In fact, we consider models where the input is of functional kind and the output is a scalar. Then, we establish the asymptotic normality of the nonparametric local linear estimators of (i) the conditional distribution function and (ii) the successive derivatives of the conditional density. Moreover, as by product we deduce the asymptotic normality of the local linear estimator of the conditional mode. Finally, in order to show interests of our results, on the practical point of view, we have conducted a computational study, first on a simulated data and, then on some real data concerning the forage quality.

**Key Words .** Functional Data Analysis (FDA) · Local linear estimation · Conditional cumulative distribution · Derivatives of the conditional density · Conditional mode · Asymptotic normality · Small ball probability · Forage quality.

AMS Classification : 62G05, 62G08, 62G20, 62G35, 62G07, 62G32, 62G30. Secondary : 62H12.

## 2.1 Introduction and motivations

The purpose of this paper is linked to FDA, and more precisely, to statistical models for which co-variables are of functional type (eventually of infinite dimension) and responses are scalars. It should be noticed that the scope of this topic is no longer justified given the large number of scientific publications dealing with all aspects related to the FDA. We therefore refer to Ramsay and Silverman [27], Ferraty and Vieu [17], Demongeot et al. [9]

and [10] and references therein, for a summary on theoretical and practical motivations on this subject.

Recall that the study of conditional models, in nonparametric FDA, is among the most important statistical analysis methods. Such investigation is essentially used to explore and to model the relationship between a scalar response variable and a functional regressor. Moreover, the consistency and the asymptotic normality of some statistical parameters estimators are two fundamental asymptotic properties.

Thus, we aim to extend studies on the local linear functional nonparametric regression estimation method. This latter is essentially based on an extension of the well known Nadaraya-Watson (N-W) kernel estimator to the FDA case. Recall that the nonparametric kernel estimator when functional data are available has been considered by many authors. Ferraty et al. [15] derived the mean squared convergence and the asymptotic distribution for the N-W estimator, while Ferraty et al. [18] established its uniform almost-complete convergence (a.co.<sup>1</sup>). Meanwhile, the nonparametric  $k$  Nearest-Neighbours ( $k$ NN) estimator, in FDA, has also been investigated (see, for instance, recent works by Kudraszow and Vieu [20], Kara-Zaïtri et al. [21, 22] and references therein). The case when both the response and the predictor variables are of functional kind is also investigated (see Ferraty et al., [19]).

It should be noticed that, to the best of our knowledge, first results on the local linear modeling in the FDA setup were established by Baïllo and Grané [1] and Berline et al. [4] to cite a few. These papers focus on the local linear estimation of the regression operator when the regressor takes values in a Hilbert space. The case where the explanatory variable takes values in a semi-metric space, has been considered in Barrientos et al. [2] and El methni and Rachdi [11]. In these two last works, authors obtained the almost-complete convergence with rates, of the proposed estimator. Then, we can cite Boj et al. [5] for an other alternative version of the functional local linear modeling.

In this paper, we are concerned with the asymptotic normality of local linear estimators of some statistical parameters i.e., the conditional cumulative distribution, conditional den-

---

1. Let  $(z_n)_{n \in \mathbb{N}}$  be a sequence of real random variables. We say that  $(z_n)$  converges almost-completely (a.co.) toward zero if, and only if, for all  $\epsilon > 0$ ,  $\sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon) < \infty$ . Moreover, we say that the rate of the almost-complete convergence of  $(z_n)$  to zero is of order  $u_n$  (with  $n \rightarrow 0$ ) and we write  $z_n = O_{a.co.}(u_n)$  if, and only if, there exists  $\epsilon > 0$  such that  $\sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon u_n) < \infty$ . This kind of convergence implies both the almost-sure convergence and the convergence in probability.

sity derivatives and the conditional mode. Recall that, consistencies, in mean square and a.co., of these estimators have been shown in Demongeot et al. [9] and [10] and Laksaci et al. [23], whereas the asymptotic normality of the N-W kernel estimators of the conditional mode was established and confidence bands were proposed (see Ezzahrioui and Ould-Saïd [12] for independent and identically distributed data (i.i.d.) and Louani and Ould-Saïd [24], Masry [25] and Delsol [8] for the strong mixing case). Notice also that Ezzahrioui and Ould-Saïd [13] and [14]), studied the asymptotic normality of the N-W kernel estimator of the conditional quantile for i.i.d. and strong mixing cases.

This paper is organized as follows. In section 2, we study the asymptotic normality of the conditional distribution function. In section 3, we focus on the conditional density and its successive derivatives. The asymptotic normality of the local linear estimator of the conditional mode is treated in section 4. Then, in order to show the practical interest of our results, we have conducted, in section 5, a computational study on (i) a simulated data and, (2) on a real data concerning the forage quality. Finally, the proofs of the main results are given in Section 6 and the appendix is reserved to some technical auxiliary results.

## 2.2 Conditional cumulative distribution estimation

Let  $(X_i, Y_i)$  be a sequence of i.i.d. random vectors where the random variable (r.v.)  $X_i$  belongs to a semi-metric space  $(\mathfrak{F}, d)$  and  $Y_i$  is a real-valued r.v. For a fixed  $x \in \mathfrak{F}$ , we denote the conditional probability distribution of  $Y_i$  given  $X_i = x$ , by

$$\text{for all } y \in \mathbb{R}, \quad F^x(y) = \mathbb{P}(Y_i \leq y | X_i = x).$$

The functional local linear estimator of  $F^x(y)$  is based on the minimization, with respect to  $(a, b)$ , of

$$\sum_{i=1}^n (H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K(h_K^{-1}\delta(x, X_i)), \quad (2.1)$$

where  $\beta(., .)$  and  $\delta(., .)$  are known functions from  $\mathfrak{F} \times \mathfrak{F}$  into  $\mathbb{R}$ ,  $K$  is a kernel,  $H$  is a distribution function and  $h_K := h_{K,n}$  and  $h_H = h_{H,n}$  are chosen as a sequences of positive real numbers. More precisely, the local linear estimator  $\hat{F}^x(y)$  of  $F^x(y)$  is  $\hat{a}$  which is the first component of the pair  $(a, b)$  solution of the minimization problem (2.1). However, if



the bi-functional operator  $\beta$  is such that, for all  $z \in \mathfrak{F}$ ,  $\beta(z, z) = 0$ , then  $\widehat{F}^x(y)$  is explicitly defined by

$$\widehat{F}^x(y) = \frac{\sum_{i,j=1}^n w_{ij} H(h_H^{-1}(y - Y_j))}{\sum_{i,j=1}^n w_{ij}} = \frac{\sum_{j=1}^n \Delta_j K_j H_j}{\sum_{j=1}^n \Delta_j K_j}, \quad (2.2)$$

where  $w_{ij} = \beta_i(\beta_i - \beta_j) K_i K_j$  and  $\Delta_j = K_j^{-1} \left( \sum_{i=1}^n w_{ij} \right) = \sum_{i=1}^n \beta_i^2 K_i - \left( \sum_{i=1}^n \beta_i K_i \right) \beta_j$ , with  $\beta_i = \beta(X_i, x)$ ,  $K_i = K(h_K^{-1}(\delta(x, X_i)))$  and  $H_j = H(h_H^{-1}(y - Y_j))$ .

### 2.2.1 Assumptions and notations

In what follows we denote by  $x$  (resp.  $y$ ) a fixed point in  $\mathfrak{F}$  (resp.  $\mathbb{R}$ ),  $\mathcal{N}_x$  (resp.  $\mathcal{N}_y$ ) a fixed neighborhood of  $x$  (resp. of  $y$ ) and  $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$ . Then, in order to establish our results the following assumptions will be needed.

(H1) For any  $r > 0$ ,  $\phi_x(r) := \phi_x(-r, r) > 0$ , and there exists a function  $\chi_x(\cdot)$  such

$$\text{that, for all } t \in [-1, 1], \lim_{h_K \rightarrow 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \chi_x(t).$$

(H2) The conditional distribution function  $F^x$  satisfies that there exist some positive constants  $b_1$  and  $b_2$ , such that : for all  $(x_1, x_2, y_1, y_2) \in \mathcal{N}_x \times \mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_y$  :

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C \left( |\delta(x_1, x_2)|^{b_1} + |y_1 - y_2|^{b_2} \right),$$

where  $C$  is a positive constant depending on  $x$ .

(H3) The functions  $\delta(\cdot, \cdot)$  and  $\beta(\cdot, \cdot)$  are such that :

(i) for all  $z \in \mathfrak{F}$ ,  $|\delta(x, z)| = d(x, z)$  and  $C_1 |\delta(x, z)| \leq |\beta(x, z)| \leq C_2 |\delta(x, z)|$  where  $C_1$  and  $C_2 > 0$ .

(ii)  $\sup_{u \in B(x, r)} |\beta(u, x) - \delta(x, u)| = o(r)$ , where  $B(x, r) = \{z \in \mathfrak{F} / |\delta(z, x)| \leq r\}$  denotes

the closed-ball centered at  $x$  and of radius  $r$ .

(H4) The kernel  $K$  is a positive function which is supported within  $[-1, 1]$ , differentiable and for which the first derivative  $K'$  satisfies :

$$K^2(1) - \int_{-1}^1 (K^2(u))' \chi_x(u) du > 0.$$

(H5) The kernel  $H$  is a differentiable function and its first derivative  $H'$  is symmetric and such that :  $\int |t|^{b_2} H'(t) dt < \infty$ .

(H6) The bandwidths are such that : there exists a positive integer  $n_0$  for which

$$-\frac{1}{\phi_x(h_K)} \int_{-1}^1 \phi_x(zh_K, h_K) \frac{d}{dz} (z^2 K(z)) dz > C_3 > 0 \text{ for } n > n_0.$$

$$\lim_{n \rightarrow \infty} h_K = \lim_{n \rightarrow \infty} h_H = 0 \text{ and } \lim_{n \rightarrow \infty} \frac{\log n}{n \phi_x(h_K)} = 0,$$

and

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

where  $dP(u)$  is the cumulative distribution of  $X$ .

### 2.2.2 Comments on the assumptions

The assumptions listed above are standard in the FDA context. In particular, assumptions (H1) and (H2) are adaptations of assumptions H1 and H3 in Ferraty et al. [15], when one replaces the semi-metric  $d$  by some bi-functional operator  $\delta$ . The assumption (H1) characterizes the concentration property of the probability measure of the functional variable  $X$ , which permits to control the effect of the topological structure in the asymptotic results (see Ferraty et al. [16]), while the assumption (H2) is a regularity condition which characterizes the functional space, of our model, and is needed to evaluate the bias term in the asymptotic results. Then, assumption (H3) has been introduced and commented, first, in Barrientos et al. [2] and it plays an important role in our methodology, particularly when we will compute exact constant terms involved in the asymptotic result. Moreover, the first and the third parts of the assumption (H6) are common in the setting of functional local linear fitting (see for instance Laksaci et al. [10]) and Rachdi et al. [26]). The rest of the hypotheses are imposed for a sake of brevity of our results's proofs. Moreover, one could find in Ferraty and Vieu [17] some examples of kernels  $K$  and  $H$  satisfying assumptions (H4) and (H5).

### 2.2.3 Main results

Before enunciating our main results, we introduce the quantities  $M_j$  and  $N(a, b)$ , which will appear in the bias and variance dominant terms :

$$M_j = K^j(1) - \int_{-1}^1 (K^j(u))' \chi_x(u) du \text{ where } j = 1, 2,$$

$$\text{and for all } a > 0 \text{ and } b = 2, 4, \quad N(a, b) = K^a(1) - \int_{-1}^1 (u^b K^a(u))' \chi_x(u) du.$$

Moreover, let

$$\widehat{F}^x(y) = \frac{\widehat{F}_N^x(y)}{\widehat{F}_D^x},$$

$$\text{where } \widehat{F}_N^x(y) = \frac{1}{n\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_j K_j H_j \text{ and } \widehat{F}_D^x = \frac{1}{n\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_j K_j.$$

Then, we have the following theorem.

**Theorem 1** *Under assumptions (H1)-(H6), we obtain*

$$\sqrt{n\phi_x(h_K)}(\widehat{F}^x(y) - F^x(y) - B_n(x, y)) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, y)),$$

where

$$V_{HK}(x, y) = \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)), \quad (2.3)$$

and

$$B_n(x, y) = \frac{\mathbb{E}(\widehat{F}_N^x(y))}{\mathbb{E}(\widehat{F}_D^x)} - F^x(y), \quad (2.4)$$

with  $\xrightarrow{D}$  denoting the convergence in distribution.

**Remark 2** *As mentioned in Demongeot et al. [9], the function  $\phi_x(t)$  can be empirically estimated by*

$$\widehat{\phi}_x(t) = \frac{\#\{i : |\delta(X_i, x)| \leq t\}}{n},$$

where  $\#(A)$  denotes the cardinality of the set  $A$ . So, if we take advantage of the following assumption,

$$(H\gamma) \quad \lim_{n \rightarrow \infty} \sqrt{n\phi_x(h_K)} B_n(x, y) = 0,$$

we can cancel the bias term and obtain the following corollary.

**Corollary 3** *When the assumptions (H1)-(H7) hold, we have the following asymptotic result*

$$\sqrt{\frac{n\hat{\phi}_x(h_K)}{V_{HK}(x,y)}} (\hat{F}^x(y) - F^x(y)) \xrightarrow{D} \mathcal{N}(0,1).$$

## 2.3 Conditional density and its derivatives

In this section, similar results, to Theorem 1, will be derived for the local linear estimator of the  $j^{th}$  order derivative,  $f^{x(j)}(y)$ , of the conditional density  $f^x(y)$  of the real r.v.  $Y = y$  given the r.v.  $X = x \in \mathfrak{F}$ .

The local linear estimator of  $f^{x(j)}(y)$  is given by  $\hat{f}^{x(j)}(y) = \hat{a}$  corresponding to the first component of the solution vector of the minimization problem :

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n \left( h_H^{-j-1} H^{(j+1)}(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1}\delta(x, X_i)).$$

This estimator is explicitly given by :

$$\hat{f}^{x(j)}(y) = \frac{\sum_{i \neq k}^n W_{ik}(x) H^{(j+1)}(h_H^{-1}(y - Y_k))}{h_H^{(j+1)} \sum_{i \neq k}^n W_{ik}(x)} = \frac{\sum_{k=1}^n \Delta_k K_k H_k^{(j+1)}}{h_H^{(j+1)} \sum_{k=1}^n \Delta_k K_k}.$$

Notice that this estimator was already introduced by Demongeot et al. [9] who established its a.co. (pointwise and uniform) consistency when data are i.i.d. Then, a spatial version of this estimator was studied by Laksaci et al. [10].

### 2.3.1 Assumptions and main result

In order to establish the asymptotic convergence of  $\hat{f}^{x(j)}(y)$  we need some notations and assumptions.

Firstly we denote, for  $l \in \{j, j+2\}$ , by

$$\varphi_l(\cdot, y) = \frac{\partial^l f(\cdot, y)}{\partial y^l} \text{ and } \psi_l(s) = \mathbb{E}[\varphi_l(X, y) - \varphi_l(x, y) | \beta(X, x) = s].$$

Secondly, in addition to the assumptions introduced along the previous section, we need the following conditions :

- (H1') The second derivative at 0, for  $l \in \{j, j+2\}$ ,  $\psi_l^{(2)}(0)$  exists.  
(H6') Assumption (H6) is changed by replacing the last part by

$$\lim_{n \rightarrow +\infty} \frac{\log n}{nh_H^{2j+1} \phi_x(h_K)} = 0.$$

- (H7) (i) There exists a positive constant  $C$  such that, for all  $(y_1, y_2) \in \mathbb{R}^2$ ,

$$|H^{(j+1)}(y_1) - H^{(j+1)}(y_2)| \leq C|y_1 - y_2|$$

- (ii) There exists a positive constant  $\kappa$  such that, for all  $j' \leq j$ ,

$$\lim_{y \rightarrow \infty} |y|^{1+\kappa} |H^{(j'+1)}(y)| = 0.$$

- (iii)  $H^{(j+1)}$  is a bounded function, and  $\int (H^{(1)}(t))^\nu dt < \infty$  where  $\nu > 1$ .

**Remark 4** Remark that, assumption (H1') enables to get exact expressions of the constants in the asymptotic results, and assumption (H7) is a technical condition (see Ezzahrioui and Ould-Saïd [16] for more comments).

Similarly to the conditional cumulative function estimator, let

$$\hat{f}^{x(j)}(y) = \frac{\hat{f}_N^{x(j)}(y)}{\hat{F}_D^x} \text{ where } \hat{f}_N^{x(j)}(y) = \frac{1}{nh_H^{(j+1)} \mathbb{E}(\Delta_1 K_1)} \sum_{i=1}^n \Delta_i K_i H_i^{(j+1)}.$$

Then, we have the following asymptotic result.

**Theorem 5** Under assumptions (H1'), (H2'), (H3), (H4), (H5), (H6') and (H7), we have

$$\sqrt{nh_H^{2j+1} \phi_x(h_k)} (\hat{f}^{x(j)}(y) - f^{x(j)}(y) - B_n(x, y)) \xrightarrow{D} \mathcal{N}(0, V_{H,K}(x, y)),$$

where

$$V_{H,K}(x, y) = \frac{M_2}{M_1^2} f^x(y) \int (H^{(j+1)}(t))^2 dt,$$

and

$$B_n(x, y) = \frac{\mathbb{E}(\hat{f}_N^{x(j)}(y))}{\mathbb{E}(\hat{F}_D^x)} - f^{x(j)}(y) \text{ with } \mathbb{E}(\hat{F}_D^x) = 1.$$

**Remark 6** *If we impose the following additional assumption*

$$(H8) \quad \lim_{n \rightarrow \infty} \sqrt{nh_H^{2j+1} \phi_x(h_K)} B_n(x, y) = 0,$$

*we can cancel the bias term and obtain the following corollary.*

**Corollary 7** *Assume that hypotheses of Theorem 5 are satisfied. If assumption (H8) holds, then we obtain*

$$\sqrt{\frac{nh_H^{2j+1} \widehat{\phi}_x(h_K)}{V_{H,K}(x, y)}} (\widehat{f}^{x(j)}(y) - f^{x(j)}(y)) \xrightarrow{D} \mathcal{N}(0, 1),$$

where  $\widehat{\phi}_x(h_K)$  denotes the empirical counterpart of  $\phi_x(h_K)$ .

## 2.4 Conditional mode estimation

We focus on the local linear estimation of the conditional mode. For this aim, we assume that  $f^x(\cdot)$  is unimodal. Then, its conditional mode, that we denote by  $\theta(x)$ , is defined by

$$f^x(\theta(x)) = \sup_{y \in S} f^x(y),$$

where  $S$  is a ‘‘compact’’ set such that the mode  $\theta(x)$  of  $f^x$  is assumed to be uniquely defined within.

A local linear estimator  $\widehat{\theta}(x)$ , of  $\theta(x)$ , is defined by

$$\widehat{f}^x(\widehat{\theta}(x)) = \sup_{y \in S} \widehat{f}^x(y)$$

In order to establish the asymptotic normality of  $\widehat{\theta}(x)$ , with a suitable normalization, we need to introduce the following additional smoothness conditions.

(H9) There exists some integer  $j > 0$ , such that,  $f^x$  is  $j$ -times continuously differentiable with respect to  $y$  on  $(\theta - \xi, \theta + \xi)$  for  $\xi > 0$ .

$$(H10) \quad \begin{cases} f^{x(l)}(\theta) = 0 & \text{if } 1 \leq l < j, \\ |f^{x(j)}(\theta)| > 0. \end{cases}$$

(H11) The bandwidth  $h_H$  satisfies  $\lim_{n \rightarrow +\infty} nh_H^3 \phi_x(h_K) = 0$ .

**Theorem 8** *Under the hypotheses of Theorem 5, if assumptions (H9) and (H10) hold, we obtain*

$$\left( \frac{nh_H^3 \phi_x(h_K) (f^{x(2)}(\theta(x)))^2}{V_{H,K}(x, \theta(x))} \right)^{1/2} \left( \widehat{\theta}(x) - \theta(x) - B_n(x, \theta(x)) \right) \xrightarrow{D} \mathcal{N}(0, 1),$$

where

$$B(x, \theta(x)) = B_H(x, \theta(x))h_H^2 + B_K(x, \theta(x))h_K^2 + o(h_H^2) + o(h_K^2),$$

with

$$B_H(x, \theta(x)) = \frac{1}{2} f^{x(3)}(\theta(x)) \int t^2 H^{(1)}(t) dt,$$

$$B_K(x, \theta(x)) = \frac{1}{2} \psi_1^{(2)}(0) \frac{N(1, 2)}{M_1},$$

and

$$V_{H,K}(x, \theta(x)) = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt.$$

**Remark 9** *Under additional assumptions on the function  $\phi_x(\cdot)$  and the bandwidths  $h_K$  and  $h_H$ , we can get rid of the bias term  $B_n(x, y)$ .*

**Corollary 10** *Under the hypotheses of Theorem 8, if the bandwidths  $h_K$  and  $h_H$  and the function  $\phi_x(\cdot)$  satisfy  $\lim_{n \rightarrow \infty} (h_K + h_H) \sqrt{nh_H^3 \phi_x(h_K)} = 0$ , we obtain*

$$\left( \frac{nh_H^3 \phi_x(h_K) (f^{x(2)}(\theta(x)))^2}{V_{H,K}(x, \theta(x))} \right)^{1/2} \left( \widehat{\theta}(x) - \theta(x) \right) \xrightarrow{D} \mathcal{N}(0, 1).$$

## 2.5 Computational studies

We aim to evaluate, on a finite sample, performances of the asymptotic normality of both functional models (the conditional distribution and the conditional density functions) on simulated data and on a real data-set. In particular, our main purpose is to show how we can implement easily and fastly these estimators in practice. Of course, the applicability of our asymptotic normality result requires a practical estimation of the asymptotic bias and variance. For a sake of shortness, we neglect the bias term and we use a plug-in approach to construct an estimator of the asymptotic variance. For the conditional distribution function :

$$\widehat{V}_{HK}(x, y) = \frac{n \widehat{\phi}_x(h_K) \sum_{i=1}^n K_i^2 \widehat{F}^x(y) (1 - \widehat{F}^x(y))}{\left( \sum_{i=1}^n K_i \right)^2},$$

and, for the conditional density function :

$$\widehat{V}_{H,K}(x, y) = \frac{n\widehat{\phi}_x(h_K) \sum_{i=1}^n K_i^2 \widehat{f}^x(y)}{(\sum_{i=1}^n K_i)^2} \int (H^{(1)}(t))^2 dt.$$

### 2.5.1 On simulated data

In order to test the effectiveness of the two asymptotic normality results, we generate the data by the model  $Y = r(x) + \varepsilon$ , where  $\varepsilon \sim \mathcal{N}(0, 0.5)$ . The explanatory functional variables are constructed by :

$$X_i(t) = 2 * \cos(t * W_{1_i}) / (2 + \cos(t * W_{2_i})) + 0.2 * W_{3_i}, \text{ for } i = 1, \dots, n = 100,$$

where  $W_1$  (resp.  $W_2$ ) is a uniformly distributed random variable on  $[0, \pi/2]$  (rep.  $[\pi/2, \pi]$ ) and  $W_3$  is generated from a gaussian distribution  $\mathcal{N}(0, 0.2)$ . All the curves  $X_i$ 's are generated from 100 equispaced measurements in  $(0, 1)$ , and are plotted in Fig. 2.1.

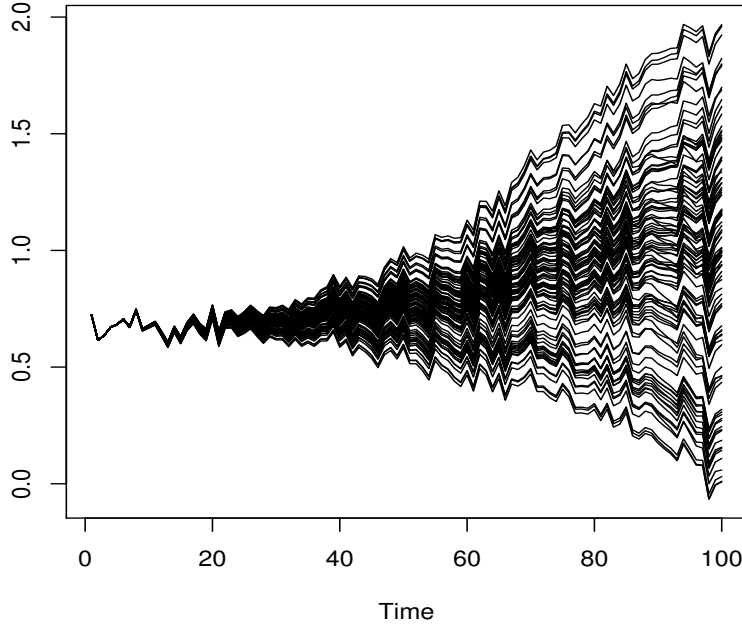


FIGURE 2.1 – A sample of 100 curves

On the other hand, for  $i = 1, \dots, n = 100$ , the scalar response  $Y_i$  is computed by considering the following operator :

$$r(x) = \int_0^1 \frac{dt}{1 + |x(t)|}.$$



Recall that, the conditional distribution of  $Y$  given  $X = x$  corresponding to this model is explicitly given by the law of  $\varepsilon_i$  shifted by  $\int_0^1 \frac{dt}{1 + |x(t)|}$ . Then, the corresponding conditional density  $f^x(y)$  is

$$f^x(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - r(x))^2\right),$$

and the conditional mode is

$$\theta(x) = r(x) = \int_0^1 \frac{dt}{1 + |x(t)|}.$$

Elsewhere, as it is well known in FDA, the choice of the metric and the smoothing parameters have crucial roles in the computational issues. To optimize these choices in this illustration, we use (i) the local cross-validation procedure method on the  $k$ NN for choosing smoothing parameters, (ii) regarding the shape of the curves  $X_i$ , it is clear that the PCA-type semi-metric (see Benhenni et al., [3]), is well adapted to this kind of data. Then, we point out that, we opted for a quadratic kernel which is supported within  $(0, 1)$  and taken  $K = H'$ . To sum things up, we have followed the following steps for the three models :

- Step 1 : We generate  $m$  independent replications of  $(X_i, Y_i)_{i=1, \dots, n}$
- Step 2 : We compute the mean curve  $\mathcal{X}$  of  $(X_i)_{i=1, \dots, n \times m}$  by the routine *func.mean* of the R-package *fda.usc*.
- Step 3 : We compute the asymptotic variances  $\widehat{V}_{H,K}(\mathcal{X}, 0)$  and  $\widehat{V}_{HK}(\mathcal{X}, 0)$  for both models
- Step 4 : We compute the quantities

$$\sqrt{\frac{n\phi_x(h_K)}{\widehat{V}_{HK}(x, y)}}(\widehat{F}^x(y) - F^x(y)) \quad \text{and} \quad \sqrt{\frac{nh_H\phi_x(h_K)}{\widehat{V}_{H,K}(x, y)}}(\widehat{f}^x(y) - f^x(y))$$

- Step 5 : We estimate the densities for both  $m$ -samples by the kernel method
- Step 6 : We compare these densities to the standard Gaussian distribution

The obtained results are shown in Fig. 2.2.

It appears clearly that, asymptotic distributions have good behaviors with respect to the standard normal distribution. This conclusion is confirmed by the Kolmogorov-Smirnov test which, for  $n = m = 100$ , gives 0.84 as a P-value for the first model and 0.77 for the second one.

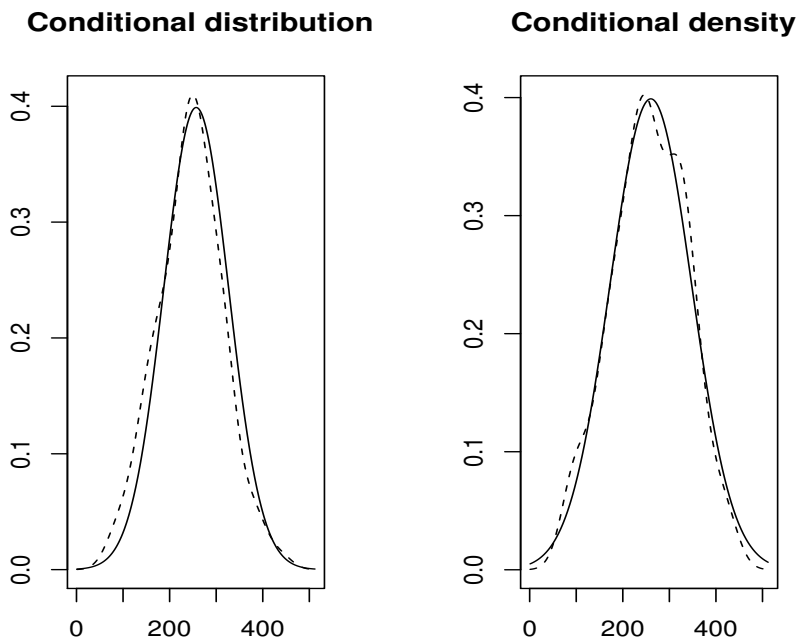


FIGURE 2.2 – The asymptotic distribution : conditional distribution (left plot) and the conditional density (right plot)

## 2.5.2 Real data application

We apply our approach to the forage quality data. In practice, the analysis of this kind of data is very important for the food industry. In fact, it intervenes in many food products such as milk quality, dairy products, meat quality, . . . . Notice that, there are three fundamental parameters in the study of the forage quality that are (i) the concentration of crude protein (CP), (ii) the acid detergent fiber (ADF) and (iii) the neutral detergent fiber (NDF). The classical analytical procedures are time-consuming and very expensive. That is why we propose a new approach based on the functional local linear approach. More Precisely, we use the conditional mode to predict the level of the concentration of crude protein (CP) given the spectrometric curves of the sheepgrass hay. Moreover, we use the conditional distribution function and the conditional density to construct a predictive region of this parameter. Notice that data are collected from sheepgrass fields of the hay factories in Heilongjiang Province of Northeast China. We refer to Chen et al. (2015), for a complete description of this data-set. We consider a sample of size 150 and the near-infrared spectra were recorded at 5 *nm* intervals from 950 to 1650 *nm*. These functional curves are plotted in Fig. 2.3.

In order to highlight our models, we consider two prediction problems (i) the single-point prediction by the conditional mode and (ii) the predictive region. For the first purpose,

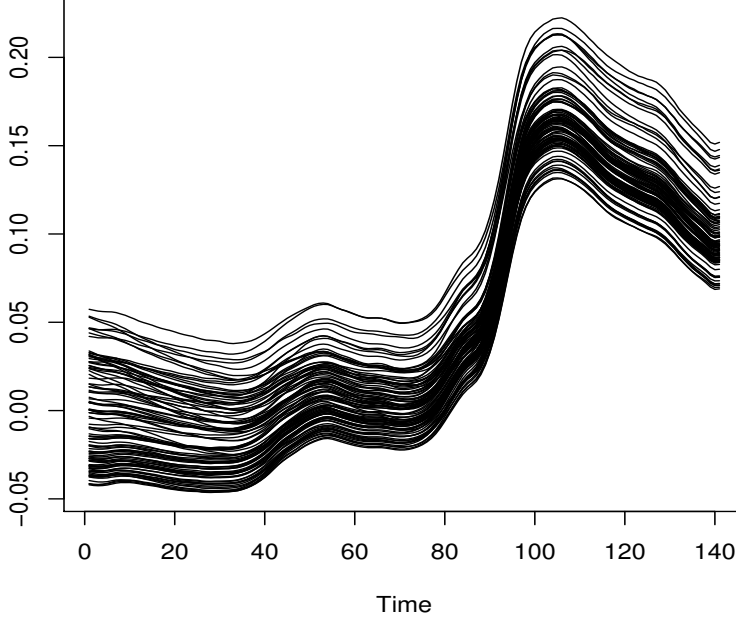


FIGURE 2.3 – Near-infrared spectra curves

we compare the local linear approach method to the kernel method in the prediction by the conditional mode function, while for the second method we propose two predictive regions. The first one results from the asymptotic normality of the conditional mode which is expressed by :

$$[a_1, b_1] \quad \text{with} \quad \begin{cases} a_1 = \widehat{\theta}_x - t_{1-\zeta/2} \times \left( \frac{\widehat{V}_{H,K}(x, \widehat{\theta}(x))}{nh_H^3 \phi_x(h_K) \widehat{f}^{x(2)}(\widehat{\theta}(x))^2} \right)^{1/2} \\ \text{and} \\ b_1 = \widehat{\theta}_x + t_{1-\zeta/2} \times \left( \frac{\widehat{V}_{H,K}(x, \widehat{\theta}(x))}{nh_H^3 \phi_x(h_K) \widehat{f}^{x(2)}(\widehat{\theta}(x))^2} \right)^{1/2} \end{cases}$$

where  $t_{1-\zeta/2}$  denotes the  $1-\zeta/2$  quantile of the standard normal distribution. The shortest conditional modal interval is :

$$[a_2, b_2] = \arg \min \left\{ \text{Leb}([c, d]) \text{ such that } \widehat{F}^x(d) - \widehat{F}^x(c) \geq \alpha \right\},$$

where  $\text{Leb}(C)$  denotes the Lebesgue measure of the set  $C$ .

For a practical purpose, we randomly split our data into two subsets. The first sample, of size  $n = 130$ , will be used to calculate our estimators on the 20 remaining curves.

The estimators are computed by using the  $B$ -spline semi-metric of the first derivative and the bandwidths parameters are selected by the local cross-validation technique on the number of nearest neighbors. On the other hand, we consider the quadratic kernel  $K(x) = 0.75(1 - x^2)$  if  $x \in [-1, 1]$ , which is supported within  $[-1, 1]$ , and the function  $H$  is chosen as the primitive of the kernel  $K$ . Then, the single-point prediction results are presented in Fig. 2.4.

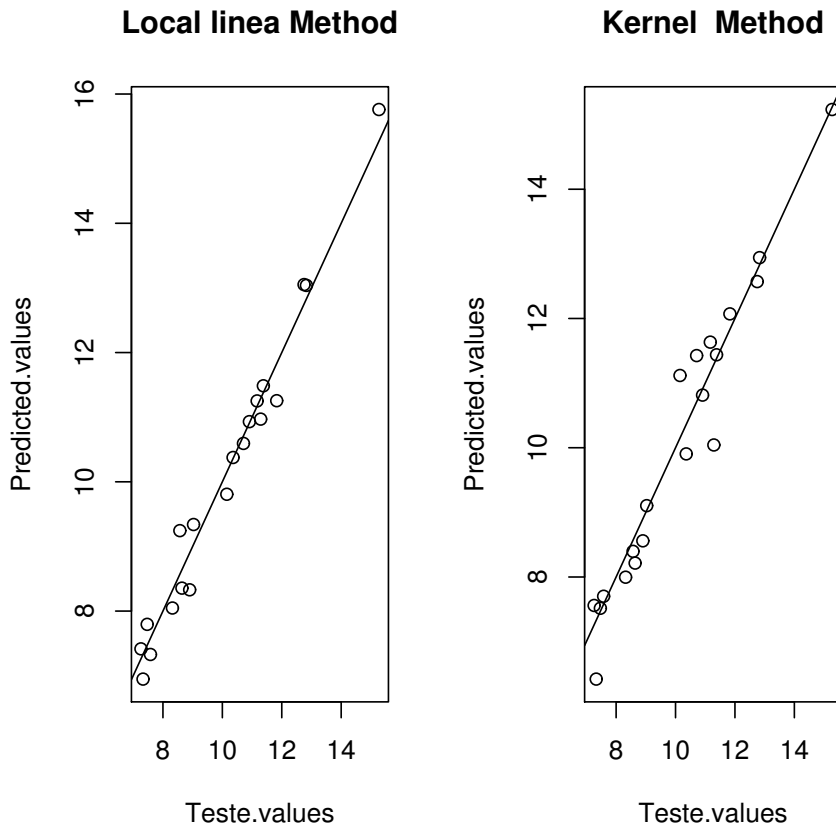


FIGURE 2.4 – Predicted values ( $y$ -axis) versus Test values ( $x$ -axis) results : local linear method (left plot) and the kernel smoothing approach (right plot)

The comparison of both scatterplots indicates that the local linear approach (on the left plot) gives better prediction results than the classical kernel method (on the right plot). This is confirmed by the mean squared error which is equal 0.27 versus 0.46. On the other hand, we give in Table 2.1 the 90% predictive intervals of the concentrations of crude protein of the 10 values in the sample test.

The true value	$[a_1 - b_1]$	$[a_2 - b_2]$
7.26	[6.68 8.90]	[7.42 8.56]
7.33	[6.78 8.76]	[6.91 8.23]
7.58	[6.77 8.40]	[7.02 8.36]
7.47	[7.58 8.19]	[7.36 8.55]
15.27	[12.29 18.05]	[13.98 17.34]
12.75	[10.37 12.98]	[11.02 12.56]
10.15	[8.58 11.75]	[9.71 11.16]
10.71	[9.038 12.76]	[10.76 11.03]
11.38	[11.58 14.02]	[10.85 13.82]
12.83	[10.43 13.54]	[11.99 13.51]
11.83	[11.18 13.39]	[12.01 13.37]

TABLE 2.1 – The 90% predictive intervals

It appears clear that both predictive intervals are equivalent. This statement is confirmed by the coverage probability (C.P.) where, we have 10% for the first one versus 11% for the second predictive interval. This conclusion shows the good performance of our asymptotic normality because, it is well known that, the shortest conditional modal interval is of the smallest width among all predictive intervals.

## 2.6 Proofs

In what follows, when there is no confusion, we denote by  $C$  or/and  $C'$  any generic positive constants. For a sake of shortness, we give full details for the proof concerning Theorem 1. Proofs of the other cases follow the same lines and are therefore presented in a much more synthetic way, but they may be obtained on request.

**Proof of Theorem 1.** Remark that

$$\widehat{F}^x(y) - F^x(y) - B_n(x, y) = \frac{\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - \widehat{F}_D^x B_n(x, y)}{\widehat{F}_D^x}.$$

If we denote by

$$\begin{aligned} Q_n(x, y) &= \widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - \mathbb{E}\left(\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x\right) \\ &= \widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - B_n(x, y), \end{aligned} \quad (2.5)$$

since

$$\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x = Q_n(x, y) + B_n(x, y),$$

then, the proof of this theorem will be complete from the following expression

$$\widehat{F}^x(y) - F^x(y) - B_n(x, y) = \frac{Q_n(x, y) - B_n(x, y)(\widehat{F}_D^x - \mathbb{E}(\widehat{F}_D^x))}{\widehat{F}_D^x}, \quad (2.6)$$

and the following auxiliary results which play a main role and for which proofs are given in the appendix.

**Lemma 2.6.1** (see [2] and [37]).

Under assumptions (H1)-(H6), we have

- (a)  $\mathbb{E}(K_1^j) = M_j \phi_x(h_K)$  for  $j = 1, 2$ .
- (b)  $\mathbb{E}(K_1^a \beta_1) = o(h_K \phi_x(h_K))$  for all  $a > 0$ .
- (c)  $\mathbb{E}(K_1^a \beta_1^b) = N(a, b) h_K^b \phi_x(h_K) + o(h_K^b \phi_x(h_K))$  for all  $a > 0$  and  $b = 2, 4$ .
- (d)  $\frac{\mathbb{E}(K_1 \Delta_1)}{n-1} = \mathbb{E}(w_{12}) = N(1, 2) M_1 h_K^2 \phi_x^2(h_K) + o(h_K \phi_x(h_K))$ .
- (e) For all  $(k, l) \in \mathbb{N}^* \times \mathbb{N}$ ,  $\mathbb{E}\left(K_1^k |\beta_1|^l\right) \leq C h_K^l \phi_x(h_K)$ .
- (f)  $\mathbb{E}\left(K_1 \beta_1^2\right) \leq C h_K^2 \phi_x(h_K)$ .

**Lemma 2.6.2** Under assumptions (H1)-(H5), we have

$$\widehat{F}_D^x \xrightarrow{P} \mathbb{E}(\widehat{F}_D^x) = 1,$$

where  $\xrightarrow{P}$  denotes the convergence in probability.

**Lemma 2.6.3** Under assumptions (H2), (H4) and (H5), as  $n \rightarrow \infty$ , we have

$$\mathbb{E}\left(K_1^2 \text{Var}\left(H\left(\frac{y - Y_1}{h}\right) \mid X_1\right)\right) \rightarrow \mathbb{E}(K_1^2)F^x(y)(1 - F^x(y)).$$

**Lemma 2.6.4** (see [10]). Under assumptions (H1)-(H5), we have

$$I \quad \mathbb{E}(\widehat{F}_N^x(y)) - F^x(y) = B_H(x, y)h_H^2 + B_K(x, y)h_K^2 + o(h_K^2) + o(h_K^2),$$

$$II \quad \text{Var}(\widehat{F}^x(y)) = \frac{V_{HK}(x, y)}{n\phi_x(h_K)} + o\left(\frac{1}{n\phi_x(h_K)}\right),$$

$$III \quad \text{Cov}(\widehat{F}_N^x(y), \widehat{F}_D^x(y)) = o\left(\frac{1}{n\phi_x(h_K)}\right),$$

$$IV \quad \text{Var}(\widehat{F}_D^x(y)) = o\left(\frac{1}{n\phi_x(h_K)}\right),$$

where

$$B_H(x, y) = \frac{1}{2} \frac{\partial^2 F^x(y)}{\partial y^2} \int t^2 H'(t) dt, \quad B_K(x, y) = \frac{1}{2} \Psi_0^{(2)}(0) \frac{N(1, 2)}{M_1},$$

and

$$V_{HK}(x, y) = F^x(y)(1 - F^x(y)) \frac{M_2}{M_1^2}.$$

So, Lemma 2.6.2, implies that  $\widehat{F}_D^x \rightarrow 1$ . Moreover,  $B_n(x, y) = o(1)$  as  $n \rightarrow \infty$  because of the continuity of  $F^x$ . Then, we obtain that

$$\widehat{F}^x(y) - F^x(y) - B_n(x, y) = \frac{Q_n(x, y)}{\widehat{F}_D^x} (1 + o_p(1)).$$

So, it suffices to show that

$$\sqrt{n\phi_x(h_K)} Q_n(x, y) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, y)), \quad (2.7)$$

where  $V_{HK}(x, y)$  is defined by (2.3). For this, notice that on one side, we have

$$\begin{aligned} & \sqrt{n\phi_x(h_K)}Q_n(x, y) \\ &= \frac{\sqrt{n\phi_x(h_K)}}{n\mathbb{E}(\Delta_1 K_1)} \left( \sum_{j=1}^n \Delta_j K_j (H_j - F^x(y)) - \mathbb{E} \left( \sum_{j=1}^n \Delta_j K_j (H_j - F^x(y)) \right) \right), \end{aligned}$$

which, combined with (2.5) implies that

$$\begin{aligned} & \sqrt{n\phi_x(h_K)}Q_n(x, y) \\ &= \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)) \\ & \quad - \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)) \\ & \quad - \mathbb{E} \left( \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)) \right) \\ & \quad + \mathbb{E} \left( \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)) \right). \end{aligned}$$

Denote by

$$\begin{aligned} T_1 &= \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i \quad , \quad T_2 = \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)) \\ T_3 &= \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i \quad \text{and} \quad T_4 = \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)). \end{aligned}$$

Then

$$\begin{aligned} \sqrt{n\phi_x(h_K)}Q_n(x, y) &= (T_1 T_2 - T_3 T_4 - \mathbb{E}(T_1 T_2 - T_3 T_4)) \\ &= (T_1 T_2 - \mathbb{E}(T_1 T_2)) - (T_3 T_4 - \mathbb{E}(T_3 T_4)). \end{aligned}$$

Hence, by the Slutsky's theorem, to show (2.5), it suffices to prove the following two claims :

$$T_1 T_2 - \mathbb{E}(T_1 T_2) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, y)) \quad (2.8)$$

$$T_3 T_4 - \mathbb{E}(T_3 T_4) \xrightarrow{P} 0, \quad (2.9)$$



*Proof of (2.8).* We can write that

$$T_1 T_2 - \mathbb{E}(T_1 T_2) = (T_2 - \mathbb{E}(T_2)) + ((T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)).$$

Again by the Slutsky's Theorem, (2.8) can be deduced from the too following intermediate results,

$$(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2) \xrightarrow{P} 0, \quad (2.10)$$

and

$$T_2 - E(T_2) \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, y)). \quad (2.11)$$

Concerning the proof of (2.10), by applying the Bienaymé-Tchebychev's inequality, we obtain for all  $\varepsilon > 0$

$$\mathbb{P}(|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)| > \varepsilon) \leq \frac{\mathbb{E}(|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)|)}{\varepsilon}.$$

Then, the Cauchy-Schwarz inequality implies that

$$\mathbb{E}|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)| \leq 2\sqrt{\mathbb{E}((T_1 - 1)^2)}\sqrt{\mathbb{E}(T_2^2)}.$$

On one side, by using Lemma 2.6.1's result, we obtain

$$\begin{aligned} \mathbb{E}((T_1 - 1)^2) &= \text{Var}(T_1) = \frac{1}{n^2 \mathbb{E}^2(\beta_1^2 K_1)} n \text{Var}(\beta_1^2 K_1) \\ &\leq \frac{1}{n(O(h^4 \phi_x^2(h_K)))} \mathbb{E}(\beta_1^4 K_1^2) = O\left(\frac{1}{n\phi_x(h_K)}\right), \end{aligned}$$

and on the other side, we obtain

$$\begin{aligned} \mathbb{E}((T_2)^2) &= \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E}\left(\sum_{j=1}^n K_j(H_j - F^x(y))\right)^2 \\ &= \frac{n}{(n-1)^2 O(\phi_x(h_K))} (nO(\phi_x(h_K)) + n(n-1)o(\phi_x^2(h_K))) \\ &= O(1) + o(n\phi_x(h_K)). \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E}|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T)| &\leq 2\sqrt{\mathbb{E}((T_1 - 1)^2)}\sqrt{\mathbb{E}(T_2^2)} \\ &\leq 2\sqrt{O\left(\frac{1}{n\phi_x(h_K)}\right)}(O(1) + o(n\phi_x(h_K))) = o(1), \end{aligned}$$

which implies that  $(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2) = o_p(1)$ . Then, as  $n \rightarrow \infty$ , we get

$$\mathbb{P}(|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)| > \varepsilon) \leq \frac{\mathbb{E}(|(T_1 - 1)T_2 - \mathbb{E}((T_1 - 1)T_2)|)}{\varepsilon} \rightarrow 0.$$

Therefore, to prove (2.8), we just need to prove (2.11). For that we denote

$$\begin{aligned} R_n &= T_2 - \mathbb{E}(T_2) \\ &= \frac{\sqrt{n\phi_x(h_K)\mathbb{E}(\beta_1^2 K_1)}}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j(H_j - F^x(y)) - \mathbb{E}(K_j(H_j - F^x(y))) \\ &= \sum_{j=1}^n \varepsilon_{nj}(x, y), \end{aligned}$$

where

$$\varepsilon_{nj} = \frac{\sqrt{n\phi_x(h_K)\mathbb{E}(\beta_1^2 K_1)}}{\mathbb{E}(\Delta_1 K_1)} K_j(H_j - F^x(y)) - \mathbb{E}(K_j(H_j - F^x(y))).$$

By the fact that  $\varepsilon_{nj}(x, y)$  are i.i.d., it follows that

$$\begin{aligned} \text{Var}(R_n(x, y)) &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \text{Var}(\varepsilon_{n1}(x, y)) \\ &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E}(\varepsilon_{n1}^2(x, y)). \end{aligned}$$

Thus

$$\begin{aligned} \text{Var}(R_n(x, y)) & \tag{2.12} \\ &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} (\mathbb{E}((H_1 - F^x(y))^2 K_1^2) - (\mathbb{E}((H_1 - F^x(y))K_1))^2). \end{aligned}$$

Concerning the second term on the right hand side of (2.12), we have

$$\begin{aligned} (\mathbb{E}((H_1 - F^x(y))K_1))^2 &= (\mathbb{E}(\mathbb{E}(H_1 - F^x(y))K_1|X_1))^2 \\ &= (\mathbb{E}(K_1 \mathbb{E}((H_1|X_1) - F^x(y))))^2, \end{aligned}$$

where

$$\mathbb{E}((H_1|X_1) - F^x(y)) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{2.13}$$

Now, let us return to the first term of the right hand side of (2.12). We have

$$\begin{aligned}
 & \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \left( \mathbb{E} \left( (H_1 - F^x(y))^2 K_1^2 \right) \right) \\
 &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} \left( \mathbb{E} \left( \mathbb{E} \left( (H_1 - F^x(y))^2 | X_1 \right) K_1^2 \right) \right) \\
 &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E} \left( \text{Var}(H_1 | X_1) K_1^2 \right) \\
 & \quad + \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \left( \mathbb{E} \left( \mathbb{E} \left( (H_1 | X_1) - F^x(y) \right)^2 K_1^2 \right) \right).
 \end{aligned}$$

By using (2.13), that allows to have, as  $n \rightarrow \infty$

$$\frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \left( \mathbb{E} \left( \mathbb{E} \left( (H_1 | X_1) - F^x(y) \right)^2 K_1^2 \right) \right) \rightarrow 0.$$

Combining Lemma 2.6.1 with Lemma 2.6.2's results, we obtain, as  $n \rightarrow \infty$

$$\mathbb{E} \left( \text{Var}(H_1 | X_1) K_1^2 \right) \rightarrow \mathbb{E}(K_1^2) F^x(y) (1 - F^x(y)) = M_2 F^x(y) (1 - F^x(y)) \phi_x(h_K).$$

Therefore, by using Lemma 2.6.1's result, equation (2.12) becomes

$$\begin{aligned}
 \text{Var}(R_n(x, y)) &= \frac{n^2 \phi_x(h_K) (N(1, 2) h_K^2 \phi_x(h_K))^2}{((n-1) N(1, 2) M_1 h_K^2 \phi_x^2(h_K))^2} M_2 F^x(y) (1 - F^x(y)) \phi_x(h_K) \\
 &= \frac{n^2 M_2}{(n-1)^2 M_1^2} F^x(y) (1 - F^x(y)) \\
 &\rightarrow \frac{M_2}{M_1^2} F^x(y) (1 - F^x(y)) = V_{HK}(x, y) \quad \text{as } n \rightarrow \infty.
 \end{aligned}$$

Now, in order to end the proof of (2.8), we focus on the central limit theorem. So, the proof of (2.8) is completed if the Lindeberg's condition is verified. In fact, the Lindeberg's condition holds since, for any  $\eta > 0$

$$\sum_{j=1}^n \mathbb{E} \left( \varepsilon_{nj}^2 \mathbb{1}_{(|\varepsilon_{nj}| > \eta)} \right) = n \mathbb{E} \left( \varepsilon_{n1}^2 \mathbb{1}_{(|\varepsilon_{n1}| > \eta)} \right) = \mathbb{E} \left( (\sqrt{n} \varepsilon_{n1})^2 \mathbb{1}_{(|\sqrt{n} \varepsilon_{n1}| > \sqrt{n} \eta)} \right) \rightarrow 0,$$

as

$$\mathbb{E} \left( (\sqrt{n} \varepsilon_{n1})^2 \right) = n \mathbb{E}(\varepsilon_{n1}^2) \rightarrow \frac{M_2}{M_1^2} F^x(y) (1 - F^x(y)).$$

This last equation allows to finish the proof of (2.8).

*Proof of (2.9).* To use same arguments as those invoked to prove (2.8), let us write

$$T_3T_4 - \mathbb{E}(T_3T_4) = (T_4 - \mathbb{E}(T_4)) + ((T_3 - 1)T_4 - \mathbb{E}((T_3 - 1)T_4)).$$

By applying the Bienaymé-Tchebychev's inequality we obtain, for all  $\varepsilon > 0$

$$\mathbb{P}(|T_3T_4 - \mathbb{E}(T_3T_4)| > \varepsilon) \leq \frac{\mathbb{E}(|T_3T_4 - \mathbb{E}(T_3T_4)|)}{\varepsilon},$$

and the Cauchy-Schwarz's inequality implies that

$$\mathbb{E}|(T_3 - 1)T_4 - \mathbb{E}((T_3 - 1)T_4)| \leq 2\mathbb{E}|(T_3 - 1)T_4| \leq 2\sqrt{\mathbb{E}((T_3 - 1)^2)}\sqrt{\mathbb{E}(T_4^2)}.$$

By the same kind of arguments, and by using Lemma 2.6.1's result, we get

$$\begin{aligned} \mathbb{E}((T_3 - 1)^2) &= \text{Var}(T_3) = \frac{n}{n^2\mathbb{E}^2(\beta_1K_1)} \text{Var}(\beta_1K_1) \\ &\leq \frac{1}{n(O(h_K^4\phi_x^2(h_K)))} \mathbb{E}(\beta_1^4K_1^2) = O\left(\frac{1}{n\phi_x(h_K)}\right). \end{aligned}$$

On the other hand

$$\begin{aligned} \mathbb{E}(T_4^2) &= \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1K_1)}{E^2(\Delta_1K_1)} \mathbb{E}\left(\sum_{j=1}^n \beta_jK_j(H_j - F^x(y))\right)^2 \\ &= \frac{n\phi_x(h_K)O(h_K^2\phi_x^2(h_K))}{(n-1)^2O(h_K^4\phi_x^4(h_K))} \left(n\mathbb{E}(\beta_1K_1(H_1 - F^x(y)))^2\right. \\ &\quad \left.+ n(n-1)\mathbb{E}^2(\beta_1K_1(H_1 - F^x(y)))\right) \\ &= o(1) + o(n\phi_x(h_K)). \end{aligned}$$

It remains to show

$$\mathbb{E}|(T_3 - 1)T_4 - \mathbb{E}((T_3 - 1)T_4)| \leq 2\sqrt{\mathbb{E}((T_3 - 1)^2)}\sqrt{\mathbb{E}(T_4^2)} = o(1),$$

which implies that  $\mathbb{E}(|(T_3 - 1)T_4 - \mathbb{E}((T_3 - 1)T_4)|) = o_p(1)$ . Therefore,

$$\mathbb{P}(|T_3T_4 - \mathbb{E}(T_3T_4)| > \varepsilon) \leq \frac{\mathbb{E}(|T_3T_4 - \mathbb{E}(T_3T_4)|)}{\varepsilon} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

So, to prove (2.9), it suffices to show  $T_4 - \mathbb{E}(T_4) = o(1)$ , while

$$\mathbb{E}(T_4 - \mathbb{E}(T_4))^2 = \text{Var}(T_4) = \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \text{Var}(\beta_1 K_1 (H_1 - F^x(y))).$$

We arrive finally at

$$\text{Var}(\beta_1 K_1 (H_1 - F^x(y))) = F^x(y)(1 - F^x(y)) \mathbb{E}(\beta_1^2 K_1^2).$$

This last result together with Lemma 2.6.1 lead directly to

$$\begin{aligned} \mathbb{E}(T_4 - \mathbb{E}(T_4))^2 &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1 K_1)}{E^2(\Delta_1 K_1)} (F^x(y)(1 - F^x(y))) \mathbb{E}(\beta_1^2 K_1^2) \\ &= (F^x(y)(1 - F^x(y))) o(1), \end{aligned}$$

which allows to finish the proof of Theorem 1. ■

**Proof of Theorem 5.** The proof of this theorem rely on the following auxiliary lemmas for which the proofs are given in the appendix.

**Lemma 2.6.5** *Under the hypotheses of Theorem 5, we have*

$$\mathbb{E}(\widehat{f}_N^{x(j)}(y)) - f^{x(j)}(y) = B_H(x, y) h_H^2 + B_K(x, y) h_K^2 + o(h_H^2) + o(h_K^2),$$

where

$$B_H(x, y) = \frac{1}{2} \frac{h_H^2 \partial^{j+2} f^x(y)}{\partial y^{j+2}} \int t^2 H^{(1)}(t) dt \quad \text{and} \quad B_K(x, y) = \frac{1}{2} \psi_j^{(2)}(0) \frac{N(1, 2)}{M_1}.$$

**Lemma 2.6.6** *Under the hypotheses of Theorem 5, we have*

$$\text{Var}(\widehat{f}_N^{x(j)}(y)) = \frac{V_{H,K}(x, y)}{n h_H^{2j+1} \phi_x(h_K)} + o\left(\frac{1}{n h_H^{2j+1} \phi_x(h_K)}\right).$$

**Lemma 2.6.7** *Under assumptions (H1'), (H5) and (H7), as  $n \rightarrow \infty$ , we have*

$$\frac{1}{h_H} \mathbb{E}\left(K_1^2 \text{Var}(H^{(j+1)}\left(\frac{y - Y_1}{h}\right) | X_1)\right) \rightarrow M_1^2 f^x(y) \int (H^{(j+1)}(t))^2 dt.$$

Let

$$S_n = \sum_{i=1}^n (Z_i - \mathbb{E}(Z_i)) \text{ where } Z_i = \frac{\sqrt{nh_H^{2j+1}\phi_x(h_K)}}{nh_H^{j+1}\mathbb{E}(\Delta_1 K_1)} \Delta_i K_i (H^{(j+1)} - h_H^{j+1} f^x(y)).$$

Obviously, we have

$$S_n = \sqrt{nh_H^{2j+1}\phi_x(h_K)} \left( \widehat{f}_N^x(y) - f^x(y) \widehat{F}_D^x - \mathbb{E} \left( \widehat{f}_N^x(y) - f^x(y) \widehat{F}_D^x \right) \right). \quad (2.14)$$

Thus, to show the asymptotic normality

$$S_n \xrightarrow{D} \mathcal{N}(0, V_{H,K}(x, y)), \quad (2.15)$$

on one side, we have

$$\begin{aligned} S_n &= \frac{\sqrt{nh_H^{2j+1}\phi_x(h_K)}}{nh_H^{j+1}\mathbb{E}(\Delta_1 K_1)} \left( \sum_{k=1}^n \Delta_k K_k (H_k^{(j+1)} - h_H^{j+1} f^x(y)) \right) \\ &\quad - \frac{\sqrt{nh_H^{2j+1}\phi_x(h_K)}}{nh_H^{j+1}\mathbb{E}(\Delta_1 K_1)} \mathbb{E} \left( \sum_{k=1}^n \Delta_k K_k (H_k^{(j+1)} - h_H^{j+1} f^x(y)) \right). \end{aligned} \quad (2.16)$$

Combining (2.14) with (2.16) imply that

$$\begin{aligned} S_n &= (S_1 S_2 - S_3 S_4 - \mathbb{E}(S_1 S_2 - S_3 S_4)) \\ &= (S_1 S_2 - \mathbb{E}(S_1 S_2)) - (S_3 S_4 - \mathbb{E}(S_3 S_4)), \end{aligned}$$

where

$$\begin{aligned} S_1 &= \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{k=1}^n \beta_k^2 K_k, \\ S_2 &= \frac{\sqrt{nh_H^{2j+1}\phi_x(h_K)\mathbb{E}(\beta_1^2 K_1)}}{h_H^{j+1}\mathbb{E}(\Delta_1 K_1)} \sum_{k=1}^n K_k (H_k^{(j+1)} - h_H^{j+1} f^x(y)), \\ S_3 &= \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{k=1}^n \beta_k K_k, \\ \text{and } S_4 &= \frac{\sqrt{nh_H^{2j+1}\phi_x(h_K)\mathbb{E}(\beta_1 K_1)}}{h_H^{j+1}\mathbb{E}(\Delta_1 K_1)} \sum_{k=1}^n \beta_k K_k (H_k^{(j+1)} - h_H^{j+1} f^x(y)). \end{aligned}$$

So, to show the asymptotic normality (2.15), we only need to show the following two claims :

$$S_1 S_2 - \mathbb{E}(S_1 S_2) \xrightarrow{D} \mathcal{N}(0, V_{H,K}(x, y)) \quad \text{and} \quad S_3 S_4 - \mathbb{E}(S_3 S_4) \xrightarrow{P} 0,$$

which follows the same proof's steps as for proving claims (2.8) and (2.9), but to save space, these details are omitted, nevertheless they can be obtained on request. This is enough/sufficient to complete the proof of Theorem 5. ■

**Proof of Theorem 8.** From assumption (H7), we can use a Taylor expansion of  $\widehat{f}^{x(1)}(\cdot)$  in the neighborhood of  $\theta(x)$ . By assuming  $f^{x(2)}(\theta(x)) \neq 0$  and  $\widehat{f}^{x(2)}(\widehat{\theta}(x)) \neq 0$ , we get

$$\widehat{f}^{x(1)}(\widehat{\theta}(x)) = \widehat{f}^{x(1)}(\theta(x)) + (\widehat{\theta}(x) - \theta(x))\widehat{f}^{x(2)}(\theta'(x)),$$

where  $\theta'(x)$  is between  $\widehat{\theta}(x)$  and  $\theta(x)$ .

By using the fact that  $\widehat{f}^{x(1)}(y) = \widehat{f}_N^{x(1)}(y)/\widehat{F}_D^x$ , we can write

$$\widehat{\theta}(x) - \theta(x) = -\frac{\widehat{f}_N^{x(1)}(\theta(x))}{\widehat{f}_N^{x(2)}(\theta'(x))} = -\frac{\widehat{f}_N^{x(1)}(\theta(x)) - \mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right)}{\widehat{f}_N^{x(2)}(\theta'(x))} - \frac{\mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right)}{\widehat{f}_N^{x(2)}(\theta'(x))}.$$

Then

$$\begin{aligned} & \sqrt{nh_H^3 \phi_x(h_K) (\widehat{f}^{x(2)}(\theta(x)))^2 (\theta(x) - \widehat{\theta}(x))} \\ &= \sqrt{nh_H^3 \phi_x(h_K)} \left( \widehat{f}_N^{x(1)}(\theta(x)) - \mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right) \right) \\ &+ \sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right). \end{aligned} \quad (2.17)$$

Hence, to show (2.17), by the Slutsky's Theorem it suffices to prove the following two claims :

$$\sqrt{nh_H^3 \phi_x(h_K)} \left( \widehat{f}_N^{x(1)}(\theta(x)) - \mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right) \right) \xrightarrow{D} \mathcal{N}(0, V_{H,K}(x, y)) \quad (2.18)$$

and

$$\sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E}\left(\widehat{f}_N^{x(1)}(\theta(x))\right) \xrightarrow{P} 0 \quad (2.19)$$

*Proof of (2.18).* Let

$$S_n = \sum_{i=1}^n (A_i - \mathbb{E}(A_i)) \text{ where } A_i = \frac{\sqrt{nh_H^3 \phi_x(h_K)}}{nh_H^2 \mathbb{E}(\Delta_1 K_1)} \Delta_i K_i H_i^{(2)}.$$

Obviously, we have

$$S_n = \sqrt{nh_H^3 \phi_x(h_K)} \left( \widehat{f}_N^{x(1)}(\theta(x)) - \mathbb{E} \left( \widehat{f}_N^{x(1)}(\theta(x)) \right) \right).$$

Thus, to prove the asymptotic normality of  $S_n$ , it is sufficient to apply the Lindeberg's central limit theorem on  $A_i$  for which it suffices to show that

$$\frac{1}{\text{Var}(S_n)} \sum_{i=1}^n \mathbb{E} \left( A_i^2 \mathbb{1}_{|A_i| > \varepsilon} \text{Var}(S_n) \right) \rightarrow 0, \text{ for all } \varepsilon > 0.$$

Since

$$|A_i| \leq \frac{C \sqrt{nh_H^{(3)} \phi_x(h_K)}}{n(n-1)h_H^2 \mathbb{E}(\Delta_{12})} \text{ and } \mathbb{E}(\Delta_{12}) = O(h_K^2 \phi_x^2(h_K)),$$

$$\text{Var}(S_n) = nh_H^{(3)} \phi_x(h_K) \text{Var}(\widehat{f}_N^{x(1)}(\theta(x))),$$

and by Lemma 2.6.6's result, we obtain

$$\text{Var}(\widehat{f}_N^{x(1)}(\theta(x))) = \frac{V_{H,K}(x, \theta(x))}{nh_H^3 \phi_x(h_K)} + o\left(\frac{1}{nh_H^3 \phi_x(h_K)}\right).$$

Then

$$\text{Var}(S_n) = V_{H,K}(x, \theta(x)),$$

and since

$$|A_i| \leq \frac{C \sqrt{nh_H^{(3)} \phi_x(h_K)}}{n(n-1)h_H^2 O(h_K^2 \phi_x^2(h_K))},$$

that leads to

$$\frac{1}{\text{Var}(S_n)} \sum_{i=1}^n \mathbb{E} \left( A_i^2 \mathbb{1}_{|A_i| > \varepsilon} \text{Var}(S_n) \right) \leq C \sum_{i=1}^n \mathbb{P}(|A_i| > \varepsilon \text{Var}(S_n)).$$

On the other hand, when  $n \rightarrow \infty$ , we have

$$\frac{|A_i|}{\text{Var}(S_n)} \leq \frac{C}{V_{H,K}(x, \theta(x)) h_H^{-1/2} (n(n-1)^2 O(h_K^4 \phi_x^3(h_K)))^{1/2}} \rightarrow 0$$



So, for any  $\varepsilon > 0$ , if  $n$  is great enough, then  $\mathbb{P}(|A_i| > \varepsilon \text{Var}(S_n)) = 0$ , which implies that

$$\frac{1}{\text{Var}(S_n)} \sum_{i=1}^n \mathbb{E} \left( A_i^2 \mathbb{1}_{|A_i| > \varepsilon} \text{Var}(S_n) \right) = 0,$$

which yields to the proof of (2.18).

*Proof of (2.19).* On one hand we have

$$\mathbb{E}(\widehat{f}_N^{x(1)}(\theta(x))) = \frac{1}{nh_H^2 \mathbb{E}(\Delta_1 K_1)} \sum_{i=1}^n \mathbb{E}(\Delta_i K_i H_i^{(2)}) = \frac{\mathbb{E}(\Delta_1 K_1 \mathbb{E}(H_1^{(2)} | X_1))}{h_H^2 \mathbb{E}(\Delta_1 K_1)}.$$

On the other hand, by using an integration by parts followed by a change of variable, we get

$$\begin{aligned} \mathbb{E} \left( H^{(2)} \left( \frac{\theta(x) - Y_1}{h_H} \right) | X_1 \right) &= \int_{\mathbb{R}} H^{(2)} \left( \frac{\theta(x) - u}{h_H} \right) f^{X_1}(u) du \\ &= h_H^2 \int_{\mathbb{R}} H^{(1)}(t) f^{(1)(X_1)}(\theta(x) - th_H) dt. \end{aligned}$$

Now, by using a two order Taylor expansion of  $f^{(1)(X_1)}$  around  $\theta(x)$  we can deduce that

$$\begin{aligned} \sqrt{(nh_H^3 \phi_x(h_k))} \mathbb{E}(\widehat{f}_N^{x(1)}(\theta(x))) &= \sqrt{(nh_H^3 \phi_x(h_k))} \int_{\mathbb{R}} H^{(1)}(t) [f^{(1)(X_1)}(\theta(x)) dt] \\ &\quad - \sqrt{(nh_H^3 \phi_x(h_k))} \int_{\mathbb{R}} H^{(1)}(t) [th_H f^{(2)(X_1)}(\theta(x)) dt] \\ &\quad + \sqrt{(nh_H^7 \phi_x(h_k))} \int_{\mathbb{R}} t^2 H^{(1)}(t) f^{(3)(X_1)}(\theta'(x)) dt. \end{aligned}$$

Then (2.19) is a straightforward consequence of this last result with assumptions (H10) and (H11). ■

## 2.7 Appendix

**Proof of Lemma 2.6.2.** By applying the Bienaymé-Tchebychev's inequality, as  $n \rightarrow \infty$ , we obtain, for all  $\varepsilon > 0$

$$\mathbb{P}(|\widehat{F}_D^x - \mathbb{E}(\widehat{F}_D^x)| > \varepsilon) \leq \frac{\text{Var}(\widehat{F}_D^x)}{\varepsilon^2} \leq \varepsilon^{-2} O\left(\frac{1}{n\phi_x(h_K)}\right) \rightarrow 0.$$

■

**Proof of Lemma 2.6.5.** We start by writing

$$\begin{aligned} \mathbb{E} \left( \widehat{f}_N^{x(j)}(y) \right) &= \frac{n}{nh_H^{j+1} E(\Delta_1 K_1)} \mathbb{E} \left( \Delta_1 K_1 H_1^{(j+1)} \right) \\ &= \frac{1}{h_H^{j+1} \mathbb{E}((n-1)W_{12})} \mathbb{E} \left( (n-1)W_{12} \mathbb{E} \left( H_1^{(j+1)} | X_1 \right) \right), \end{aligned} \quad (2.20)$$

where

$$\mathbb{E} \left( H_1^{(j+1)} | X_1 \right) = h_H^{j+1} \int_{\mathbb{R}} H^{(1)}(t) f^{X_1(j)}(y - h_H t) dt.$$

Then, by using a Taylor expansion, of order two, of  $f^{X_1(j)}$  and under assumption (H2'), we obtain

$$\mathbb{E} \left( H_1^{(j+1)} | X_1 \right) = h_H^{j+1} \left( f^{X_1(j)}(y) + \frac{h_H^2 t^2}{2} \int H_1^{(1)}(t) dt \frac{\partial^{j+2} f^{X_1}(y)}{\partial y^{j+2}} + o(h_H^2) \right).$$

Then

$$\begin{aligned} \mathbb{E} \left( \widehat{f}_N^{x(j)}(y) \right) &= \frac{1}{\mathbb{E}(W_{12})} \left( \mathbb{E} \left( W_{12} f^{X_1(j)}(y) \right) \right. \\ &\quad \left. + \frac{h_H^2 t^2}{2} \int H_1^{(1)}(t) dt \mathbb{E} \left( W_{12} \frac{\partial^{j+2} f^{X_1}(y)}{\partial y^{j+2}} \right) + o(h_H^2) \right), \end{aligned}$$

and by assumption (H1'), we get

$$\begin{aligned} \mathbb{E} \left( \widehat{f}_N^{x(j)}(y) \right) &= \frac{1}{\mathbb{E}(W_{12})} \left( \mathbb{E} (W_{12} \varphi_j(X_1, y)) \right. \\ &\quad \left. + \frac{h_H^2 t^2}{2} \int H_1^{(1)}(t) dt \mathbb{E} (W_{12} \varphi_{j+2}(X_1, y)) + o(h_H^2) \right). \end{aligned}$$

Since  $\mathbb{E}(\beta_1 W_{12}) = 0$ , from assumption (H1') with  $l \in \{j, j+2\}$ , and the fact that  $\psi_l(0) = 0$ , we obtain

$$\begin{aligned} &\mathbb{E} (W_{12} \varphi_l(X_1, y)) \\ &= \varphi_l(x, y) \mathbb{E} (W_{12}) + \mathbb{E} (W_{12} \mathbb{E} (\varphi_l(X_1, y) - \varphi_l(x, y) | \beta(X_1, x))), \\ &= \varphi_l(x, y) \mathbb{E} (W_{12}) + \frac{1}{2} \psi_l^{(2)}(0) \mathbb{E} (W_{12} (\beta(X_1, x)^2)) + o(\mathbb{E} (W_{12} (\beta(X_1, x))))). \end{aligned}$$

Thus

$$\begin{aligned} & \mathbb{E} \left( \widehat{f}_N^{x(j)}(y) \right) \\ &= f^{x(j)}(y) + \frac{h_H^2}{2} \frac{\partial^{j+2} f^x(y)}{\partial y^{j+2}} \int t^2 H^{(1)}(t) dt + o \left( h_H^2 \frac{\mathbb{E}(\beta(X_1, x)W_{12})}{\mathbb{E}(W_{12})} \right) \\ & \quad + \frac{1}{2} \psi_j^{(2)}(0) \frac{\mathbb{E}(\beta^2(X_1, x)W_{12})}{\mathbb{E}(W_{12})} + o \left( \frac{\mathbb{E}(\beta(X_1, x)W_{12})}{\mathbb{E}(W_{12})} \right). \end{aligned}$$

Elsewhere, it is clear that  $\mathbb{E}(\beta^2(X_1, x)W_{12}) = \mathbb{E}(K_1\beta_1^2)^2 - \mathbb{E}(K_1\beta_1)\mathbb{E}(K_1\beta_1^3)$ . Then, by using Lemma 2.6.1's result, we find

$$\mathbb{E}(\beta^2(X_1, x)W_{12}) = N(1, 2)^2 h_K^4 \phi_x(h_K)^2 - o(h_K \phi_x(h_K)) O(h_K^3 \phi_x(h_K)), \quad (2.21)$$

$$\text{and} \quad \mathbb{E}(W_{12}) = N(1, 2) M_1 h_K^2 \phi_x^2(h_K). \quad (2.22)$$

Finally, by combining (2.21) and (2.22) we obtain

$$\frac{\mathbb{E}(\beta^2(X_2, x)W_{12})}{\mathbb{E}(W_{12})} = h_K^2 \frac{N(1, 2)}{M_1} + o(h_K^2).$$

■

**Proofs of Lemmas 2.6.3 and 2.6.7.** We have

$$\begin{aligned} & \frac{1}{h_H^l} \mathbb{E} \left( K_1^2 \text{Var} \left( H^{(j+1)} \left( \frac{y - Y_1}{h} \right) \middle| X_1 \right) \right) \\ &= \frac{1}{h_H^l} \mathbb{E} \left( K_1^2 \mathbb{E} \left( \left( H^{(j+1)} \left( \frac{y - Y_1}{h_H} \right) \right)^2 \middle| X_1 \right) \right) \end{aligned} \quad (2.23)$$

$$- h_H^l \mathbb{E} \left( K_1^2 \mathbb{E}^2 \left( \frac{1}{h_H^l} H^{(j+1)} \left( \frac{y - Y_1}{h_H} \right) \middle| X_1 \right) \right). \quad (2.24)$$

Concerning the term (2.23), under assumptions (H5) (for  $j = -1$  and  $l = 0$ ) and (H7) (for  $j \geq 0$  and  $l = 1$ ) and by an integration by parts followed by a change of variable, we get

$$\begin{aligned} \frac{1}{h_H^l} \mathbb{E} \left( \left( H^{(j+1)} \left( \frac{y - Y_1}{h_H} \right) \right)^2 \middle| X_1 \right) &= h_H^{-l} \int_{\mathbb{R}} \left( H^{(j+1)} \left( \frac{y - z}{h_H} \right) \right)^2 f^{X_1}(z) dz \\ &= h_H^{1-l} \int_{\mathbb{R}} (H^{(j+1)}(t))^2 f^{X_1}(y - th_H) dt \\ &= h_H^{1-l} \int_{\mathbb{R}} (H^{(j+1)}(t))^2 dF^{X_1}(y - th_H) dt. \end{aligned}$$

*Proof of Lemma 2.6.3.* For  $j = -1$  and  $l = 0$ , by noting  $H^{(0)}\left(\frac{y - Y_1}{h_H}\right) = H\left(\frac{y - Y_1}{h_H}\right)$ , we have

$$\begin{aligned} \mathbb{E}\left(H^2\left(\frac{y - Y_1}{h_H}\right) \middle| X_1\right) &= \int_{\mathbb{R}} H^2(t) dF^{(X_1)}(y - th_H) \\ &= \int_{\mathbb{R}} 2H(t)^{(1)} H(t) \left(F^{(X_1)}(y - th_H) - F^x(y)\right) dt \\ &\quad + \int_{\mathbb{R}} 2H(t)^{(1)} H(t) F^x(y) dt. \end{aligned}$$

Since  $\int_{\mathbb{R}} 2H^{(1)}(t) H(t) F^x(y) dt = F^x(y)$ , as  $n \rightarrow \infty$ , we deduce that

$$\mathbb{E}\left(K_1^2 H^2\left(\frac{y - Y_1}{h_H}\right) \middle| X_1\right) \rightarrow \mathbb{E}(K_1^2) F^x(y),$$

and

$$\mathbb{E}\left(H\left(\frac{y - Y_1}{h_H}\right) \middle| X_1\right) - F^x(y) \rightarrow 0.$$

So, the term (2.24) tends to  $(F^x(y))^2$  as  $n$  tends to infinity. Then

$$\mathbb{E}\left(K_1^2 \mathbb{E}^2\left(H\left(\frac{y - Y_1}{h_H}\right) \middle| X_1\right)\right) \rightarrow \mathbb{E}\left(K_1^2 (F^x(y))^2\right) = \mathbb{E}\left(K_1^2\right) (F^x(y))^2.$$

Finally, as  $n \rightarrow \infty$ , we have

$$\mathbb{E}\left(K_1^2 \text{var}\left(H\left(\frac{y - Y_1}{h_H}\right) \middle| X_1\right)\right) \rightarrow \mathbb{E}(K_1^2) F^x(y) (1 - F^x(y)).$$

*Proof of Lemma 2.6.7.* For  $j \geq 0$  and  $l = 1$ , we have

$$\begin{aligned} \frac{1}{h_H^l} \mathbb{E}\left(\left(H^{(j+1)}\left(\frac{y - Y_1}{h_H}\right)\right)^2 \middle| X_1\right) &= \int_{\mathbb{R}} (H^{(j+1)}(t))^2 \left(f^{X_1}(y - th_H) - f^x(y)\right) dt \\ &\quad + \int_{\mathbb{R}} (H^{(j+1)}(t))^2 f^x(y) dt. \end{aligned}$$

Remark that  $(f^{X_1}(y - th_H) - f^x(y)) \rightarrow o(1)$  as  $n \rightarrow \infty$ . Then, as  $n \rightarrow \infty$ , we deduce that

$$\frac{1}{h_H^l} \mathbb{E}\left(\left(H^{(j+1)}\left(\frac{y - Y_1}{h_H}\right)\right)^2 \middle| X_1\right) \rightarrow f^x(y) \int_{\mathbb{R}} (H^{(j+1)}(t))^2 dt,$$

and

$$\begin{aligned}
 \mathbb{E} \left( H_1^{(j+1)}(y) | X_1 \right) &= \int_{\mathbb{R}} H^{(j+1)}(h_H^{-1}(y-z)) f^{X_1}(z) dz \\
 &= - \sum_{l=1}^j h_H^l \left[ H^{(j-l-1)}(h_H^{-1}(y-z)) f^{X_1^{(l-1)}}(z) \right]_{-\infty}^{+\infty} \\
 &\quad + h_H^j \int_{\mathbb{R}} H^{(1)}(h_H^{-1}(y-z)) f^{X_1^{(j)}}(z) dz. \tag{2.25}
 \end{aligned}$$

So, assumption (H7) allows to cancel the first term on the right side of (2.25). Then, by using a Taylor expansion followed by a change of variable, we obtain

$$\frac{1}{h_H} \mathbb{E} \left( H_1^{(j+1)}(y) | X_1 \right) - h_H^j f^{X_1^{(j)}}(y) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof of Lemma 2.6.6.* We have that

$$\begin{aligned}
 \text{Var} \left( \widehat{f}_N^{x(j)}(y) \right) &= \frac{1}{\left( n(n-1) h_H^{j+1} \mathbb{E}(W_{12}) \right)^2} \text{Var} \left( \sum_{i \neq k=1}^n W_{ik} H_k^{(j+1)} \right) \\
 &= \frac{1}{\left( n(n-1) h_H^{j+1} \mathbb{E}(W_{12}) \right)^2} \left( n(n-1) \mathbb{E} \left( W_{12}^2 H_2^{(j+1)^2} \right) \right. \\
 &\quad + n(n-1) \mathbb{E} \left( W_{12} W_{21} H_1^{(j+1)} H_2^{(j+1)} \right) \\
 &\quad + n(n-1)(n-2) \mathbb{E} \left( W_{12} W_{13} H_2^{(j+1)} H_3^{(j+1)} \right) \\
 &\quad + n(n-1)(n-2) \mathbb{E} \left( W_{12} W_{23} H_2^{(j+1)} H_3^{(j+1)} \right) \\
 &\quad + n(n-1)(n-2) \mathbb{E} \left( W_{12} W_{31} H_2^{(j+1)} H_1^{(j+1)} \right) \\
 &\quad + n(n-1)(n-2) \mathbb{E} \left( W_{12} W_{32} H_2^{(j+1)^2} \right) \\
 &\quad \left. - n(n-1)(4n-6) \left( \mathbb{E} \left( W_{12} H_2^{(j+1)} \right) \right)^2 \right). \tag{2.26}
 \end{aligned}$$

Observe that

$$\begin{aligned}
 \mathbb{E} \left( W_{12} H_1^{(j+1)} \right) &= \mathbb{E} \left( (\beta_1^2 K_1 K_2 - \beta_1 \beta_2 K_1 K_2) H_1^{(j+1)} \right) \\
 &\leq \mathbb{E}(\beta_1^2 K_1) \cdot \mathbb{E}(H_1^{j+1} K_1) = O(h_H^{j+1} h_K^2 \phi_x^2(h_K)). \tag{2.27}
 \end{aligned}$$

Then

$$\frac{1}{h_H^{2j+2}} \frac{\mathbb{E} \left( W_{12} H_1^{(j+1)} \right)^2}{\mathbb{E} (W_{12})} = \frac{1}{h_H^{2j+2}} \frac{O(h_H^{2j+2} h_K^4 \phi_x^4(h_K))}{O(h_K^4 \phi_x^4(h_K))} = O(1).$$

By some simple manipulations and by using Lemma 2.6.1's result, we get

$$\left\{ \begin{array}{l} \mathbb{E} \left( W_{12}^2 H_2^{(j+1)^2} \right) = O(h_K^4 h_H \phi_x^2(h_K)), \\ \mathbb{E} \left( W_{12} W_{21} H_1^{(j+1)} H_2^{(j+1)} \right) = O(h_K^4 h_H^{2j+2} \phi_x^2(h_K)), \\ \mathbb{E} \left( W_{12} W_{13} H_2^{(j+1)} H_3^{(j+1)} \right) = O(h_K^4 h_H^{2j+2} \phi_x^3(h_K)), \\ \mathbb{E} \left( W_{12} W_{23} H_2^{(j+1)} H_3^{(j+1)} \right) = O(h_K^4 h_H^{2j+2} \phi_x^3(h_K)), \\ \mathbb{E} \left( W_{12} W_{31} H_2^{(j+1)} H_1^{(j+1)} \right) = O(h_K^4 h_H^{2j+2} \phi_x^3(h_K)), \\ \mathbb{E} \left( W_{12} W_{32} H_2^{(j+1)^2} \right) = (\mathbb{E} (\beta_1^2 K_1))^2 \mathbb{E} \left( K_1^2 H_1^{(j+1)^2} \right) + O(h_K^4 h_H \phi_x^3(h_K)), \\ \mathbb{E} \left( W_{12} H_2^{(j+1)} \right)^2 = O(h_K^4 h_H^{2j+2} \phi_x^2(h_K)). \end{array} \right. \quad (2.28)$$

Therefore, the leading term in the expression of  $\text{Var} \left( \widehat{f}_N^{x(j)}(y) \right)$  is

$$\frac{n(n-1)(n-2)}{(n(n-1)h_H^{j+1} \mathbb{E}(W_{12}))^2} \mathbb{E} \left( \beta_1^2 K_1 \right)^2 \mathbb{E} \left( K_1^2 \left( H_1^{(j+1)} \right)^2 \right).$$

Combining equations (2.26)-(2.27) with (2.28), we obtain

$$\text{Var} \left( \widehat{f}_N^{x(j)}(y) \right) = \frac{\mathbb{E} \left( K_1^2 \left( H_1^{(j+1)} \right)^2 \right)}{nh_H^{2j+2} (\mathbb{E} (K_1))^2} + o \left( \frac{1}{nh_H^{2j+1} \phi_x(h_K)} \right). \quad (2.29)$$

Remark that

$$\mathbb{E} \left( K_1^2 \left( H_1^{(j+1)} \right)^2 \right) = \mathbb{E} \left( K_1^2 \mathbb{E} \left( \left( H_1^{(j+1)} \right)^2 \mid X \right) \right).$$

Hence, from Lemma 2.6.4's proof, as  $n \rightarrow \infty$ , we obtain

$$\mathbb{E} \left( K_1^2 \left( H_1^{(j+1)} \right)^2 \left( \frac{y - Y_1}{h_H} \right) \mid X_1 \right) \rightarrow h_H \mathbb{E} (K_1^2) f^x(y) \int H^{(j+1)}(t) dt. \quad (2.30)$$

Combining equations (2.29) and (2.30), leads to

$$\text{Var} \left( \widehat{f}_N^{x(j)}(y) \right) = \frac{f^x(y) \int \left( H^{(j+1)}(t) \right)^2 dt}{nh_H^{2j+1} \phi_x(h_K)} \frac{M_2}{M_1^2} + o \left( \frac{1}{nh_H^{2j+1} \phi_x(h_K)} \right). \quad (2.31)$$

# Bibliographie

- [1] Baïllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102-111.
- [2] Barrientos, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632.
- [3] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data. *Comput. Statist.*, **22**, 353-369.
- [4] Berlinet, A., Elamine, A. and Mas, A. (2011). Local linear regression for functional data. *Inst. Statist. Math.*, **63**, 1047-1075.
- [5] Boj, E., Delicado, P. and Fortiana, J. (2010). Distance-based local linear regression for functional predictors. *Comput. Statist and Data Anal.*, **54**, 429-437.
- [6] Chen, J., Zhu, R., Xu, R., Zhang, W., Shen, Y. and Zhang, Y. (2015). Evaluation of *Leymus chinensis* quality using near-infrared reflectance spectroscopy with three different statistical analyses. *Peer J 3 :e1416*. <https://doi.org/10.7717/peerj.1416>
- [7] Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data. *Statistica*, **70**, 83-94.
- [8] Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional time series analysis. *Statistics*, **43**, 13-33.
- [9] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics*, **47**, 26-44.

- 
- [10] Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the Local Linear Modelization of the Conditional Distribution for Functional Data. *Sankhya*, **76**, 328-355.
- [11] El Methni, M. and Rachdi, M. (2011). Local weighted average estimation of the regression operator for functional data. *Comm. Statist.-Theory and Meth.*, **40**, 3141-3153.
- [12] Ezzahrioui, M. and Ould-Saïd, E. (2008a). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Statist.*, **20**, 3-18.
- [13] Ezzahrioui, M. and Ould-Saïd, E. (2008b). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East J. Theoretical Statist.*, **25**, 15-38.
- [14] Ezzahrioui, M. and Ould-Saïd, E. (2008c). Asymptotic results of a nonparametric conditional quantile estimator for functional time series *Comm. Statist.-Theory and Meth.*, **37**, 2735-2759.
- [15] Ferraty, F., Mas, A. and Vieu, P. (2007). Nonparametric regression on functional data : inference and practical aspects. *Australian New Zealand J. Statistics*, **49**, 267-286.
- [16] Ferraty, F., Laksaci, A. and Vieu, P. (2006a). Estimation of some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inference Stoch. Process.*, **9**, 47-76.
- [17] Ferraty, F. and Vieu, P. (2006b). *Nonparametric functional data analysis : Theory and Practice*. Springer Series in Statistics, New York.
- [18] 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**, 335-352.
- [19] Ferraty F., Laksaci A., Tadj A. and Vieu P. (2011). Kernel regression with functional response. *Elec. J. Statist.*, **5**, 159-171.
- [20] Kudraszow, N.L. and Vieu, P. (2013). Uniform consistency of kNN regressors for functional variables. *Statist and Probab. Lett.*, **83**, 1863–1870.



- [21] Kara-Zaïtri, L., Laksaci, A., Rachdi, M. and Vieu, P. (2017a). *Uniform in the smoothing parameter consistency results in functional regression*. In : Aneiros G., G. Bongiorno E., Cao R., Vieu P. (eds). *Functional Statistics and Related Fields. Contributions to Statistics*. Springer, Cham., ISBN 978-3-319-55845- 5, 161–167.
- [22] Kara-Zaïtri, L., Laksaci, A., Rachdi, M. and Vieu, P. (2017b). Data-driven kNN estimation in nonparametric functional data analysis. *J. Multivariate Anal.*, **153**, 176–188.
- [23] Laksaci, A., Rachdi, M. and Rahmani, S., (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1-23.
- [24] Louani, D. and Ould-Saïd, E. (1999). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametr. Statist.*, **11**, 413-442.
- [25] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.
- [26] Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A. and Madani, F. (2014). Theoretical and practical aspects on the quadratique error in the local linear estimation of the conditional density for functional. *Comput. Statist and Data Anal.*, **73**, 53-68.
- [27] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, New York.
- [28] Zhou, Z. and Lin, Z.-Y. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.*, **28**, 116-131.

Chapitre **3**

# Local linear conditional cumulative distribution function with mixing data

Ce chapitre fait l'objet d'une publication parue dans la revue internationale Arabian Journal of Mathematics.



### Local linear conditional cumulative distribution function with mixing data

Oussama Bouanani <sup>1</sup>, Saâdia Rahmani <sup>1</sup>, and Larbi Ait-Hennani <sup>2</sup>

<sup>1</sup> Laboratoire de Modèles Stochastiques, Statistique et Applications, Université Dr moulay Tahar, Saïda 20000, Algeria. E-mail : oussamaproba@gmail.com. E-mail : saadia.rahmani@gmail.com.

<sup>2</sup> Université Lille 2, Droit et santé, IUT C, Roubaix, France.  
E-mail :larbi.aithennani@univ-lille2.fr

**Abstract :** This paper investigates a conditional cumulative distribution of a scalar response given by a functional random variable with an  $\alpha$ -mixing stationary sample using a local polynomial technique. The main purpose of this study is to establish asymptotic normality results under selected mixing conditions satisfied by many time-series analysis models in addition to the other appropriate conditions to confirm the planned prospects.

AMS Classification : 62G05, 62G08, 62G20, 62G07, 62G30. 62H12.

## 3.1 Introduction and motivations

With the evolution of the measuring instruments and the growth of research studies mainly since the publication of the pioneer paper of Ferraty and Vieu [19], functional data analysis (FDA) has attracted the attention of many works as in the recent monograph of Horváth and Kokoszka [23]. On the other hand, alternative conditional predictions of the classical regression have also gained a considerable interest in basically all fields of statistics, especially for estimating conditional models by using the kernel approach (or local constant) as investigated in the paper of Ferraty et al. [18], Dabo-Niang and Laksaci [7] or the paper of Ezzahrioui and Ould-Saïd [15].

Reminding that in numerous nonparametric statistic problems, the estimation of a conditional distribution function (CDF) constitutes a key aspect of inference. Accordingly, the present study employs a specific CDF model for constructing prediction intervals that can be involved in many applications such as the survival analysis and reliability,... etc. Interestingly, it is well known that the CDF has the advantage to completely characterize the conditional law of the considered random variables. The determination of the CDF allows, in fact, to obtain the conditional density and conditional hazard functions. Moreover, several prediction tools can also be implemented for the nonparametric statistics

modeling, taking the example of conditional mode, median, or quantile. In addition, an extensive literature including various nonparametric approaches has taken a place in the conditional estimation of independent samples and dependent observations in finite- and infinite-dimensional spaces, (see for instance, Berline et al. [3], Honda [22], and Ferraty and Vieu [17]). In many situations, the kernel approach cannot adequately estimate the conditional models for the reason that this technique suffers from a large bias particularly at the boundary region.

However, the kernel approach can be improved using local polynomial smoothers, and especially local linear smoothers, because they correct the asymptotic bias that is adversely affected at the boundaries (see Fan and Gijbels [16] for more discussions on this subject in the real case ). In recent years, there has been a strong interest in local linear smoothers in the infinite-dimensional space (see, for instance, Baïllo and Grané [1] and Barrientos-Marina et al.[2]. And it should be noted that the last precursor work has been extended in many directions, including asymptotic properties (see Demongeot et al. [11, 12] and Zhou and Lin [37]), nature of the variables (see Demongeot et al. [14]), or the dependence type (see Demongeot et al. [10] and Laksaci et al. [27]).

In this regard, our interest in this paper is to give a result concerning the limit in distribution of the estimate of the CDF, by the local linear fit. More precisely, we consider the case when the observations  $(X_i, Y_i)_{i \geq 0}$  are strongly mixing. We prove the asymptotic normality of a local linear estimator of the CDF by utilizing the appropriate form of Bernstein's blocking arguments and a reduction analysis leading to the Lindeberg–Feller central limit Theorem. We point out that this contribution has a potential impact in practice as well as in the theory. Indeed, from a practical point of view, this asymptotic property is used to derive confidence intervals or to make statistical tests. On the other hand, from a theoretical point of view, the asymptotic normality is a basic ingredient to determine the mean quadratic error or to study the uniform integrability of the estimator.

Accordingly, this work is mainly structured as follows : Sect 2 will present the model selected for study, describe the estimation method through giving the explicit solution to the minimization problem, and provide some basic assumptions and notations. Sect 3 will state the main asymptotic normality results achieved by the conditional distribution function estimator, indicating that their accuracy will lead to interesting perspectives. Finally, Sect 4 will discuss the applicability of the provided asymptotic results on some statistical problems such determination of confidence intervals. Detailed proofs of the main results will be consequently postponed to the appendix.

## 3.2 The model

The model is defined in the following way. Assuming that  $(X_i, Y_i)_{1 \leq i \leq n}$  is stationary  $\alpha$ -mixing process. The  $X_i$  are random variables with values in a functional space  $\mathfrak{F}$  where the random variables  $Y_i$  are real. In all the sequel  $\mathfrak{F}$ , a semi-metric space endowed with a semi-metric  $d(\cdot, \cdot)$  is taken into consideration. For  $x \in \mathfrak{F}$ , the conditional probability distribution of  $Y_i$  which is given by  $X_i = x$  is classically written as :

$$\forall y \in \mathbb{R}, \quad F^x(y) = \mathbb{P}(Y_i \leq y | X_i = x).$$

This distribution is absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}$ .

### 3.2.1 The estimate

The conditional cumulative distribution function  $F^x$  is estimated by  $\hat{a}$  where the couple  $(\hat{a}, \hat{b})$  is obtained by the optimization rule :

$$\sum_{i=1}^n (H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x))^2 K(h_K^{-1}\delta(x, X_i)), \quad (3.1)$$

where  $\beta(\cdot, \cdot)$  and  $\delta(\cdot, \cdot)$  are known functions from  $\mathfrak{F}^2$  into  $\mathbb{R}$ ,  $K$  is a kernel,  $H$  is a cumulative distribution function, and  $h_K$  and  $h_H$  are the bandwidths parameters. However, if the bifunctional operator  $\beta$  is such that,  $\forall z \in \mathfrak{F}$ ,  $\beta(z, z) = 0$ , then the quantity  $\hat{F}^x(y)$  is explicitly defined by the following :

$$\hat{F}^x(y) = \frac{\sum_{i,j=1}^n w_{ij} H(h_H^{-1}(y - Y_j))}{\sum_{i,j=1}^n w_{ij}} = \frac{\sum_{j=1}^n \Delta_j K_j H_j}{\sum_{j=1}^n \Delta_j K_j}, \quad (3.2)$$

with  $w_{ij} = \beta_i(\beta_i - \beta_j) K_i K_j$ ,  $\Delta_j = K_j^{-1}(\sum_{i=1}^n w_{ij}) = \sum_{i=1}^n \beta_i^2 K_i - (\sum_{i=1}^n \beta_i K_i)\beta_j$ , where

$\beta_i = \beta(X_i, x)$ ,  $K_i = K(h_K^{-1}(\delta(x, X_i)))$ , and  $H_j = H(h_H^{-1}(y - Y_j))$ .

Several asymptotic properties of this estimation are recently obtained. It turns out that the existing literature addresses either the statement of almost-complete consistencies or a mean-square error (see Demongeot et al. [12]).

### 3.2.2 Assumptions and notations

In what follows,  $x$  (resp.  $y$ ) will denote a fixed point in  $\mathfrak{F}$  (resp. in  $\mathbb{R}$ ),  $\mathcal{N}_x$  (resp.  $\mathcal{N}_y$ ) will denote a fixed neighborhood of  $x$  (resp. of  $y$ ) and  $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$ , and let  $G$  be the real valued function defined as for any  $l \in \{0, 2\}$  :

$$G_l(s) = \mathbb{E}[g_l(X, y) - g_l(x, y) | \beta(x, X) = s] \quad \text{with } g_l(x, y) = \frac{\partial^l F^x(y)}{\partial y^l}.$$

We now state some conditions which ensure asymptotic normality of (3.2) :

(H1) (i) For any  $r > 0$ ,  $\phi_x(r) := \phi_x(-r, r) > 0$ , and there exists a function  $\Psi_x(\cdot)$  such that :

$$\text{for all } t \in [-1, 1], \quad \lim_{h_K \rightarrow 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \Psi_x(t).$$

(ii) For any  $l \in \{0, 2\}$ , the quantities  $G_l^{(2)}(0)$  exist.

(H3)  $\sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) \leq \psi_x(h_K)$ , where  $\psi_x(h_K) = O(\phi_x^2(h_K))$ .

(H4) The coefficients of  $\alpha$ -mixing sequence  $(X_i, Y_i)_{i \in \mathbb{N}^*}$  satisfy the following two conditions :

$$(i) \exists s > 1, \exists c > 0 : \forall n \in \mathbb{N}; \quad \alpha(n) \leq cn^{-s},$$

$$(ii) \sum_{k=1}^{+\infty} k^\delta (\alpha(k))^{1/p} < \infty \text{ for some } p > 0 \text{ and } \delta > \frac{1}{p}.$$

(H5) The bandwidth  $h_K$  satisfies :

$$(i) \lim_{n \rightarrow +\infty} n \phi_x^m(h_K) = +\infty, \quad \text{for } m = 1, 2,$$

$$(ii) \lim_{n \rightarrow +\infty} n^\varsigma \phi_x(h_K) = 0, \quad \text{for } \varsigma = 2, 3,$$

$$(iii) \exists C_1 > 0, \quad C_2 > 0, \quad C_2 n^{1-p} \leq \phi_x(h_K) \leq C_1 n^{\frac{1}{1-p}} \text{ for } p \geq 2.$$

(H6) The locating operator  $\beta$  satisfies the following two conditions :

$$(i) \forall z \in \mathfrak{F}, \quad C_1 |\delta(x, z)| \leq |\beta(x, z)| \leq C_2 |\delta(x, z)| \text{ where } C_1 > 0, C_2 > 0,$$

$$(ii) h_K \int_{B(x, h_K)} \beta(u, x) dP(u) = o\left(\int_{B(x, h_K)} \beta^2(u, x) dP(u)\right), \text{ where } B(x, r) = \{z \in \mathfrak{F} / |\delta(z, x)| \leq r\}$$

and  $dP(x)$  is the cumulative distribution of  $X$ .

(H7) (i)  $K$  is a positive, differentiable function with support  $[-1, 1]$ .

(ii)  $H$  is a positive, bounded, and Lipschitzian continuous function, satisfying that

$$\int |t|^{b_2} H'(t) dt < \infty,$$

$$\int H'(t) dt = 1 \text{ and } \int H^2(t) dt < \infty.$$

- (H8) Let  $(r_n)$  and  $(v_n)$  be sequences of positive integers tending to infinity, such that  $(r_n + v_n) = O(n)$ , and
- (i)  $v_n = o\left((n\phi_x(h_K))^{\frac{1}{2}}\right)$  and  $\lim_{n \rightarrow +\infty} \left(\frac{n}{\phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) = 0$ ,
  - (ii)  $q_n v_n = o\left((n\phi_x(h_K))^{\frac{1}{2}}\right)$  and  $\lim_{n \rightarrow +\infty} q_n \left(\frac{n}{\phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) = 0$ , where  $q_n$  is the largest integer, such that  $q_n(r_n + v_n) = O(n)$ .

### Comments on the assumptions

It is observed that the assumptions listed above are standard in the FDA context. In particular, hypotheses (H1) and (H6) are not unduly restrictive and are common in the setting of the functional local linear fitting (see Barrientos-Marina et al. [2], and Demongeot et al. [12] among others). Concerning the first part of (H1), the reader will find, in Ferraty and Vieu [19] a deep discussion concerning the links between this assumption, the semi-metric  $d$  and the small ball concentration properties. Moreover, this hypothesis intervenes to compute the exact constant terms involved in the asymptotic expansions. For example, the previous hypothesis is proposed to evaluate the constant  $M_j = \mathbb{E}(K_1^j)$ , where  $j \in (1, 2)$ . However, the second part of (H1) is needed to evaluate the bias of estimation in the asymptotic result. To avoid the expression of covariance in the rate of convergence, assumptions (H3) and (H4) are required ; in addition, hypothesis (H3) can be differently seen based on the idea of the maximum concentration between the quantities  $P(X_i \in B(x, h_K))$  and  $P(X_j \in B(x, h_K))$  (see Ferraty et al. [18]). Concerning the hypothesis (H4), it is used to insure the absolute convergence of the series  $\sum_{k \in \mathbb{Z}} Cov(X_0, X_k)$ . Conditions on the smoothing parameters  $h_K$ , and  $h_H$  are standard and will be stated along the theorem below. The boundedness of the Kernel  $K$  in (H7)(i) is standard ; also assumptions (H7)(ii) and (H8) constitute technical conditions for brevity proofs. Furthermore, the role of assumption (H8) is to use Bernstein's big-block and small-block techniques to prove the asymptotic normality for the  $\alpha$ -mixing sequence ; nonetheless, the choice of the sequences  $(r_n)$  and  $(v_n)$  in hypothesis (H8) is not surprising. Another choice can be found in Masry [31].

### 3.3 Main Results

Before announcing the main results, the quantities  $M_j$  and  $N(a, b)$  are introduced to provide bias and variance dominant terms.

$$M_j = K^j(1) - \int_{-1}^1 (K^j(u))' \Psi_x(u) du, \text{ where } j = 1, 2,$$

$$N(a, b) = K^a(1) - \int_{-1}^1 (u^b K^a(u))' \Psi_x(u) du, \text{ For all } a > 0, \text{ and } b = 2, 4.$$

Moreover, let  $\widehat{F}^x(y) = \frac{\widehat{F}_N^x(y)}{\widehat{F}_D^x}$ , where  $\widehat{F}_N^x(y) = \frac{1}{n \mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_j K_j H_j$

and  $\widehat{F}_D^x = \frac{1}{n \mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \Delta_j K_j$ .

**Theorem 11** *Under assumptions (H1)-(H8), we have*

$$\sqrt{n\phi_x(h_K)}(\widehat{F}^x(y) - F^x(y) - B_n(x, y)) \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)), \quad (3.3)$$

where

$$V_{HK}(x, y) = \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)), \quad (3.4)$$

and

$$B_n(x, y) = \mathbb{E}(\widehat{F}^x(y)) - F^x(y) = B_H(x, y)h_H^2 + B_K(x, y)h_K^2 + o(h_K^2) + o(h_K^2), \quad (3.5)$$

where

$$B_H(x, y) = \frac{1}{2} \frac{\partial^2 F^x(y)}{\partial y^2} \int t^2 H'(t) dt,$$

$$B_K(x, y) = \frac{1}{2} G_0^{(2)}(0) \frac{N(1, 2)}{M_1}.$$

**Remark 12** *I If we impose the additional assumption :*

$$(H9) \lim_{n \rightarrow \infty} \sqrt{n\phi_x(h_K)} B_n(x, y) = 0,$$

and, in addition, if we replace the function  $\phi_x(h_K)$  by its empirical estimator defined by the following :

$$\widehat{\phi}_x(h_K) = \frac{\#\{i : |\delta(X_i, x)| \leq h_K\}}{n},$$

the bias term can be canceled to obtain the following Corollary :



**Corollary 13** *When the assumptions (H1)-(H9) are held, the following asymptotic result is achieved :*

$$\sqrt{\frac{n\hat{\phi}_x(h_K)}{\widehat{V}_{HK}(x,y)}}(\widehat{F}^x(y) - F^x(y)) \xrightarrow{d} \mathcal{N}(0,1).$$

### 3.3.1 Proof of Theorem 11

Starting by

$$\widehat{F}^x(y) - F^x(y) - B_n(x,y) = \frac{\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - \widehat{F}_D^x B_n(x,y)}{\widehat{F}_D^x}.$$

Denote by

$$\begin{aligned} Q_n(x,y) &= \widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - \mathbb{E}(\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x) \\ &= \widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x - B_n(x,y), \end{aligned} \quad (3.6)$$

then

$$\widehat{F}^x(y) - F^x(y) - B_n(x,y) = \frac{Q_n(x,y) - B_n(x,y)(\widehat{F}_D^x - \mathbb{E}(\widehat{F}_D^x))}{\widehat{F}_D^x}. \quad (3.7)$$

The relationship (3.7) is important to establish the asymptotic normality of  $\widehat{F}^x(y)$ ; moreover, the continuity of  $F^x$  insures the asymptotic negligibility of  $B_n(x,y)$  and if  $\widehat{F}_D^x$  converges in probability to 1 as  $n \rightarrow \infty$ , then

$$\widehat{F}^x(y) - F^x(y) - B_n(x,y) = \frac{Q_n(x,y)}{\widehat{F}_D^x}(1 + o_p(1)),$$

is obtained. The proof of Theorem 11 will be completed from the above expression and the following results for which proofs are given in the appendix.

**Lemma 3.3.1** *Under the assumptions of Theorem (11), we have*

$$\sqrt{n\hat{\phi}_x(h_K)}Q_n(x,y) \xrightarrow{d} \mathcal{N}(0, V_{HK}(x,y)), \text{ as } n \rightarrow \infty,$$

where  $V_{HK}(x,y)$  is defined in (3.4).

**Lemma 3.3.2** Under the assumptions (H3), (H5), we have

$$\widehat{F}_D^x \xrightarrow{P} \mathbb{E}(\widehat{F}_D^x) = 1, \text{ as } n \rightarrow \infty.$$

**Lemma 3.3.3** (see[12]) Under the assumptions (H1), (H5) and (H7), we have the following :

$$B_n(x, y) = B_H(x, y)h_H^2 + B_K(x, y)h_K^2 + o(h_K^2) + o(h_H^2), \text{ as } n \rightarrow \infty.$$

### 3.4 Confidence Intervals

In parallel, the precise form of (3.3) is very useful to construct confidence intervals for  $F^x(y)$  based on the normal-approximation method that requires to estimate the quantities  $\widehat{M}_1$ , and  $\widehat{M}_2$  by the following empirical estimators :

$$\widehat{M}_j = \frac{1}{n\widehat{\phi}_x(h_K)} \sum_{i=1}^n K^j \left( \frac{|\delta(X_i, x)|}{h_K} \right), \quad \text{where } j = 1, 2.$$

To show the asymptotic  $(1 - \xi)$  confidence interval of  $F^x(y)$ , where  $0 < \xi < 1$ , it is necessary to consider the estimator of  $\widehat{V}_{HK}(x, y)$  as follows

$$\widehat{V}_{HK}(x, y) = \frac{\sum_{i=1}^n K_i^2 \widehat{F}^x(y)(1 - \widehat{F}^x(y))}{\widehat{\phi}_x(h_K)(\sum_{i=1}^n K_i)^2}.$$

In addition, a kernel  $K$  and a distribution function  $H$  are chosen satisfy the condition (H7) by selecting the bandwidths  $h_K$  and  $h_H$  through adapting the cross-validation method. Choosing the locating functions  $\beta(., .)$  and  $\delta(., .)$  constitutes an important parameter for the practical utilization of the employed approach. There are several ways permitting to choose the operators  $\beta(., .)$  and  $\delta(., .)$  (see Barrientos-Marina-Marina et al. [2] for some examples), but the appropriate choice is determined with respect to the shape of the curves and depends on the purpose of the statistical study. For example, if the functional data are smooth curves, one can try to use the following family of locating functions :

$$\beta(x_1, x_2) = \int_0^1 \theta(t)(x_1^{(q)}(t) - x_2^{(q)}(t))dt \quad \text{and} \quad \delta(x_1, x_2) = \sqrt{\int_0^1 (x_1^{(q)}(t) - x_2^{(q)}(t))^2 dt},$$

where  $x^{(q)}$  denotes the  $q$ th derivative of the curve  $x$  and  $\theta(t)$  is the eigenfunction of the empirical covariance operator  $\frac{1}{n} \sum_{i=1}^n (X_i^{(q)} - \overline{X^{(q)}})^t (X_i^{(q)} - \overline{X^{(q)}})$  associated with the  $q$ -greatest eigenvalue.

Finally, by Corollary (13), the asymptotic  $(1 - \xi)$  confidence interval of  $F^x(y)$  is given by the following

$$\left[ \widehat{F}^x(y) - \lambda_{\frac{\xi}{2}} \sqrt{\frac{\widehat{V}_{HK}(x, y)}{n\widehat{\phi}_x(h_K)}}, \widehat{F}^x(y) + \lambda_{\frac{\xi}{2}} \sqrt{\frac{\widehat{V}_{HK}(x, y)}{n\widehat{\phi}_x(h_K)}} \right],$$

where  $\lambda_{\frac{\xi}{2}}$  is the  $\frac{\xi}{2}$  quantile of the standard normal distribution.

## 3.5 Appendix

### 3.5.1 Proof of Lemma 3.3.1.

Let us first note that, in view of (3.6), we have

$$\sqrt{n\phi_x(h_K)}Q_n(x, y) = \frac{\sqrt{n\phi_x(h_K)}}{n\mathbb{E}(\Delta_1 K_1)} \left( \sum_{j=1}^n \Delta_j K_j (H_j - F^x(y)) - \mathbb{E} \left( \sum_{j=1}^n \Delta_j K_j (H_j - F^x(y)) \right) \right).$$

In the same way, using the definition of  $\Delta_j$ , this equality can be rewritten as follows

$$\begin{aligned} \sqrt{n\phi_x(h_K)}Q_n(x, y) &= \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)) \\ &\quad - \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)) \\ &\quad - \mathbb{E} \left( \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)) \right) \\ &\quad + \mathbb{E} \left( \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)) \right). \end{aligned}$$

Denote by

$$\begin{aligned} T_{1,i} &= \frac{1}{n\mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i, & T_{2,j} &= \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j (H_j - F^x(y)), \\ T_{3,i} &= \frac{1}{n\mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i & \text{and } T_{4,j} &= \frac{\sqrt{n\phi_x(h_K)}\mathbb{E}(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j (H_j - F^x(y)), \end{aligned}$$

and thus

$$\begin{aligned}\sqrt{n\phi_x(h_K)}Q_n(x, y) &= T_{1,i}T_{2,j} - T_{3,i}T_{4,j} - \mathbb{E}(T_{1,i}T_{2,j} - T_{3,i}T_{4,j}) \\ &= T_{1,i}T_{2,j} - \mathbb{E}(T_{1,i}T_{2,j}) - (T_{3,i}T_{4,j} - \mathbb{E}(T_{3,i}T_{4,j})).\end{aligned}$$

Finally, the rest of the proof is based on the following statements :

$$T_{1,i}T_{2,j} - \mathbb{E}(T_{1,i}T_{2,j}) \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)). \quad (3.8)$$

$$T_{3,i}T_{4,j} - \mathbb{E}(T_{3,i}T_{4,j}) \xrightarrow{P} 0. \quad (3.9)$$

**Proof of (3.8).**

Let us write the left-hand side of (3.8) as follows :

$$T_{1,i}T_{2,j} - \mathbb{E}(T_{1,i}T_{2,j}) = T_{2,j} - \mathbb{E}(T_{2,j}) + ((T_{1,i} - 1)T_{2,j} - \mathbb{E}((T_{1,i} - 1)T_{2,j})).$$

Hence, by Slutsky's Theorem (see Theorem 11.1.5 in [25]), (3.8) is straightforward consequence of the following two claims :

$$T_{2,j} - \mathbb{E}(T_{2,j}) \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)). \quad (3.10)$$

$$(T_{1,i} - 1)T_{2,j} - \mathbb{E}((T_{1,i} - 1)T_{2,j}) \xrightarrow{P} 0. \quad (3.11)$$

**Proof of (3.10).**

As a matter of fact, we need to evaluate the variance of  $(T_{2,j})$ . For this, we have the following :

$$\text{Var}(T_{2,j}) = n \text{Var}(T_{2,1}) + 2 \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \sum_{1 \leq i < j \leq n} \text{cov}(K_i(H_i - F^x(y)), K_j(H_j - F^x(y))).$$

Therefore, to prove that  $\lim_{n \rightarrow +\infty} \text{Var}(T_{2,j}) = V_{HK}(x, y)$ , it is necessary to establish the following results :

$$\lim_{n \rightarrow +\infty} n \text{Var}(T_{2,1}) = \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)). \quad (3.12)$$

$$\lim_{n \rightarrow +\infty} \left( \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \sum_{1 \leq i < j \leq n} \text{cov}(K_i(H_i - F^x(y)), K_j(H_j - F^x(y))) \right) = 0. \quad (3.13)$$

**Proof of (3.12).** One has

$$n \operatorname{Var}(T_{2,1}) = \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \left( \mathbb{E}(K_1^2 (H_1 - F^x(y))^2) - \mathbb{E}^2(K_1 (H_1 - F^x(y))) \right). \quad (3.14)$$

Concerning the second term on the right-hand side of (3.14), we have the following :

$$\mathbb{E}^2(K_1 (H_1 - F^x(y))) = \mathbb{E}^2(K_1 (\mathbb{E}(H_1/X_1) - F^x(y))),$$

and by the continuity of  $F^x$  we deduce that :

$$\mathbb{E}(H_1/X_1) - F^x(y) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty. \quad (3.15)$$

Now, we turn to the first term on the right-hand side of (3.14). Let us begin with writing :

$$\begin{aligned} \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E} \left( (H_1 - F^x(y))^2 K_1^2 \right) &= \frac{n^2 \phi_x(h) \mathbb{E}^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} \mathbb{E} \left( \mathbb{E}((H_1 - F^x(y))^2 / X_1) K_1^2 \right) \\ &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E} \left( \operatorname{Var}(H_1/X_1) K_1^2 \right) \\ &+ \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} \\ &\times \mathbb{E} \left( (E(H_1/X_1) - F^x(y))^2 K_1^2 \right). \end{aligned} \quad (3.16)$$

In view of (3.15), classical computations of the second term on the right-hand side of (3.16) give :

$$\frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} \mathbb{E} \left( (E(H_1/X_1) - F^x(y))^2 K_1^2 \right) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

Concerning the first term on the right-hand side of (3.16), we use the following definition of the conditional variance :

$$\operatorname{Var}(H_1/X_1) = \mathbb{E}(H_1^2/X_1) - \mathbb{E}^2(H_1/X_1). \quad (3.17)$$

Thus, using an integration by parts followed by a change of variable, we get :

$$\mathbb{E} \left( H^2 \left( \frac{y - Y_1}{h_H} \right) / X_1 \right) = \int_{\mathbb{R}} 2H'(t)H(t)(F^{(X_1)}(y - th_H) - F^x(y))dt + \int_{\mathbb{R}} 2H(t)'H(t)F^x(y)dt.$$

Remark that  $\int_{\mathbb{R}} 2H'(t)H(t)F^x(y)dt = F^x(y)$ , and by the continuity of  $F^x$  we deduce that

$$\mathbb{E}\left(H^2\left(\frac{y - Y_1}{h_H}\right)/X_1\right) \longrightarrow F^x(y) \quad \text{as } n \longrightarrow \infty.$$

Therefore, the second term on the right-hand side of (3.17) tends to  $(F^x(y))^2$  as  $n$  tends to infinity.

Finally, we have the following :

$$\mathbb{E}\left(K_1^2 \text{Var}\left(H\left(\frac{y - Y_1}{h_H}\right)/X_1\right)\right) \longrightarrow \mathbb{E}(K_1^2)F^x(y)(1 - F^x(y)) \quad \text{as } n \longrightarrow \infty.$$

Next, by using Lemma A.1 of [37], we get :

$$\begin{aligned} n \text{Var}(T_{2,1}) &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \mathbb{E}\left((H_1 - F^x(y))^2 K_1^2\right) \\ &= \frac{n^2 M_2}{(n-1)^2 M_1^2} F^x(y)(1 - F^x(y)) \longrightarrow \frac{M_2}{M_1^2} F^x(y)(1 - F^x(y)) \\ &= V_{HK}(x, y) \quad \text{as } n \longrightarrow \infty. \end{aligned}$$

**Proof of (3.13).**

First, the set  $E_1$  and  $E_2$  are defined by setting

$$E_1 = \{(i, j) \in \{1, 2, \dots, n\} \text{ such that } 1 \leq |i - j| \leq m_n\},$$

$$E_2 = \{(i, j) \in \{1, 2, \dots, n\} \text{ such that } m_n + 1 \leq |i - j| \leq n - 1\},$$

where  $m_n$  is a sequence of integers, such that  $m_n \longrightarrow +\infty$  as  $n \rightarrow +\infty$ .

Now, denote by

$$L_i = K_i(H_i - F^x(y)) \quad \text{and} \quad L_j = K_j(H_j - F^x(y)),$$

then

$$\begin{aligned} \text{Cov}(T_{1,i}, T_{1,j}) &= \frac{n \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \sum_{E_1} \text{Cov}(L_i, L_j) \\ &\quad + \frac{n \phi_x(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} \sum_{E_2} \text{Cov}(L_i, L_j) \\ &=: A_{1,n} + A_{2,n}. \end{aligned}$$

Having the sum of covariance over the set  $E_1$ , by stationarity :

$$\begin{aligned} \text{Cov}(L_i, L_j) &= \mathbb{E}(K_i K_j \mathbb{E}((H_i - F^x(y))(H_j - F^x(y))/(X_i, Y_j))) \quad (3.18) \\ &\quad - \mathbb{E}^2(K_1(H_1 - F^x(y))). \end{aligned}$$

Under (H7)(ii), Having  $|H_i(y) - F^x(y)| \leq 1$ , according to (H3), the following inequality is obtained as follows :

$$\begin{aligned} |Cov(L_i, L_j)| &\leq \mathbb{E}(K_i K_j) + \mathbb{E}^2(K_1), \\ &\leq C\mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) + M_1^2 \phi_x^2(h_K), \\ &\leq C\psi_x(h_K) + M_1^2 \phi_x^2(h_K). \end{aligned}$$

Now, by the application of Lemma A.1 [37], we have the following :

$$|A_{1,n}| \leq \frac{Cn^2 \phi_x(h_K) m_n}{(n-1)^2 M_1^2} \left( \left( \frac{\psi_x(h_K)}{\phi_x(h_K)^2} \right) + M_1^2 \right),$$

It follows that, by (H5)(iii) and taking  $m_n = \left( \frac{1}{(\phi_x(h_K) \log n)^{\frac{1}{p}}} \right)^{\frac{1}{\delta}}$  we obtain  $A_{1,n} = o(1)$  as  $n \rightarrow +\infty$ .

Concerning the sum over the set  $E_2$ , the proposition A.10 (ii) of [19] is used to get :

$$\sum_{E_2} |Cov(L_i, L_j)| \leq \sum_{E_2} C(\alpha|j-i|)^{\frac{1}{p}} (\mathbb{E}|L_i|^q)^{\frac{1}{q}} (\mathbb{E}|L_j|^r)^{\frac{1}{r}}.$$

First, we evaluate the quantity  $\mathbb{E}(|L_i|^q)$ . Conditioning on  $X_i$ , and using the fact that  $|H_i - F^x(y)| \leq 1$ , we obtain :

$$\begin{aligned} \mathbb{E}(|L_i|^q) &= \mathbb{E}(|K_i|^q \mathbb{E}(|H_i - F^x(y)|^q / X_1)), \\ &\leq C(\phi_x(h_K))^{\frac{1}{q}}. \end{aligned}$$

Again, using Lemma A.1 of [37], we have :

$$|A_{2,n}| \leq \frac{Cn^2 (\phi_x(h_K))^{\frac{1}{q} + \frac{1}{r} = 1 - \frac{1}{p}}}{(n-1)^2 M_1^2 \phi_x(h_K) (m_n)^\delta} \sum_{|k| > m_n} k^\delta (\alpha(|k|))^{\frac{1}{p}}.$$

Finally, the obtained result is combined with assumptions (H4)(ii) and the sequence  $m_n$  previously chosen to get :  $A_{2,n} = o(1)$ . Now, the asymptotic normality of the conditional cumulative distribution estimation is established dealing with dependent random variables :

$$\begin{aligned} T_{2,j} - \mathbb{E}(T_{2,j}) &= \frac{\sqrt{n\phi_x(h_K)} \mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n (K_j(H_j - F^x(y)) - \mathbb{E}(K_j(H_j - F^x(y)))) \\ &= \frac{\sum_{j=1}^n (Z_j - \mathbb{E}(Z_j))}{\sqrt{n}} =: S_n, \end{aligned}$$

where

$$Z_j = \frac{n\sqrt{\phi_x(h_K)}\mathbb{E}(\beta_1^2 K_1)}{\mathbb{E}(\Delta_1 K_1)} (K_j(H_j - F^x(y)) - \mathbb{E}(K_j(H_j - F^x(y)))).$$

Remark that (3.13) implies that :

$$\sum_{1 \leq i < j \leq n} \text{Cov}(Z_i, Z_j) = o(n).$$

Therefore, it suffices to show the following result :

$$S_n \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)). \quad (3.19)$$

Bernstein's big-block and small-block procedure is employed following similar arguments to those involved in Theorem 3.1 of Liang and Baek [29].  $(1, 2, \dots, n)$  is splitted into  $2\kappa_n + 1$  subsets with large blocks of size  $(r_n)$  and small blocks of size  $(v_n)$  and by putting

$$\kappa = \left\lfloor \frac{n}{r_n + v_n} \right\rfloor.$$

Assumption (H8)(ii) permits to define the large block size as follows :

$$r_n = \left\lfloor \left( \frac{n\phi_x(h_K)}{q_n} \right)^{\frac{1}{2}} \right\rfloor.$$

Moreover, some easy computations are obtained using the same hypothesis :

$$\lim_{n \rightarrow +\infty} \frac{v_n}{r_n} = 0, \quad \lim_{n \rightarrow +\infty} \frac{r_n}{n} = 0, \quad \lim_{n \rightarrow +\infty} \frac{r_n}{\sqrt{n\phi_x(h_K)}} = 0, \quad (3.20)$$

and it can easily be deduced that, as  $n \rightarrow +\infty$

$$\frac{\kappa v_n}{n} \simeq \left( \frac{n}{r_n + v_n} \right) \frac{v_n}{n} \simeq \frac{v_n}{r_n + v_n} \simeq \frac{v_n}{r_n} = 0.$$

In addition, if  $v_n$  is replaced by  $r_n$ , we obtain

$$\lim_{n \rightarrow +\infty} \frac{\kappa r_n}{n} = 1.$$

The sum  $S_n$  is then splitted in the following way :

The set  $I_j = \{j(r+v)+1, \dots, j(r+v)+r\}$  contains  $r$  elements for each  $j \in (0, 1, 2, \dots, k-$



1), and the set

$l_j = \{j(r+v)+r+1, \dots, (j+1)(r+v)\}$  contains  $v$  elements for each  $j \in (0, 1, 2, \dots, k-1)$ .

The following random variables are defined by the following :

$$\Upsilon_j = \sum_{i=j(r+v)+1}^{j(r+v)+r} Z_i, \quad \tilde{\Upsilon}_j = \sum_{i=j(r+v)+r+1}^{(j+1)(r+v)} Z_i, \quad R_n = \sum_{i=\kappa(r+v)+1}^n Z_i. \quad (3.21)$$

It is clear that :

$$\begin{aligned} S_n &= \sum_{j=0}^{\kappa-1} \frac{\Upsilon_j}{\sqrt{n}} + \sum_{j=0}^{\kappa-1} \frac{\tilde{\Upsilon}_j}{\sqrt{n}} + \frac{R_n}{\sqrt{n}} \\ &=: S_{1,n} + S_{2,n} + S_{3,n}. \end{aligned}$$

Then, (3.19) will be obtained as soon as the following assertions are checked :

$$[S_{2,n} + S_{3,n}] \xrightarrow{p} 0, \quad (3.22)$$

and

$$S_{1,n} \xrightarrow{d} \mathcal{N}(0, V_{HK}(x, y)). \quad (3.23)$$

**Proof of (3.22).**

By Markov's Inequality, it remains to establish for all  $\varepsilon > 0$

$$\mathbf{P}(|S_{2,n}| > \varepsilon) \leq \frac{\mathbb{E}((S_{2,n})^2)}{\varepsilon^2}, \quad (3.24)$$

$$\mathbf{P}(|S_{3,n}| > \varepsilon) \leq \frac{\mathbb{E}((S_{3,n})^2)}{\varepsilon^2}. \quad (3.25)$$

To prove (3.24), it is clearly observed that :

$$\begin{aligned} \mathbb{E}((S_{2,n})^2) &= \frac{1}{n} \left( \sum_{j=0}^{k-1} \text{Var}(\tilde{\Upsilon}_j) + 2 \sum_{0 \leq i < j \leq k-1} \text{Cov}(\tilde{\Upsilon}_i, \tilde{\Upsilon}_j) \right) \\ &=: A_1 + A_2. \end{aligned}$$

Noting also that, by the second-order stationarity, it will be retained :

$$\begin{aligned} \text{Var}(\tilde{\Upsilon}_j) &= \text{Var} \left( \sum_{i=j(r+v)+r+1}^{(j+1)(r+v)} Z_i \right) \\ &= v_n \text{Var}(Z_1) + 2 \sum_{\substack{i \neq j \\ i \neq j}}^{v_n} \text{Cov}(Z_i, Z_j). \end{aligned}$$

Consequently :

$$\begin{aligned} \frac{A_1}{n} &= \frac{\kappa v_n}{n} \text{Var}(Z_1) + \frac{2}{n} \sum_{j=0}^{k-1} \sum_{i \neq j}^{v_n} \text{Cov}(Z_i, Z_j) \\ &\leq \frac{\kappa v_n}{n} \underbrace{\text{Var}(Z_1)}_{V_{HK}(x,y)} + \frac{2}{n} \underbrace{\sum_{i \neq j}^n \text{Cov}(Z_i, Z_j)}_{o(n)}. \end{aligned}$$

By the assumption (H8), it is clear that :

$$\lim_{n \rightarrow +\infty} \frac{A_1}{n} \rightarrow 0.$$

Concerning  $A_2$ , it has :

$$\begin{aligned} \frac{A_2}{n} &= \frac{2}{n} \sum_{|i-j|>0}^{k-1} \text{Cov}(\tilde{Y}_i, \tilde{Y}_j) \\ &= \frac{2}{n} \sum_{|i-j|>0}^{k-1} \sum_{l=1}^{v_n} \sum_{\acute{l}=1}^{v_n} \text{Cov}(Z_{\varsigma_i+l}, Z_{\varsigma_j+\acute{l}}), \end{aligned}$$

with  $\varsigma_i = i(r_n + v_n) + r_n$ . Since  $i \neq j$ , we have  $|\varsigma_i - \varsigma_j + l - \acute{l}| \geq r_n$ . It follows that

$$\frac{A_2}{n} \leq \frac{2}{n} \sum_{i=1}^n \sum_{\substack{j=1 \\ |i-j| \geq r_n}}^n |\text{Cov}(Z_i, Z_j)|,$$

which leads to

$$\lim_{n \rightarrow +\infty} \frac{A_2}{n} \rightarrow 0.$$

For (3.25), we have the following :

$$\begin{aligned} \mathbb{E}((S_{3,n})^2) &= \frac{1}{n} \text{Var}(R_n) \\ &= \frac{\mu_n}{n} \text{Var}(Z_1) + \frac{2}{n} \underbrace{\sum_{i \neq j}^{\mu_n} \text{Cov}(Z_i, Z_j)}_{o(n)}, \end{aligned}$$

where  $\mu_n = n - \kappa_n(r_n + v_n)$ , and by the definition of  $\kappa_n$  we have  $\mu_n < r_n + v_n$ . Hence

$$\mathbb{E}(S_{3,n}^2) \leq \frac{r_n + v_n}{n} \underbrace{\text{Var}(Z_1)}_{V_{HK}(x,y)} + o(1).$$

And, again hypothesis (H8), we get

$$\mathbb{E}((S_{3,n})^2) \longrightarrow 0 \quad \text{as } n \longrightarrow +\infty.$$

**Proof of (3.23).**

Making use of Volkonskii and Rozanov's Lemma [36] and the fact that the process  $(X_i, Y_i)$  is strong mixing; and that  $\Upsilon_a$  is  $\mathcal{A}_{i_a}^{j_a}$  measurable with  $i_a = a(r_n + v_n) + 1$  and  $j_a = a(r_n + v_n) + r_n$ ; hence with  $V_j = \exp\left(\frac{it\Upsilon_j}{\sqrt{n}}\right)$  we have the following :

$$\left| \mathbb{E} \left( \exp \left( it \frac{S_{1,n}}{\sqrt{n}} \right) \right) - \prod_{j=0}^{\kappa-1} \mathbb{E} \left( \exp \left( it \frac{\Upsilon_j}{\sqrt{n}} \right) \right) \right| \longrightarrow 0 \quad \text{as } n \rightarrow +\infty. \quad (3.26)$$

Consequently, according to formula (3.26),  $\Upsilon_j$  are asymptotically independent. Therefore, for the variance of  $S_{1,n}$  we have the following :

$$\text{Var}(S_{1,n}) = \frac{\kappa r_n}{n} \underbrace{\text{Var}(Z_1)}_{V_{HK}(x,y)}.$$

Furthermore, from assumption (H8),  $\frac{\kappa r_n}{n} \longrightarrow 1$  as  $n \longrightarrow +\infty$ .

Finally, we get

$$\frac{1}{n} \sum_{j=0}^{\kappa-1} \mathbb{E}[\Upsilon_j^2] \longrightarrow V_{HK}(x, y) \quad \text{as } n \rightarrow +\infty.$$

Now, to end the proof of (3.23), we focus on the central limit Theorem due to Linderberg. More precisely, by applying the Linderberg's version of central limit Theorem on  $\Upsilon_j$ , it suffices to show that for any  $\varepsilon > 0$ ,

$$\frac{1}{n} \sum_{j=0}^{\kappa-1} \mathbb{E} \left[ \Upsilon_j^2 \mathbb{1}_{|\Upsilon_j| > \varepsilon \sqrt{n V_{HK}(x,y)}} \right] \longrightarrow 0 \quad \text{as } n \longrightarrow +\infty.$$

In view of the first summation of (3.21), classical computations give

$$\left| \frac{\Upsilon_j}{n} \right| \leq \frac{r_n}{n} |Z_1|.$$

Next, the application of Lemma A.1 of [37] together with (H5), leads to  $|Z_1| \longrightarrow 0$  as  $n \longrightarrow +\infty$ , and noting that this last result combined with (3.20) ensures that  $\left| \frac{\Upsilon_j}{n} \right| \longrightarrow 0$

as  $n \rightarrow +\infty$ . Therefore, for all  $\varepsilon$  and if  $n$  is great enough, the set  $\{|\Upsilon_j| > \varepsilon \sqrt{nV_{HK}(x, y)}\}$  becomes empty; and the proof of (3.23) is therefore, complete.

**Proof of (3.11).**

By Bienaymé–Tchebychev’s Inequality, it is sufficient to show that, for all  $\varepsilon > 0$

$$\mathbb{P}(|(T_{1,i} - 1)T_{2,j} - \mathbb{E}((T_{1,i} - 1)T_{2,j})| \geq \varepsilon) \leq \frac{\mathbb{E} |(T_{1,i} - 1)T_{2,j} - \mathbb{E}((T_{1,i} - 1)T_{2,j})|}{\varepsilon}.$$

In addition, Cauchy–Schwarz’s Inequality entails

$$\mathbb{E} |(T_{1,i} - 1)T_{2,j} - \mathbb{E}((T_{1,i} - 1)T_{2,j})| \leq 2\mathbb{E} |(T_{1,i} - 1)T_{2,j}| \leq \sqrt{E((T_{1,i} - 1)^2)} \sqrt{\mathbb{E}(T_{2,j}^2)}.$$

Then, (3.11) is a straightforward consequence of the following results :

$$\mathbb{E}((T_{1,i} - 1)^2) \rightarrow 0, \quad \text{as } n \rightarrow +\infty, \quad (3.27)$$

$$\mathbb{E}(T_{2,j}^2) \rightarrow V_{HK}(x, y), \quad \text{as } n \rightarrow +\infty. \quad (3.28)$$

**Proof of (3.27).**

First, we can write

$$E((T_{1,i} - 1)^2) = n \text{Var}(T_{1,1}) + 2 \text{Cov}(T_{1,i}, T_{1,l})$$

For the first term on the right–hand side of this equality, we have the following :

$$\begin{aligned} n \text{Var}(T_{1,1}) &= \frac{n \text{Var}(\beta_1^2 K_1)}{n^2 \mathbb{E}(\beta_1^2 K_1)} \\ &= O\left(\frac{1}{n\phi_x(h_K)}\right). \end{aligned}$$

Concerning the second term of the previous equality, we have the following :

$$\text{Cov}(T_{1,i}, T_{1,l}) = \frac{1}{n^2 \mathbb{E}^2(\beta_1^2 K_1)} \sum_{\substack{i=1 \\ i \neq l}}^n \sum_{l=1}^n \text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l). \quad (3.29)$$

The proof of this result is very close to the proof of (3.13). Specifically, by keeping the same notations as those used in (3.13), and by splitting the sum into two separate summations over the sets  $E_1$  and  $E_2$  :

$$E_1 = \{(i, l) \in \{1, 2, \dots, n\} \text{ such that } 1 \leq |i - l| \leq m_n\},$$

$$E_2 = \{(i, l) \in \{1, 2, \dots, n\} \text{ such that } m_n + 1 \leq |i - l| \leq n - 1\},$$

where the sequence  $m_n$  is chosen, such that  $m_n \rightarrow +\infty$  as  $n \rightarrow +\infty$ . Denoting now by  $A_{1,n}$  and  $A_{2,n}$  the sum of covariances over  $E_1$  and  $E_2$  respectively, so

$$A_{1,n} = \frac{1}{n^2 \mathbb{E}^2(\beta_1^2 K_1)} \sum_{E_1} \text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l),$$

$$A_{2,n} = \frac{1}{n^2 \mathbb{E}^2(\beta_1^2 K_1)} \sum_{E_2} \text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l).$$

By stationarity, we have :

$$\text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l) = \mathbb{E}(\beta_i^2 K_i \beta_l^2 K_l) - \mathbb{E}^2(\beta_1^2 K_1),$$

then

$$|\text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l)| \leq |\mathbb{E}(\beta_i^2 K_i \beta_l^2 K_l)| + \mathbb{E}^2(\beta_1^2 K_1).$$

Moreover, assumptions (H1) and (H6) imply that :

$$K_i |\beta_i|^2 h_K^{-2} \leq C K_i |\delta(x, X_i)|^2 h_K^{-2} \leq C K_i.$$

On the other hand, we may apply Jensen's Inequality and assumption (H3) to obtain :

$$\begin{aligned} |\mathbb{E}(\beta_i^2 K_i \beta_l^2 K_l)| &\leq \mathbb{E}|\beta_i^2 K_i \beta_l^2 K_l|, \\ &\leq C h_K^4 \mathbb{E}(K_i K_l), \\ &\leq C h_K^4 \psi_x(h_K). \end{aligned}$$

In the next step, we use the technical Lemma A.1 of [2] to get :

$$\begin{aligned} |A_{1,n}| &\leq n m_n (C h_K^4 \psi_x(h_K) + \mathbb{E}^2(\beta_1^2 K_1)) \frac{1}{n^2 \mathbb{E}^2(\beta_1^2 K_1)}, \\ &\leq \frac{C m_n}{n O(1)} \frac{\psi_x(h_K)}{\phi_x^2(h_K)} + \frac{m_n}{n}. \end{aligned}$$

Then, choosing  $m_n = \sqrt{n}$  and since  $\frac{\psi_x(h_K)}{\phi_x^2(h_K)}$  is bounded from assumption (H3), we arrive at :

$$A_{1,n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Let us now treat the sum over  $E_2$ . The application of the inequality for bounded mixing processes [see Proposition A.10(i) in [19]], for all  $l \neq i$  leads to :

$$|\text{Cov}(\beta_i^2 K_i, \beta_l^2 K_l)| \leq C h_K^4 \alpha(|i - l|).$$

On the other hand, using the fact that  $\sum_{j \geq x+1} j^{-s} \leq \int_{u \geq x} u^{-s} = \left( (1-s)u^{s-1} \right)^{-1}$ , and

under (H4)(i), it is easy to get :

$$\sum_{i=1}^n \sum_{m_n+1 \leq |i-l| \leq n-1} \alpha(|i-l|) \leq \frac{n(m_n)^{1-s}}{s-1}.$$

Finally, we have :

$$|A_{2,n}| \leq \frac{C(m_n)^{1-s}}{s-1} \frac{1}{O(n\phi_x^2(h_K))}.$$

We use the same choice of  $m_n$  as before, and using assumption (H5)(i), we obtain :

$$A_{2,n} \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

**Proof of (3.28).**

We start by writing

$$\mathbb{E}((T_{2,j})^2) = \text{Var}(T_{2,j}) + \mathbb{E}^2(T_{2,j}). \quad (3.30)$$

The first term on the right-hand side of (3.30) tends to  $V_{HK}(x, y)$  as  $n$  tends to infinity, and the proof of this result was shown in (3.13). Concerning the second term on the right-hand side of (3.30), we have have the following :

$$\mathbb{E}^2(T_{2,j}) = \frac{n^3 \phi_x(h_K) E^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} E^2(K_1(H_1 - F^x(y))),$$

where

$$\mathbb{E}[H_1/X_1] = \int_R H(h_H^{-1}(y-z)) f^{X_1}(z) dz.$$

with an integration by parts followed by the change of variables  $t = \frac{y-z}{h_H}$ , allows to write :

$$\mathbb{E}(H_1/X_1) = \int_R H'(t) F^{X_1}(y - h_H t) dt.$$

Moreover, the latter integral can be rewritten as follows :

$$\mathbb{E}(H_1/X_1) = \int_R H'(t) \left( F^{X_1}(y - h_H t) - F^x(y) \right) + \int_R H'(t) F^x(y) dt.$$

Now, under assumption (H7)(ii), and by the continuity of  $F^x$  we have the following :

$$\mathbb{E}(H_1/X_1) - F^x(y) \longrightarrow 0 \text{ as } n \longrightarrow \infty. \quad (3.31)$$

In addition, by applying the technical Lemma A.1 in [37], we get  $\lim_{n \rightarrow +\infty} \frac{E^2(\beta_1^2 K_1)}{E^2(\Delta_1 K_1)} = 0$ .

Finally, assumption (H5)(ii) allows us to deduce that :

$$\lim_{n \rightarrow +\infty} \mathbb{E}^2(T_{2,j}) = 0.$$

**Proof of (3.9).** By following the same ideas as those used in (3.11), we show that :

$$\mathbb{E} |T_{4,j} - \mathbb{E}(T_{4,j})| \xrightarrow{L^1} 0 \quad \text{as } n \rightarrow +\infty, \quad (3.32)$$

$$\mathbb{E} |(T_{3,i} - 1)T_{4,j} - \mathbb{E}((T_{3,i} - 1)T_{4,j})| \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \quad (3.33)$$

**Proof (3.32).**

To show the required result (3.32), it suffices to prove the  $L^2$ -consistency of  $T_{4,j}$

$$\begin{aligned} E \left( (T_{4,j} - \mathbb{E}(T_{4,j}))^2 \right) &= n \text{Var}(T_{4,1}) \\ &+ 2 \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1^2 K_1)}{h_H \mathbb{E}^2(\Delta_1 K_1)} \sum_{1 \leq i < j \leq n} \text{Cov}(\beta_i K_i (H_i - F^x(y)), \beta_j K_j (H_j - F^x(y))). \end{aligned} \quad (3.34)$$

Concerning the first term on the right-hand side of (3.34), we have the following :

$$\begin{aligned} n \text{Var}(T_{4,1}) &= \frac{n\phi_x(h_K)\mathbb{E}^2(\beta_1 K_1)}{\mathbb{E}^2(\Delta_1 K_1)} n \text{Var}(\beta_1 K_1 (H_1 - F^x(y))), \\ &= \frac{n\phi_x(h_K) o(h_K^2 \phi_x^2(h_K))}{(n-1)^2 O(h_K^4 \phi_x^4(h_K))} n F^x(y) (1 - F^x(y)) O(h_K^2 \phi_x(h_K)), \\ &= F^x(y) (1 - F^x(y)) o(1). \end{aligned}$$

On the other hand, by exactly the same arguments at (3.13), the second term on the right-hand side of (3.34) tends to 0 as  $n$  tends to the infinity, and the desired result (3.32) is obtained.

**Proof (3.33).**

The *Cauchy-Schwarz's* inequality implies that :

$$\mathbb{E} |(T_{3,i} - 1)T_{4,j} - \mathbb{E}((T_{3,i} - 1)T_{4,j})| \leq 2\mathbb{E} |(T_{3,i} - 1)T_{4,j}| \leq 2\sqrt{\mathbb{E}((T_{3,i} - 1)^2)} \sqrt{\mathbb{E}(T_{4,j}^2)}.$$

In a first attempt, we have the following :

$$\mathbb{E}((T_{3,i} - 1)^2) = n \text{Var}(T_{3,1}) + 2 \text{Cov}(T_{3,i}, T_{3,1}).$$

Hence :

$$\begin{aligned}
 n \operatorname{Var}(T_{2,1}) &= \frac{n \operatorname{Var}(\beta_1 K_1)}{n^2 \mathbb{E}^2(\beta_1 K_1)}, \\
 &\leq \frac{\mathbb{E}(\beta_1^2 K_1^2)}{n \mathbb{E}^2(\beta_1 K_1)}, \\
 &= O\left(\frac{1}{n \phi_x(h_K)}\right).
 \end{aligned}$$

Next, we use similar ideas as those used in (3.29) to deduce that :

$$\operatorname{Cov}(T_{3,i}, T_{3,l}) \longrightarrow 0 \quad \text{as } n \longrightarrow +\infty.$$

### 3.5.2 Proof of Lemma 3.3.2

By the definition of  $\widehat{F}_D^x$  we have the following :

$$\begin{aligned}
 \widehat{F}_D^x &= \frac{1}{n} \sum_{i=1}^n \beta_i^2 K_i \frac{1}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n K_j - \frac{1}{n} \sum_{i=1}^n \beta_i K_i \frac{1}{\mathbb{E}(\Delta_1 K_1)} \sum_{j=1}^n \beta_j K_j \\
 &= \frac{n \mathbb{E}(\beta_1^2 K_1) \mathbb{E}(K_1)}{\mathbb{E}(\Delta_1 K_1)} \frac{1}{n \mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i - \frac{1}{n \mathbb{E}(K_1)} \sum_{j=1}^n K_j \\
 &\quad - \frac{n \mathbb{E}^2(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)} \frac{1}{n \mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i - \frac{1}{n \mathbb{E}(\beta_1 K_1)} \sum_{j=1}^n \beta_j K_j
 \end{aligned}$$

Let us write

$$\widehat{F}_D =: A_1 T_{1,i} T_{2,j} - A_2 T_{3,i} T_{4,j},$$

where

$$\begin{aligned}
 T_{1,i} &= \frac{1}{n \mathbb{E}(\beta_1^2 K_1)} \sum_{i=1}^n \beta_i^2 K_i, & T_{2,j} &= \frac{1}{n \mathbb{E}(K_1)} \sum_{j=1}^n K_j, \\
 T_{3,i} &= \frac{1}{n \mathbb{E}(\beta_1 K_1)} \sum_{i=1}^n \beta_i K_i, & T_{4,j} &= \frac{1}{n \mathbb{E}(\beta_1 K_1)} \sum_{j=1}^n \beta_j K_j, \\
 A_1 &= \frac{n \mathbb{E}(\beta_1^2 K_1) \mathbb{E}(K_1)}{\mathbb{E}(\Delta_1 K_1)}, & A_2 &= \frac{n \mathbb{E}^2(\beta_1 K_1)}{\mathbb{E}(\Delta_1 K_1)}.
 \end{aligned}$$

Finally, the claimed result will be obtained as soon as the two following claims have been checked :

**claim 1**

$$T_{l,i} \xrightarrow{P} 1 \quad \text{as } n \longrightarrow +\infty \quad \text{for } l \in \{1, 3, 4\}.$$



**claim 2**

$$T_{2,j} \xrightarrow{P} 1 \text{ as } n \longrightarrow +\infty, \quad A_1 \xrightarrow{P.S} 1 \text{ as } n \longrightarrow +\infty,$$

$$\text{and } A_2 = o(1) \text{ as } n \longrightarrow +\infty.$$

**Proof of claim 1**

By combining the sufficient convergence condition in probability with (3.27), we obtain the following :

$$\lim_{n \rightarrow +\infty} \mathbb{E}(T_{l,i}) = 1 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \text{Var}(T_{l,i}) = 0 \text{ for } l \in \{1, 3, 4\}.$$

**Proof of claim 2**

By following the same ideas as those used in claim 1, we show that :  $\lim_{n \rightarrow +\infty} \text{Var}(T_{2,j}) = 0$ .

First, we have the following :

$$\text{Var}(T_{2,j}) = \text{Var}(T_{2,1}) + 2 \text{Cov}(T_{2,i}, T_{2,j}),$$

where

$$\begin{aligned} \text{Var}(T_{2,1}) &= \frac{1}{n\mathbb{E}^2(K_1)} \left( \mathbb{E}(K_1^2) - \mathbb{E}^2(K_1) \right) \\ &= \frac{\mathbb{E}(K_1^2)}{n\mathbb{E}^2(K_1)} - \frac{1}{n}. \end{aligned}$$

Second, we use the technical Lemma A.1 of [2], to get :

$$\frac{\mathbb{E}(K_1^2)}{n\mathbb{E}^2(K_1)} = \frac{M_2}{M_1^2 n \phi_x(h_K)},$$

then by the assumption (H5)(i) we have the following :

$$\lim_{n \rightarrow +\infty} \text{Var}(T_{2,1}) = 0.$$

Moreover :

$$\text{Cov}(T_{2,i}, T_{2,j}) = \frac{1}{n^2 \mathbb{E}^2(K_1)} \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \text{Cov}(K_i, K_j).$$

Let us now define the sets  $E_1$  and  $E_2$  as follows :

$$E_1 = \{(i, j) \in \{1, 2, \dots, n\} \text{ such that } 1 \leq |i - j| \leq m_n\},$$

$$E_2 = \{(i, j) \in \{1, 2, \dots, n\} \text{ such that } m_n + 1 \leq |i - j| \leq n - 1\},$$

where the sequence  $m_n$  is chosen, such that  $m_n \rightarrow +\infty$  as  $n \rightarrow +\infty$ , and we denote by  $A_{1,n}$  and  $A_{2,n}$  the sum of covariances over  $E_1$  and  $E_2$  respectively, then

$$A_{1,n} = \frac{1}{n^2 \mathbb{E}^2(K_1)} \sum_{E_1} Cov(K_i, K_j).$$

By stationarity and assumption (H3), we have the following : :

$$\begin{aligned} |Cov(K_i, K_j)| &\leq |\mathbb{E}(K_i K_j)| + \mathbb{E}^2(K_1), \\ &\leq C\psi_x(h_K) + \mathbb{E}^2(K_1). \end{aligned}$$

Now, we use the technical lemma A.1 of [37], to obtain :

$$\begin{aligned} |A_{1,n}| &\leq n m_n (\psi_x(h_K) + \mathbb{E}^2(K_1)) \frac{1}{n^2 \mathbb{E}^2(K_1)} \\ &\leq \frac{C m_n \psi_x(h_K)}{M_1^2 n \phi_x^2(h_K)} + \frac{m_n}{n}. \end{aligned}$$

The fact that  $\frac{\psi_x(h_K)}{\phi_x^2(h_K)}$  is bounded by assumption (H3), and by the choice  $m_n = \sqrt{n}$  permits to get :

$$A_{1,n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Concerning the sum over  $E_2$ , by following the same ideas as those used in (3.29) we get :

$$|A_{2,n}| = \frac{1}{n^2 \mathbb{E}^2(K_1)} \sum_{E_2} |Cov(K_i, K_j)| \leq \frac{C (m_n)^{1-s}}{s-1} \frac{1}{M_1^2 n \phi_x^2(h_K)}.$$

We use the same choice of  $m_n$  as before, and by assumption (H5)(i), we obtain the following :

$$A_{2,n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Finally, to show that  $A_1 \xrightarrow{P.S} 1$  and  $A_2 = o(1)$  as  $n \rightarrow +\infty$ , it suffices to apply the technical Lemma A.1 of [37].

### Acknowledgements

The authors would like to thank the Editor, the Associate Editor, and the two anonymous reviewers for their valuable comments and suggestions which improved substantially the quality of an earlier version of this paper.

# Bibliographie

- [1] Baïllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *J. of Multivariate Analysis*, **100**, 102-111.
- [2] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally modelled regression and functional data. *J. of Nonparametric Statist.*, **22**, 617-632.
- [3] Berline, A., Gannoun, A. and Matzner-Lober, E. (1998) Propriétés asymptotiques d'estimateurs convergents des quantiles conditionnels. *C. R. Acad. Sci. Paris I*, **326**, 611-614.
- [4] Boj, E., Delicado, P. and Fortiana, J. (2010). Distance-based local linear regression for functional predictors. *Computational Statistics and Data Analysis*, **54**, 429-437.
- [5] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2018). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*, **46**, 199-233.
- [6] Bozebda, S., Chaouch, N. and Laïb, N. (2016). Limiting law results for a class of conditional mode estimates for functional stationary ergodic data. *Mathematical Methods of statistics*, **25**, 168-195.
- [7] Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227-236.
- [8] Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data. *Statistica*, **70**, 83-94.
- [9] Dabo-Niang, S., Kaid, Z. and Laksaci, A. (2012). On spatial conditional mode estimation for a functional regressor. *Statist. Probab. Lett.*, **82**, 1413-1421.

- 
- [10] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2011). *A fast functional locally modeled conditional density and mode for functional time-series*. Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, 2011, 85-90, DOI : 10.1007/978-3-7908-2736-1\_13, Physica-Verlag/Springer.
- [11] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics*, **47**, 26-44.
- [12] Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the local linear modelization of the conditional distribution for functional data. *Sankhya*, **76**, 328-355.
- [13] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2016). Estimation locale linéaire de la fonction de régression pour des variables hilbertiennes. *C.R. Math. Acad. Sci. Paris*, **354**, 847-850.
- [14] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2017). Local linear regression modelization when all variables are curves. *Statistics and Probability Letters*, **121**, 37-44.
- [15] Ezzahrioui, M. and Ould-Said, E. (2008). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East J. Theoretical Statist.*, **25**, 15-38.
- [16] Fan, J., Gijbels, I. (1996). *Local polynomial modelling and its applications*. London, Chapman and Hall.
- [17] Ferraty, F. and Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist.*, **17**, 545-564.
- [18] Ferraty, F., Laksaci, A. and Vieu, F. (2005). Functional time series prediction via the conditional mode estimation. *C. R., Math., Acad. Sci. Paris*, **340**, 389-392.
- [19] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis : Theory and Practice*. Springer Series in Statistics, New York.
- [20] Ferraty, F. and Romain, Y. (2010). *The oxford handbook of functional data analysis*, Oxford University Press.
- [21] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2011). Kernel regression with functional response. *Electronic Journal of Statistics*, **5**, 159-171.

- 
- [22] Honda, T. (2000). Nonparametric estimation of a conditional quantile for a mixing process, *Ann. Inst. Statist. Math.* **52**, 459-470.
- [23] Horváth, L. and Kokoszka, P. (2012). Inference for functional data with applications. *Springer Series in Statistics*, Springer, New York.
- [24] Khardani, S., Lemdani, M. and Ould-Saïd, E. (2011). Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series. *J. Statist. Plann. and Inf.*, **141**, 3426-3436.
- [25] Kokoszka, P. and Reimherr, M. (2017). *Introduction to functional data analysis*. Texts in statistical science. Chapman & Hall/CRC.
- [26] Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Ann. I.S.U.P.*, **51**, 69-80.
- [27] Laksaci, A., Rachdi, M. and Rahmani, S., (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1-23.
- [28] Laksaci, A., Nacéri, A. and Rachdi, M. (2015). Exact quadratic error of the local linear regression operator estimator for functional co-variates. *Functional Statistic and Application*, Springer, 79-90.
- [29] Liang, H-Y. and Ould-Saïd, E. (2016). A weighted estimator of conditional hazard rate with left-truncated and dependent data. *Ann. Inst. Stat. Math.*, **69**, 1-35.
- [30] Louani, D. and Ould-saïd, E. (2007). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. of Nonparametric Statistics*. **11**, 413-442
- [31] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.
- [32] Ould-Saïd, E. (1993). Estimation nonparamétrique du mode conditionnel. Application à la prévision. *C. R. Acad. Sci. Paris., Série I*, **316**, 943-947.
- [33] Ould-Saïd, E. (1997). A note on ergodic processes prediction via estimation of the conditional mode function. *Scandinavian J. Statist.* **24**, 231-239.
- [34] Ould-Saïd, E. and Tatachak, A. (2011). A nonparametric conditional mode estimate under RLT model and strong mixing condition. *Int. J. Stat. Econ.*, **6**, 76-92.

- [35] Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A., Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data, *Computational Statistics and Data Analysis.* **73**, 53-68.
- [36] Volkonskii, Yu.A. Rozanov. (1959). Some limit theorems for random functions, *Theory Probab. Appl.* **4**, 178-197.
- [37] Zhou, Z. and Lin, Z.-Y. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *Nonparametric Stat.***28**, 116-131.

# Asymptotic normality of conditional mode estimator for functional dependent data

Ce chapitre fait l'objet d'une publication parue dans le revue internationale Indian Journal of Pure and Applied Mathematics (IJPAM).51(2) : 465-481, June 2020.

**Asymptotic normality of conditional mode estimator for functional dependent data**

Oussama Bouanani <sup>1</sup>, Saâdia Rahmani <sup>1</sup>, Ali Laksaci <sup>2</sup>, and Mustapha Rachdi <sup>3</sup>

<sup>1</sup> Laboratoire de Modèles Stochastiques, Statistique et Applications, Université Dr moulay Tahar, Saïda 20000, Algeria. E-mail : oussamaproba@gmail.com. E-mail : saadia.rahmani@gmail.com.

<sup>2</sup> Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia. E-mail : alilak@yahoo.fr.

<sup>3</sup> Université Grenoble Alpes (France), Laboratoire AGEIS, EA 7407, AGIM Team, UFR SHS, BP. 47, F38040 Grenoble Cedex 09, France. E-mails : mustapha.rachdi@univ-grenoble-alpes.fr

**Abstract :** Based on the local polynomial smoother idea, we construct a local linear estimator of the conditional mode for dependent functional covariables. Precisly, observations are assumed to be a sequence of stationary  $\alpha$ -mixing random variables. Then, we establish the asymptotic normality of the constructed estimator.

**Key Words .** Functional dependent data · Local linear estimation · conditional cumulative Derivatives of the conditional density · Conditional mode · asymptotic normality.

AMS Classification : 62G05, 62G08, 62G20, 62G07, 62G30.

Secondary : 62H12.

## 4.1 Introduction and motivations

Nonparametric estimation in models containing functional data has been the subject of many studies over the last decade. This area of statistical research, called non-parametric functional data analysis (NFDA), is concerned with non-parametric modeling of data in the form of curves, images or objects. For a general overview on this subject, one can refer to the monographs by Ferraty and Romain [20] and Ferraty and Vieu [19] as well as to the references therein.

It is very common in NFDA that the statistical prediction of a scalar response from a functional explanation variable is performed by estimating the conditional expectation of  $Y$  given  $X$ . However, the regression function is not efficient enough in some situations. For example, when the conditional density is asymmetric or it is multi-modal. In this situation, the conditional mode would be more efficient than the regression function. Motivated by this importance, the study of the conditional mode has attracted the attention of many



researchers (see Ferraty et al. [18], Khardani et al. [21] and Ould-Saïd and Tatachak [30], and more recently Dabo-Niang et al. [8], for more discussion and motivation). Let's point out that in the above works the classical Nadaraya-Watson estimator of the conditional mode has been considered, but it is well known that the local linear estimation method has more advantages than the latter (see Fan and Gijbels [17] in the finite dimensional framework). In fact, this type of estimator makes it possible, among other things, to improve the term of the bias of the Nadaraya-Watson estimator and to avoid its boundary effects. Notice that the local linear estimation of the conditional mode was introduced in NFDA by Demongeot et al. [9], who proved the almost-complete convergence of the local linear estimator of the conditional mode for independent and identically distributed (i.i.d.) data. We also return to Laksaci et al. [22] for the case of spatial functional data. Notice also that some authors have been interested in the linear local estimation of the regression operator in NFDA (see, for instance, Baillo and Grané [1] who dealt with the case where the regressor is in a Hilbert space, and El Methni and Rachdi [14], Ouassou and Rachdi [?] and Boj et al. [3] for alternative versions of the linear local estimator).

On the other hand, Barrientos et al. [2] studied a so-called fast functional version of the linear local estimator when the regressor belongs to a semi-metric space. In this paper, we will use this last approach to construct a conditional mode estimator. Our main goal would be to study its asymptotic properties when data are dependent. In addition, in order to use the conditional model as a predictor in time series analysis, we will focus on the strongly mixing data framework. This case has attracted the attention of several researchers in the finite dimensional statistics (see, for instance, Collomb et al. [6] for the strong convergence of the kernel estimator of the conditional mode, Ould-Saïd [29] for the conditional mode prediction for ergodic processes, as well as Bouzebda et al. [5], for the latest advances and references).

In fact, in this article, we prove, under certain standard conditions, the asymptotic normality of the local linear estimator when we approach the problem of the prediction of a functional time series by estimating the conditional mode. For this purpose we will construct, in section 2, the linear local estimator of the conditional model. In section 3, we will introduce and discuss necessary conditions for establishing the asymptotic normality of the estimator. Finally, details of technical lemmas of the proof of the main result are given in the appendix (see Section 4).

## 4.2 Local linear estimator construction

Let  $(X_i, Y_i)_{1 \leq i \leq n}$  be a stationary  $\alpha$ -mixing process taking values in  $\mathfrak{F} \times \mathbb{R}$  where  $\mathfrak{F}$  is a semi-metric space equipped with the semi-metric  $d(\cdot, \cdot)$ . We assume that there exists a regular version of the conditional probability of  $Y$  given  $X$  and that, for a given  $x \in \mathfrak{F}$ , there is some compact subset  $S = [\theta - \xi, \theta + \xi]$ ,  $\xi > 0$ , such that the conditional density of  $Y$  given  $X = x$  has a unique mode  $\theta(x)$  in  $S$ , which is defined by

$$f^x(\theta(x)) = \sup_{y \in S} f^x(y).$$

A local linear estimator of  $\theta(x)$  is defined by

$$\hat{f}^x(\hat{\theta}(x)) = \sup_{y \in S} \hat{f}^x(y), \quad (4.1)$$

where the conditional density function estimator is “ $a$ ”, the minimizer of the following criterion

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n \left( h_H^{-1} H^{(1)}(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1}\delta(x, X_i)), \quad (4.2)$$

where  $\beta(\cdot, \cdot)$  and  $\delta(\cdot, \cdot)$  are a known bi-functional operators which are defined from  $\mathfrak{F}^2$  into  $\mathbb{R}$  such that  $|\delta(x, z)| = d(x, z)$  and  $\delta = \varphi(\beta)$  where  $\varphi$  is a measurable function,  $K$  is a kernel function,  $H^{(1)}$  is the first derivative of a given distribution function  $H$  and  $h_K = h_{K,n}$  (resp.  $h_H = h_{H,n}$ ) is a sequence of positive real numbers.

More precisely, the local linear estimator  $\hat{f}^x(y)$ , of  $f^x(y)$ , is then  $\hat{a}$  which is the first component of the pair  $(a, b)$  solution of the minimization problem (4.2). However, if for all  $z \in \mathfrak{F}$ ,  $\beta(z, z) = 0$ , then  $\hat{f}^x(y)$  is explicitly defined by

$$\hat{f}^x(y) = \frac{\sum_{i=1}^n \sum_{k=1}^n W_{ik}(x) H^{(1)}(h_H^{-1}(y - Y_k))}{h_H \sum_{i=1}^n \sum_{k=1}^n W_{ik}(x)}, \quad (4.3)$$

where  $W_{ik}(x) = \beta(X_i, x) (\beta(X_i, x) - \beta(X_k, x)) K(h_K^{-1}\delta(x, X_i)) K(h_K^{-1}\delta(x, X_k))$ , with the convention  $0/0 = 0$  (see Barrientos et al. (2010), for some details).

By reducing the double sum in (4.3) into a single one, and set

$$\Delta_k := K_k^{-1} \sum_{i=1}^n W_{ik} = \sum_{i=1}^n \beta_i^2 K_i - \sum_{i=1}^n \beta_i K_i \beta_k,$$

with  $\beta_i = \beta(X_i, x)$ ,  $K_i = K(h_K^{-1} \delta(x, X_i))$  and  $H_k^{(1)} = H^{(1)}(h_H^{-1}(y - Y_k))$ , we get the following simple formula of  $\widehat{f}^x(y)$

$$\widehat{f}^x(y) = \frac{\sum_{k=1}^n \Delta_k K_k H_k^{(1)}}{h_H \sum_{k=1}^n \Delta_k K_k}.$$

Furthermore, by defining

$$\widehat{f}_N^x(y) = \frac{1}{nh_H \mathbb{E}[\Delta_1 K_1]} \sum_{k=1}^n \Delta_k K_k H_k^{(1)} \text{ and } \widehat{f}_D^x = \frac{1}{n \mathbb{E}[\Delta_1 K_1]} \sum_{k=1}^n \Delta_k K_k,$$

we can write

$$\widehat{f}^x(y) = \frac{\widehat{f}_N^x(y)}{\widehat{f}_D^x}.$$

## 4.3 Main result

### 4.3.1 Assumptions and notations

In what follows,  $x$  (resp.  $y$ ) is a fixed functional element in the space  $\mathfrak{F}$  (resp.  $\mathbb{R}$ ),  $\mathcal{N}_x$  (resp.  $\mathcal{N}_y$ ) will denote a fixed neighborhood of  $x$  (resp. of  $y$ ) and

$$\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1).$$

Moreover, we will set, for any  $l \in \{0, 2\}$

$$\Lambda_l(s) = \mathbb{E}[\lambda_l(X, y) - \lambda_l(x, y) | \beta(x, X) = s] \text{ where } \lambda_l(x, y) = \frac{\partial^l f^x(y)}{\partial y^l}.$$

In order to establish our asymptotic results we need the following hypotheses.

**(H1)**

- (i) For any  $r > 0$ ,  $\phi_x(r) := \phi_x(-r, r) > 0$  where  $\lim_{r \rightarrow 0} \phi_x(r) = 0$ .

(ii) There exist a function  $\Psi_x(\cdot)$  such that

$$\lim_{h_K \rightarrow 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \Psi_x(t) \text{ for all } t \in [-1, 1].$$

**(H2)** For any  $l \in \{0, 2\}$ , the quantities  $\Lambda_l^{(2)}(0)$  exist.

**(H3)**

$$\sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) \leq \psi_x(h_K),$$

where  $\psi_x(h_K)$  is such that there exists  $\epsilon \in \left]0, \frac{a+1}{a-1}\right]$  for which

$$0 < \psi_x(h_K) = O\left(\phi_x^{1+\epsilon}(h_K)\right).$$

**(H4)** The stochastic process  $(X_i, Y_i)_{i \in \mathbb{N}}$  is  $\alpha$ -mixing whose coefficient verifies that there exist  $a > 1$ ,  $C > 0$  such that for all  $n \in \mathbb{N}$

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

**(H5)** There exist some positive constants  $C_1$  and  $C_2$  such that the bandwidths satisfy

$$C_2 n^{1-a} \leq h_H \phi_x(h_K) \leq C_1 n^{\frac{1}{1-a}} \text{ and } \lim_{n \rightarrow +\infty} \frac{\log(n)}{n h_H^3 \phi_x(h_K)} = 0,$$

with  $a > 1$ .

**(H6)** The bi-functions  $\delta$  and  $\beta$  satisfy that, there exist some positive constants  $C_1$  and  $C_2$  such that, for all  $z \in \mathfrak{F}$

$$C_1 |\delta(x, z)| \leq |\beta(x, z)| \leq C_2 |\delta(x, z)|.$$

**(H7)**

(i) The kernel  $K$  is a positive and differentiable function, for which the support is within  $(-1, 1)$ .

(ii) The  $j$ th order derivatives  $H^{(j)}$  for  $j = 1, 2$ , are bounded and Lipschitzian functions.

Before enouncing the main result, we introduce the following quantities in order to provide bias and variance dominant terms of  $\hat{f}^x(y)$ .

$$M_j = K^j(1) - \int_{-1}^1 (K^j(u))' \Psi_x(u) du \text{ where } j = 1, 2,$$

$$N(a, b) = K^a(1) - \int_{-1}^1 (u^b K^a(u))' \Psi_x(u) du \text{ for all } a > 0 \text{ and } b = 2, 4.$$

### 4.3.2 Comments on the assumptions

Observe that assumptions (H1)-(H2) and (H6) are not unduly restrictive, and are common in the setting of functional local linear fitting (see, for instance, Barrientos et al. [2] and Demongeot et al. [11] among others). Concerning the first part of the assumption (H1), the reader will find in the book by Ferraty and Vieu [19] a deeper discussion on the links between this assumption, the semi-metric  $d$  and the small ball concentration properties, whereas the second part, of this assumption, will play a key role in our methodology, in particular when we will have to compute the exact constant terms involved in our asymptotic result. In order to quantify the expression of the covariance term we need the assumption (H3). Notice also that the conditions on the smoothing parameters  $h_K$  and  $h_H$  are standard. On the other hand, the boundedness of the kernel  $K$  in the assumption (H7)(i) is also standard; and assumptions (H5)-(H7)(ii) are some technical conditions which make the theorem proof fast and brief.

We are now ready for enouncing the asymptotic normality of the estimator  $\hat{\theta}(x)$ . Notice that despite there are already various asymptotic results on the asymptotic behaviour of the classical kernel method (see, for instance, Dobo-Niang and Laksaci [7], Ezzahrioui and Ould-Saïd [16] or Ferraty et al. [18]), to the best of our knowledge this constitutes the first result on the asymptotic normality of the local linear estimator of the conditional mode for functional strongly mixing processes.

**Theorem 14** *Under assumptions (H1)-(H8), we have*

$$\left( \frac{n h_H^3 \phi_x(h_K) (f^{x(2)}(\theta(x)))^2}{V_{HK}(x, \theta(x))} \right)^{1/2} (\hat{\theta}(x) - \theta(x)) \xrightarrow{D} \mathcal{N}(0, 1),$$

where

$$V_{HK}(x, \theta(x)) = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt,$$

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

### 4.3.3 Proof of Theorem 14.

Based on the Taylor expansion of  $\hat{f}^{x(1)}(\cdot)$  in the neighborhood of  $\theta(x)$  and according to the assumptions (H2) and (H3), we have

$$\hat{\theta}(x) - \theta(x) = -\frac{\hat{f}^{x(1)}(\theta(x))}{\hat{f}^{x(2)}(\bar{\theta}(x))} = -\frac{\hat{f}_N^{x(1)}(\theta(x))}{\hat{f}_N^{x(2)}(\bar{\theta}(x))}, \quad (4.4)$$

where  $\bar{\theta}(x)$  is between  $\hat{\theta}(x)$  and  $\theta(x)$ .

By using (4.4), we obtain the following decomposition

$$\begin{aligned} & \sqrt{nh_H^3 \phi_x(h_K)} \left( \hat{f}_N^{x(2)}(\bar{\theta}(x)) \right)^2 (\theta(x) - \hat{\theta}(x)) \\ &= \sqrt{nh_H^3 \phi_x(h_K)} \left[ \hat{f}_N^{x(1)}(\theta(x)) - \mathbb{E}[\hat{f}_N^{x(1)}(\theta(x))] \right] + \sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E} \left[ \hat{f}_N^{x(1)}(\theta(x)) \right] \\ &=: J + \hat{J}. \end{aligned} \tag{4.5}$$

Then, the rest of the proof of this theorem is based on the following lemmas for which proofs are given in the appendix.

**Lemma 4.3.1** (see Rachdi et al. [31]) Under the assumptions (H2), (H3) and (H5)(ii)-(H7)(ii), we have

$$\hat{J} \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

**Lemma 4.3.2** Under the assumptions of Theorem 14, we have

$$J \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, \theta(x))).$$

**Lemma 4.3.3** Under assumptions (H5), (H7)(ii) and (H3)(i), we have

$$\hat{f}_N^{x(2)}(\bar{\theta}(x)) \xrightarrow{P} f^{x(2)}(\theta(x)),$$

where  $\xrightarrow{P}$  denotes the convergence in probability.

## 4.4 Appendix

*Proof of Lemma 4.3.2.* Denote

$$\begin{aligned} & \sqrt{nh_H^3 \phi_x(h_K)} \left[ \hat{f}_N^{x(1)}(\theta(x)) - \mathbb{E}[\hat{f}_N^{x(1)}(\theta(x))] \right] \\ &= \frac{(nh_H^3 \phi_x(h_K))^{\frac{1}{2}}}{nh_H^2 \mathbb{E}(\Delta_1 K_1)} \sum_{i=1}^n L_i(x, \theta(x)) = S_n, \end{aligned}$$

where

$$L_i(x, \theta(x)) = \Delta_i K_i H_i^{(2)} - \mathbb{E}[\Delta_i K_i H_i^{(2)}].$$

Then, we consider the following decomposition

$$S_n = T_{1,k}T_{1,i} - \mathbb{E}[T_{1,k}T_{1,i}] - (T_{2,k}T_{2,i} - \mathbb{E}[T_{2,k}T_{2,i}]),$$

where

$$T_{1,k} = \frac{1}{n\mathbb{E}[\beta_1^2 K_1]} \sum_{k=1}^n \beta_k^2 K_k, \quad T_{1,i} = \frac{\sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E}[\beta_1^2 K_1]}{h_H^2 \mathbb{E}[\Delta_1 K_1]} \sum_{i=1}^n K_i H_i^{(2)},$$

$$T_{2,k} = \frac{1}{n\mathbb{E}[\beta_1 K_1]} \sum_{k=1}^n \beta_k K_k \quad \text{and} \quad T_{2,i} = \frac{\sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E}[\beta_1 K_1]}{h_H^2 \mathbb{E}[\Delta_1 K_1]} \sum_{i=1}^n \beta_i K_i H_i^{(2)}.$$

Obviously, the result in Lemma 4.3.2 can be deduced directly from the following statements.

*Claim 1.*

$$T_{1,k}T_{1,i} - \mathbb{E}[T_{1,k}T_{1,i}] \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, \theta(x))).$$

*Claim 2.*

$$T_{2,k}T_{2,i} - \mathbb{E}[T_{2,k}T_{2,i}] \xrightarrow{P} 0.$$

The rest of this proof will be devoted to the demonstration of these two claims.

*Proof of Claim 1.* We can write

$$\begin{aligned} & T_{1,k}T_{1,i} - \mathbb{E}[T_{1,k}T_{1,i}] \\ &= T_{1,i} - \mathbb{E}[T_{1,i}] + ((T_{1,k} - 1)T_{1,i} - \mathbb{E}[(T_{1,k} - 1)T_{1,i}]). \end{aligned}$$

So, it suffices to prove the following two results

$$((T_{1,k} - 1)T_{1,i} - \mathbb{E}[(T_{1,k} - 1)T_{1,i}]) \xrightarrow{P} 0. \quad (4.6)$$

$$T_{1,i} - \mathbb{E}[T_{1,i}] \xrightarrow{D} \mathcal{N}(0, V_{HK}(x, \theta(x))). \quad (4.7)$$

*Proof of (4.6).* Firstly, by using the Cauchy-Schwartz inequality we obtain that

$$\begin{aligned} \mathbb{E} |(T_{1,k} - 1)T_{1,i} - \mathbb{E}[(T_{1,k} - 1)T_{1,i}]| &\leq 2\mathbb{E} |(T_{1,k} - 1)T_{1,i}| \\ &\leq \sqrt{E[(T_{1,k} - 1)^2]} \sqrt{\mathbb{E}[T_{1,i}^2]}. \end{aligned}$$

Then, all it remains to prove is

$$\mathbb{E}[(T_{1,k} - 1)^2] \rightarrow 0, \text{ as } n \rightarrow +\infty. \quad (4.8)$$

$$\mathbb{E}[T_{1,i}^2] \rightarrow V_{HK}(x, \theta(x)), \text{ as } n \rightarrow +\infty. \quad (4.9)$$

*Proof of (4.8).* Firstly, we can write

$$\mathbb{E}[(T_{1,k} - 1)^2] = n \text{Var}[T_{1,1}] + 2 \text{Cov}(T_{1,k}, T_{1,l}). \quad (4.10)$$

For the first term on the right hand side of (4.10), we have

$$\begin{aligned} n \text{Var}[T_{1,1}] &= \frac{1}{n^2 \mathbb{E}^2[\beta_1^2 K_1]} n \text{Var}[\beta_1^2 K_1] \\ &= O\left(\frac{1}{n \phi_x(h_K)}\right). \end{aligned}$$

Next, before computing the second term of (4.10) we treat the general case of  $\text{Cov}(T_{j,k}, T_{j,l})$  where  $j \in \{1, 2\}$ . So, by denoting for  $c = 2, j = 1$  and for  $c = 1, j = 2$ , we have

$$\frac{1}{n \mathbb{E}[\beta_1^c K_1]} \sum_{k=1}^n \beta_k^c K_k = T_{j,k}.$$

Then

$$\text{Cov}(T_{j,k}, T_{j,l}) = \frac{1}{n^2 \mathbb{E}^2[\beta_1^c K_1]} \sum_{k \neq l} \text{Cov}(\beta_k^c K_k, \beta_l^c K_l). \quad (4.11)$$

In order to adopt the same technique as in Masry [26], we define the sets  $E_1$  and  $E_2$  as follows

$$E_1 = \{(k, l) \in \{1, 2, \dots, n\} \text{ such that } 1 \leq |k - l| \leq m_n\},$$

and

$$E_2 = \{(k, l) \in \{1, 2, \dots, n\} \text{ such that } m_n + 1 \leq |k - l| \leq n - 1\},$$

where the sequence  $m_n$  is chosen such that  $m_n \rightarrow +\infty$  as  $n \rightarrow +\infty$ .

Let  $A_{1,n}$  and  $A_{2,n}$  be the sum of covariances over the sets  $E_1$  and  $E_2$  respectively

$$A_{1,n} = \frac{1}{n^2 \mathbb{E}^2[\beta_1^c K_1]} \sum_{E_1} \text{Cov}(\beta_k^c K_k, \beta_l^c K_l),$$

and

$$A_{2,n} = \frac{1}{n^2 \mathbb{E}^2[\beta_1^c K_1]} \sum_{E_2} \text{Cov}(\beta_k^c K_k, \beta_l^c K_l).$$



For the sum  $A_{1,n}$ , from the stationarity property, we have

$$|Cov(\beta_k^c K_k, \beta_l^c K_l)| \leq |\mathbb{E}[\beta_k^c \beta_l^c K_l K_k]| + \mathbb{E}^2[\beta_1^c K_1].$$

Moreover, assumption (H6) implies that

$$\text{for all } (k, c) \in (\mathbb{N}^*, \mathbb{N}) : K_i^k |\beta_i|^c h_K^{-c} \leq CK_1^k |\delta(x, X_i)|^c h_K^{-c},$$

and from the Jensen's inequality and under assumption (H3), it follows that

$$\begin{aligned} |\mathbb{E}[\beta_k^c \beta_l^c K_l K_k]| &\leq \mathbb{E}[|\beta_k^c \beta_l^c K_l K_k|], \\ &\leq Ch_K^{2c} \psi_x(h_K). \end{aligned}$$

Then

$$|Cov(\beta_k^c K_k, \beta_l^c K_l)| \leq Ch_K^{2c} \psi_x(h_K) + \mathbb{E}^2[\beta_1^c K_1].$$

This last result together with the technical lemma A.1 in [2] lead directly to

$$\begin{aligned} |A_{1,n}| &\leq C n m_n \left( h_K^{2c} \psi_x(h_K) + \mathbb{E}^2[\beta_1^c K_1] \right) \frac{1}{n^2 \mathbb{E}^2[\beta_1^c K_1]} \\ &\leq \frac{C m_n}{n \phi_x(h_K)^{1-\epsilon}}. \end{aligned}$$

For the sum  $A_{2,n}$ , we use the inequality for the bounded mixing processes (see Proposition A.10(i) in [19]). This leads, for all  $l \neq k$ , to

$$|Cov(\beta_k^c K_k, \beta_l^c K_l)| \leq Ch_K^{2c} \alpha(|k - l|).$$

On the other hand, by using the inequality  $\sum_{j \geq x+1} j^{-a} \leq \int_{u \geq x} u^{-a}$ , we get

$$\sum_{k=1}^n \sum_{E_2} \alpha(|k - l|) \leq \frac{C n (m_n)^{1-a}}{a - 1}. \quad (4.12)$$

Finally, we arrive at

$$|A_{2,n}| \leq \frac{C (m_n)^{1-a}}{n \phi_x^2(h_K)}.$$

Then, by choosing  $m_n = (\phi_x^{1+\epsilon}(h_K))^{-1/a}$ , we get

$$A_{1,n} \rightarrow 0 \text{ and } A_{2,n} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Consequently, in the case where  $c = 2$  and  $j = 1$ , we have

$$\lim_{n \rightarrow +\infty} Cov(T_{1,k}, T_{1,l}) = 0.$$

*Proof of (4.9).* We start by writing

$$\mathbb{E}[T_{1,i}^2] = Var[T_{1,i}] + \mathbb{E}^2[T_{1,i}]. \quad (4.13)$$

The first term on the right hand side of (4.13) tends to  $V_{HK}(x, \theta(x))$  as  $n \rightarrow \infty$  and details will be given in the proof of result (4.7). For the second term on the right hand side of (4.13), we use similar ideas as those used by Ezzahrioui and Ould-Saïd [15] (see Lemma 2) to deduce that

$$\sqrt{nh_H^3 \phi_x(h_K)} \mathbb{E} \left[ \frac{K_1 H_1^{(2)}}{h_H^2 \phi_x(h_K)} \right] \rightarrow 0 \text{ as } n \rightarrow +\infty. \quad (4.14)$$

Thus, it remains to check that

$$\lim_{n \rightarrow +\infty} \mathbb{E}[T_{1,i}^2] = V_{HK}(x, \theta(x)).$$

*Proof of (4.7).* First, we calculate  $Var(T_{1,i})$

$$Var[T_{1,i}] = n Var[T_{1,1}] + 2 \frac{n \phi_x(h_K) \mathbb{E}^2[\beta_1^2 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \sum_{i \neq j} Cov(K_i H_i^{(2)}, K_j H_j^{(2)}).$$

So, to prove  $\lim_{n \rightarrow +\infty} Var[T_{1,i}] = V_{HK}(x, \theta(x))$ , it is necessary to establish the following results

$$\lim_{n \rightarrow +\infty} Var[T_{1,1}] = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt, \quad (4.15)$$

$$\lim_{n \rightarrow +\infty} \frac{n \phi_x(h_K) \mathbb{E}^2[\beta_1^2 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \sum_{i \neq j} Cov(K_i H_i^{(2)}, K_j H_j^{(2)}) = 0. \quad (4.16)$$

Notice that the proof of (4.16) is very close to the proof of (4.8). So, we focus only on (4.15). To do that we use the technical Lemma A.1 in Zou and Lin [37], to have

$$\begin{aligned} n Var[T_{1,1}] &= \frac{n^2 \phi_x(h_K) \mathbb{E}^2[\beta_1^2 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \mathbb{E} \left[ \left( K_1 H_1^{(2)} \right)^2 \right] \\ &\quad - \frac{\mathbb{E}^2[\beta_1^2 K_1] n \phi_x^2(h_K)}{\mathbb{E}^2[\Delta_1 K_1]} \mathbb{E}^2 \left[ \sqrt{nh_H^3 \phi_x(h_K)} \frac{K_1 H_1^{(2)}}{h_H^2 \phi_x(h_K)} \right]. \end{aligned} \quad (4.17)$$

Then, it follows from (4.14), that the second term on the right hand side of (4.17) tends to 0. Hence

$$n \operatorname{Var}[T_{1,1}] = \frac{n^2 \phi_x(h_K) \mathbb{E}^2[\beta_1^2 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \mathbb{E} \left[ K_1^2 \mathbb{E}[(H_1^{(2)})^2 | X] \right].$$

On one hand we have

$$\begin{aligned} & \frac{1}{h_H} \mathbb{E} \left[ \left( H^{(2)} \left( \frac{y - Y_1}{h_H} \right) \right)^2 \mid X_1 \right] \\ &= \int_{\mathbb{R}} (H^{(2)}(t))^2 [f^{(X_1)}(y - th_H) - f^x(y)] dt + \int_{\mathbb{R}} (H^{(2)}(t))^2 f^x(y) dt, \end{aligned}$$

and by the continuity of  $f^x$  we deduce that

$$\mathbb{E} \left[ K_1^2 \left( H^{(2)} \left( \frac{y - Y_1}{h_H} \right) \right)^2 \mid X_1 \right] \rightarrow \mathbb{E}[K_1^2] f^x(y) \int (H^{(2)}(t))^2 dt \text{ as } n \rightarrow \infty.$$

On the other hand, by using again the technical Lemma A.1 in Zou and Lin [32] we obtain

$$\frac{n^2 \phi_x(h_K) \mathbb{E}[\beta_1^2 K_1]}{\mathbb{E}^2[\Delta_1 K_1]} \rightarrow \frac{1}{\phi_x(h_K) M_1^2} \text{ as } n \rightarrow +\infty.$$

Finally, we arrive at

$$\lim_{n \rightarrow +\infty} \operatorname{Var}[T_{1,i}] = \frac{M_2}{M_1^2} f^x(\theta(x)) \int (H^{(2)}(t))^2 dt.$$

Now to establish the asymptotic normality of the conditional mode estimator dealing with dependent random variables, we start by writing

$$\begin{aligned} T_{1,i} - \mathbb{E}(T_{1,i}) &= \frac{\sqrt{nh_H^3 \phi_x(h_K) \mathbb{E}[\beta_1^2 K_1]}}{h_H^2 \mathbb{E}[\Delta_1 K_1]} \sum_{i=1}^n \left( K_i H_i^{(2)} - \mathbb{E}[K_i H_i^{(2)}] \right) \\ &= \sum_{i=1}^n \tilde{L}_i(x, \theta(x)), \end{aligned}$$

where

$$\tilde{L}_i(x, \theta(x)) = \left( \frac{n}{(n-1)^2 h_H \phi_x(h_K)} \right)^{\frac{1}{2}} \frac{1}{M_1} \left( K_i H_i^{(2)} - \mathbb{E}[K_i H_i^{(2)}] \right).$$

We base ourselves on the CLT by Liebscher [24] (see Corollary 2.2, Page 196), which rests on the asymptotic behavior of the quantity

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E}[\tilde{L}_i^2(x, \theta(x))], \quad (4.18)$$

in addition to the assumptions

$$\left\{ \begin{array}{l} \text{There exists a sequence } \tau_n = o(\sqrt{n}) \text{ such that} \\ \tau_n \leq \left( \max_{i=1, \dots, n} C_i \right)^{-1} \text{ where } C_i = \operatorname{ess\,sup}_{\omega \in \Omega} |\tilde{L}_i(x, \theta(x))| \\ \text{and } \frac{n}{\tau_n} \alpha(\epsilon \tau_n) \rightarrow 0 \text{ for all } \epsilon > 0, \end{array} \right. \quad (4.19)$$

and

$$\left\{ \begin{array}{l} \text{There exists a sequence } (m_n) \text{ of positive integers tending to } \infty \text{ such that} \\ nm_n \gamma_n = o(1) \text{ where } \gamma_n := \max_{1 \leq i \neq j \leq n} \left( \mathbb{E}[|\tilde{L}_i(x, \theta(x)) \tilde{L}_j(x, \theta(x))|] \right) \\ \text{and } \left( \sum_{j=m_n+1}^{\infty} \alpha(j) \right) \sum_{i=1}^n C_i = o(1). \end{array} \right. \quad (4.20)$$

It follows that

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E}[\tilde{L}_i^2(x, \theta(x))] = V_{HK}(x, \theta(x)). \quad (4.21)$$

Concerning (4.19), the boundness of  $H$  and  $K$  allows to obtain  $C_i = O\left(\frac{1}{\sqrt{n} h_H \phi_x(h_K)}\right)$ .

Therefore, we can take  $\tau_n = \sqrt{\frac{n h_H \phi_x(h_K)}{\log n}}$ .

Furthermore, this choice gives, for all  $\epsilon > 0$

$$\begin{aligned} \frac{n}{\tau_n} \alpha(\epsilon \tau_n) &\leq C \left( n^{1-(a+1)/2} (h_H \phi_x(h_K))^{-(a+1)/2} (\log n)^{(a+1)/2} \right) \\ &\leq C n^{1-(a+1)/2 + (a+1)/2(a-1)} (\log n)^{(a+1)/2} \\ &\leq C n^{(3a-a^2)/2(a-1)} (\log n)^{(a+1)/2} \rightarrow 0 \text{ since } a > 3. \end{aligned}$$

Let us derive (4.20). On one hand, by using assumption (H5) and since  $a > 1$ , we obtain

$$\gamma_n = \max_{1 \leq i \neq j \leq n} \left( \mathbb{E}[|\tilde{L}_i(x, \theta(x)) \mathbb{E} \tilde{L}_j(x, \theta(x))|] \right) = O\left(\frac{h_H \phi_x(h_K)}{n}\right).$$

Next, by using the fact that

$$\sum_{j \geq x+1} j^{-a} \leq \int_{u \geq x} u^{-a} = \left[ (a-1)x^{a-1} \right]^{-1} \quad (4.22)$$

we get

$$\sum_{j=m_n+1}^{\infty} \alpha(j) \leq \sum_{j=m_n}^{\infty} \alpha(j) \leq \int_{t \geq m_n} t^{-a} dt = \frac{m_n^{1-a}}{a-1},$$

thus,

$$\left( \sum_{j=m_n+1}^{\infty} \alpha(j) \right) \sum_{i=1}^n C_i = O \left( \frac{m_n^{1-a}}{a-1} \sqrt{\frac{n}{h_H \phi_x(h_K)}} \right).$$

We choose  $m_n = \left\lfloor \left( \frac{h_H \phi_x(h_K)}{n \log n} \right)^{1/(2(1-a))} \right\rfloor$ , where  $\lfloor \cdot \rfloor$  denotes the function integer part. It is clear that under assumption (H5),  $m_n \rightarrow \infty$ . In addition, if we replace  $m_n$  by its expression, we obtain

$$\sum_{j=m_n+1}^{\infty} \alpha(j) \sum_{i=1}^n C_i = O(\log n)^{-1/2} = o(1),$$

and again, under assumption (H5), we have

$$\begin{aligned} m_n \gamma_n &\leq C n^{-1-1/(2(1-a))} (h_H \phi_x(h_K))^{1+1/(2(1-a))} (\log n)^{-1/(2(1-a))} \\ &\leq n^{(-3+2a)/(2(1-a))} (h_H \phi_x(h_K))^{(3-2a)/(2(1-a))} (\log n)^{-1/(2(1-a))} \\ &\leq n^{-(2-3a)/(2(1-a))} (\log n)^{-1/(2(1-a))} = o(n^{-1}). \end{aligned}$$

Finally, Claim 1 can be easily deduced from (4.18)-(4.20) and Corollary 2.2 by Liebscher [24].

Proof of Claim 2. Following the same approach as the one used to show Claim 1, we show that

$$\mathbb{E} |T_{2,i} - \mathbb{E}[T_{2,i}]| \xrightarrow{L^1} 0 \text{ as } n \rightarrow +\infty, \quad (4.23)$$

and

$$\mathbb{E} |(T_{2,k} - 1)T_{2,i} - \mathbb{E}[(T_{2,k} - 1)T_{2,i}]| \rightarrow 0 \text{ as } n \rightarrow +\infty. \quad (4.24)$$

*Proof of (4.23).* In order to show (4.23), it suffices to prove the  $L^2$ -consistency of  $T_{2,i}$ . For this, we have

$$\begin{aligned} & \mathbb{E} \left[ (T_{2,i} - \mathbb{E}[T_{2,i}])^2 \right] \\ &= n \text{Var}[T_{2,1}] + 2 \frac{n\phi_x(h_K)\mathbb{E}^2[\beta_1^2 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \sum_{i \neq j} \text{Cov}(\beta_i K_i H_i^{(2)}, \beta_j K_j H_j^{(2)}). \end{aligned} \quad (4.25)$$

On one hand, we have

$$n \text{Var}[T_{2,1}] = \frac{n^2 \phi_x(h_K) \mathbb{E}^2[\beta_1 K_1]}{h_H \mathbb{E}^2[\Delta_1 K_1]} \left( \mathbb{E}[\beta_1^2 K_1^2 (H_1^{(2)})^2] - \mathbb{E}^2[\beta_1 K_1 H_1^{(2)}] \right).$$

Also, by using Lemma A.1 by Zhou and Lin [32] we prove that

$$\lim_{n \rightarrow +\infty} n \text{Var}[T_{2,1}] = o(1).$$

On the other hand, using exactly the same arguments as for showing (4.16), we show that the second term on the right hand side of (4.25) tends to 0 as  $n \rightarrow \infty$ . With this the result (4.23) is thus proved.

*Proof of (4.24).* Following the same steps as those used to get (4.11) and (4.6), we get

$$\mathbb{E}[(T_{2,k} - 1)^2] = \frac{1}{n \mathbb{E}^2[\beta_1 K_1]} \text{Var}[\beta_1 K_1] + 2 \text{Cov}(T_{2,k}, T_{2,l}).$$

The result (4.24) may then be directly obtained since, firstly

$$\frac{1}{n \mathbb{E}^2[\beta_1 K_1]} \text{Var}[\beta_1 K_1] \leq \frac{1}{n \mathbb{E}^2[\beta_1 K_1]} \mathbb{E}[\beta_1^2 K_1^2] = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

and secondly, by (4.11) and for  $c = 1$ ,  $j = 2$ , we get

$$\text{Cov}(T_{2,k}, T_{2,l}) \rightarrow 0 \text{ as } n \rightarrow +\infty.$$

■

**Proof of lemma 4.3.3.** Observe that for  $n$  large enough

$$\left| \widehat{f}_N^{(2)(x)}(\bar{\theta}(x)) - \widehat{f}_N^{(2)(x)}(\theta(x)) \right| \leq \sup_{y \in S} \left| \widehat{f}_N^{(2)(x)}(y) - f^{(2)(x)}(y) \right|.$$

By following the same ideas as those used by Laksaci et al. [22] in Lemma 7, we get

$$\sup_{y \in S} \left| \widehat{f}_N^{(2)(x)}(y) - f^{(2)(x)}(y) \right| \xrightarrow{P} 0 \text{ as } n \rightarrow +\infty. \quad (4.26)$$

So, the proof of this lemma is a consequence of (4.26). ■

### Acknowledgements

The authors would like to thank the Associate-Editor and the anonymous reviewers for their valuable comments and suggestions which improved substantially the quality of this paper.

# Bibliographie

- [1] Baïllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *J. of Multivariate Analysis*, **100**, 102–111.
- [2] Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *J. of Nonparametric Statist.*, **22**, 617–632.
- [3] Boj, E., Delicado, P. and Fortiana, J. (2010). Distance-based local linear regression for functional predictors. *Computational Statistics and Data Analysis*, **54**, 429–437.
- [4] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2018). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*, 1-35. <https://doi.org/10.1007/s41237-018-0057-9>
- [5] Bozebda, S., Chaouch, N. and Laïb, N. (2016). Limiting law results for a class of conditional mode estimates for functional stationary ergodic data. *Mathematical Methods of Statistics*, **25**, 168–195.
- [6] Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227–236.
- [7] Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data. *Statistica*, **70**, 83–94.
- [8] Dabo-Niang, S., Kaid, Z. and Laksaci, A. (2012). On spatial conditional mode estimation for a functional regressor. *Statist. Probab. Lett.*, **82**, 1413–1421.
- [9] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2011). *A fast functional locally modeled conditional density and mode for functional time-series*. Recent Ad-



- 
- vances in Functional Data Analysis and Related Topics, Contributions to Statistics, 2011, Pages 85–90, DOI : 10.1007/978-3-7908-2736-1\_13, Physica-Verlag/Springer.
- [10] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44.
- [11] Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the Local Linear Modelization of the Conditional Distribution for Functional Data. *Sankhya*, **76**, 328–355.
- [12] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2016). Estimation locale linéaire de la fonction de régression pour des variables hilbertiennes (in French). *C.R. Math. Acad. Sci. Paris*, **354**, 847–850.
- [13] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. (2017). Local linear regression modelization when all variables are curves. *Statist. Probab. Lett.*, **121**, 37–44.
- [14] El Methni, M. and Rachdi, M. (2011). Local weighted average estimation of the regression operator for functional data. *Commun. Stat., Theory Methods*, **40**, Issue 17, 3141–3153.
- [15] Ezzahrioui, M. and Ould-Saïd, E. (2008(a)). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. of Nonparametric Statist.*, **20**, 3–18.
- [16] Ezzahrioui, M. and Ould-Saïd, E. (2010). Some asymptotic results of a non-parametric conditional mode estimator for functional time-series data. *Statistica Neerlandica*, **64**, 171–201.
- [17] Fan, J., Gijbels, I.(1996). *Local Polynomial Modelling and its Applications*. London, Chapman and Hall.
- [18] Ferraty, F., Laksaci, A. and Vieu, F. (2005). Functional time series prediction via the conditional mode estimation. *C. R., Math., Acad. Sci. Paris* , **340**, 389–392.
- [19] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis : Theory and Practice*. Springer Series in Statistics, New York.
- [20] Ferraty, F. and Romain, Y. (2010). *The Oxford Handbook of Fuctional Data Analysis*, Oxford University Press.
-

- 
- [21] Khardani, S., Lemdani, M. and Ould-Saïd, E. (2011). Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series. *J. Statist. Plann. Inference*, **141**, 3426–3436.
- [22] Laksaci, A., Rachdi, M. and Rahmani, S. (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1–23.
- [23] Laksaci, A., Nacéri, A. and Rachdi, M. (2015). Exact Quadratic Error of the Local Linear Regression Operator Estimator for Functional Co-variates. *Functional Statistic and Application*, Springer, **2**, 79–90.
- [24] Liebscher, E. (2001). Central limit theorems for  $\alpha$ -mixing triangular arrays with applications to nonparametric statistics. *Mathematical Methods of Statistics*, **10**, 194–214.
- [25] Louani, D. and Ould-saïd, E. (2007). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. of Nonparametric Statist.*, **11**, 413–442.
- [26] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155–177.
- [27] Ouassou, I. and Rachdi, M. (2012). Regression operator estimation by delta-sequences method for functional data and its applications. *Adv. Statist. Anal.*, **96**, 451–465.
- [28] Ould-Saïd, E. (1993). Estimation nonparamétrique du mode conditionnel. Application à la prévision (in French). *C. R. Acad. Sci. Paris, Série I*, **316**, 943–947.
- [29] Ould-Saïd, E. (1997). A note on ergodic processes prediction via estimation of the conditional mode function. *Scandinavian J. Statist.*, **24**, 231–239.
- [30] Ould-Saïd, E. and Tatachak, A. (2011). A nonparametric conditional mode estimate under RLT model and strong mixing condition. *Int. J. Stat. Econ.*, **6**, 76–92.
- [31] Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A., Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data, *Comput. Statist. Data Anal.*, **73**, 53–68.
- [32] Zhou, Z. and Lin, Z.-Y. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *J. of Nonparametric Statist.*, **28**, 116–131.

# Annexe : Quelques outils de probabilités

Dans ce chapitre, nous allons rappeler quelques définitions et théorèmes utilisés dans ce manuscrit.

## 4.5 Notions de convergence

**Définition 4.1.** (Définition A.1.[38]).

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles définie sur un espace de probabilité  $(\Omega; \mathcal{A}; \mathbb{P})$ . On dit que la suite  $(X_n)_{n \in \mathbb{N}}$  converge presque complètement ( $p, co.$ ) vers  $X$  une variable aléatoire réelle définie sur le même espace, si et seulement si si

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

et nous écrivons  $X_n \xrightarrow[n \rightarrow +\infty]{p, co.} X$ .

**Définition 4.2.** (Définition A.3.[38]).

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles,  $(u_n)_{n \in \mathbb{N}}$  une suite de nombres réelles. On dit que  $X_n = O(u_n)$  en  $p, co.$  si est seulement

$$\forall \epsilon > 0, \sum_{n=0}^{\infty} \mathbb{P}(|X_n| > \epsilon u_n) < \infty.$$

**Définition 4.3.** (Proposition A.5.[38]).

Supposons que  $\lim_{n \rightarrow +\infty} u_n = 0$ ,  $X_n = O_{p, co.}(u_n)$  et  $\lim_{n \rightarrow +\infty} Y_n = l$  ( $p, co.$ ) où  $l$  est un nombre réel. Nous avons :

$$\text{I } X_n Y_n = O_{p,co.}(u_n).$$

$$\text{II } \frac{X_n}{Y_n} = O_{p,co.}(u_n), \text{ à condition que } l \neq 0.$$

**Théorème 4.1.** [25].

Soit  $(X_n)_{n \in \mathbb{N}}$  et  $(Y_n)_{n \in \mathbb{N}}$  deux suites de variables aléatoires définies sur un espace probabilisé  $(\Omega; \mathcal{A}; \mathbb{P})$  et  $X$  une variable aléatoire.

Si  $X_n$  converge en loi vers  $X$ , et si  $Y_n$  converge en probabilité vers une constante  $c$  alors

$$\text{I } X_n + Y_n \xrightarrow[n \rightarrow +\infty]{D} X + c.$$

$$\text{II } X_n \times Y_n \xrightarrow[n \rightarrow +\infty]{D} c \times X.$$

$$\text{III } \frac{X_n}{Y_n} \xrightarrow[n \rightarrow +\infty]{D} \frac{X}{c}, \text{ avec } c \neq 0.$$

## 4.6 Quelques inégalités utiles

**Théorème 4.2.** (Jensen)[70].

Soient  $X$  une variable aléatoire réelle et  $\varphi$  une fonction convexe. Alors

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X)).$$

**Théorème 4.3.** (Markov)[46].

Soient  $X$  une variable aléatoire réelle. Alors, pour tout  $a > 0$ ,

$$\mathbb{P}(|X| > a) < \frac{\mathbb{E}(|X|)}{a}.$$

**Théorème 4.4.** (Bienaymé-Tchebychev)[8].

Soient  $X$  une variable aléatoire réelle. Alors, pour tout  $a > 0$  :

$$\mathbb{P}(|X - \mathbb{E}(X)| > a) < \frac{\text{Var}(X)}{a^2}.$$

**Théorème 4.5.** (Inégalité de Hölder)[1].

Soit  $X$  et  $Y$  deux variables aléatoires telles que  $X \in L^p(\Omega; \mathcal{A}; \mathbb{P})$  et  $Y \in L^q(\Omega; \mathcal{A}; \mathbb{P})$  avec

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} \text{ et } p \geq 1, q \geq 1, \text{ Alors :}$$

$$(\mathbb{E}(|XY|^{\frac{1}{r}})) \leq \mathbb{E}(|X|^p)^{\frac{1}{p}} \mathbb{E}(|Y|^q)^{\frac{1}{q}}.$$

Si  $p = 2$  et  $q = 2$ , nous obtenons L'inégalité de Cauchy-Schwarz.

**Théorème 4.6.** (Inégalité de Minkowski)[1].

Soit  $X$  et  $Y$  deux variables aléatoires telles que  $X$  et  $Y \in L^p(\Omega; \mathcal{A}; \mathbb{P})$  avec  $p > 1$ . Alors,  $X + Y \in L^p(\Omega; \mathcal{A}; \mathbb{P})$  et,

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

**Définition 4.4.** Soit  $(X_i)_{i \geq 1}$  une suite de variables aléatoires. Pour tout entier naturel non nul  $n$ , on définit le coefficient d'alpha-mélange par :

$$\alpha(n) = \sup_{k > 0} \{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A}_1^k(X) \text{ et } B \in \mathcal{A}_{n+k}^\infty(X) \}.$$

$\mathcal{A}_i^k(X)$  désigne la tribu des évènements engendrés par les  $X_j, ; i \leq j \leq k$ . La suite est dite  $\alpha$ -mélangeante ou fortement mélangeante si le coefficient d'alpha-mélange vérifie :  $\alpha(n) \rightarrow 0$ , quand  $n \rightarrow \infty$ .

**Théorème 4.7.** Inégalités de covariance (Proposition A.10.[38]).

Soit  $(X_n)_{n \in \mathbb{Z}}$  une famille de variables aléatoires  $\alpha$ -mélangeantes. Étant donnée la variable aléatoire  $X$  (resp.  $Y$ ) qui, pour tout  $k \in \mathbb{Z}$ , est  $\sigma(X_i, -\infty < i < k)$ -mesurable (resp.  $\sigma(X_i, n + k < i < +\infty)$ -mesurable).

I Si  $X$  et  $Y$  sont bornées, alors :

$$\exists C > 0, \text{ Cov}(X, Y) \leq C\alpha(n).$$

II Si pour quelques nombres réels positive  $p, q$  et  $r$  tel que :  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ , on a

$\mathbb{E}(X^p) < \infty$  et  $\mathbb{E}(Y^p) < \infty$ , alors :

$$\exists C > 0, \text{ Cov}(X, Y) \leq \mathbb{E}(|X|^p)^{\frac{1}{p}} \mathbb{E}(|Y|^p)^{\frac{1}{p}} \alpha(n)^{\frac{1}{r}}.$$

**Lemme 4.1.** (Volkonskii et Rozanov 1959)[36].

Soit  $V_1, \dots, V_m$  une suite de variables aléatoires fortement ou  $\alpha$ -mélangeante mesurables par rapport aux tribus  $(\mathfrak{F}_{i_1}^{j_1}, \dots, \mathfrak{F}_{i_m}^{j_m})$  avec  $1 \leq i_1 < j_1 < \dots < j_m \leq n$ ,

$i_{l+1} - j_l \geq w \geq 1$  et  $|V_k| \leq 1$  pour tout  $l, k = 1, 2, \dots, m$ , alors :

$$\left| \mathbb{E} \prod_{j=1}^m V_j - \prod_{j=1}^m \mathbb{E}(V_j) \right| \leq 8(m-1)\alpha(w),$$

où  $\mathfrak{F}_a^b = \sigma\{V_i, a \leq i \leq b\}$  et  $\alpha(w)$  est le coefficient de mélange.

## 4.7 Théorème central limite

**Théorème de Lindeberg [16] :**

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires définies sur le même espace de probabilité, indépendantes. Supposons que, pour  $n \geq 1$ ,  $X_n$  ait une espérance finie  $\mu_n$  et un écart-type fini  $\sigma_n$ , et posons  $S_n^2 = \sum_{i=1}^n \sigma_i^2$ ,  $Z_n = \frac{1}{S_n} \sum_{i=1}^n (X_i - \mu_i)$ .

Si, pour tout  $\epsilon > 0$ ,  $\lim_{n \rightarrow +\infty} \frac{1}{S_n^2} \sum_{i=1}^n \mathbb{E} \left[ (X_i - \mu_i)^2 1_{\{|X_i - \mu_i| > \epsilon S_n\}} \right] = 0$ , alors la loi de  $Z_n$

converge vers une loi normale centrée réduite  $\mathcal{N}(0, 1)$ .

**Théorème de Liebscher (2001) :**

Soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de variables aléatoires centrées,  $\alpha$ -mélangeantes, et

$T_n = \sum_{i=1}^n X_i$ . Notons  $\Gamma_n := \max_{i \in \mathbb{N}^*} \mathbb{E}(X_i^2)$  et  $\gamma_n := \max_{1 \leq i, 1 \leq j, i \neq j} \text{Cov}(X_i, X_j)$ .

**Condition  $C(p)$  :** Supposons,  $2 < p < \infty$  :  $\Gamma_n = O(n^{-1})$ ,  $\sum_{k=1}^{\infty} k^{1-2/p} \alpha(k) < \infty$ , et

$\mathbb{E}(|X_i^p|)^{1/p} < \infty$  pour tout  $i \in \mathbb{N}$ . Il existe une suite de nombres réels positifs  $m_n$  qui tend vers  $\infty$  telles que :

$$n m_n \gamma_n = o(1), \text{ et } \left( \sum_{j=m_n+1}^{\infty} j^{2/p-2} \alpha(j) \right)^{1-2/p} \sum_{i=1}^n \mathbb{E}(|X_i^p|)^{2/p} = o(1).$$

**Théorème 4.8 ([61]).** .

Supposons que la condition  $C(p)$  est vérifiée pour un certain  $p > 2$  et qu'il existe une suite de nombres réels positifs  $\tau_n$  avec  $\tau_n = o(\sqrt{n})$ , tel que

$$\sum_{i=1}^n \mathbb{E}(X_i^2) \mathbb{1}_{(|X_i| > \tau_n^{-1})} = o(1),$$

et

$$\forall \epsilon > 0, \quad \frac{n}{\tau_n} \alpha(\epsilon \tau_n) \xrightarrow{n \rightarrow +\infty} 0.$$

De plus, supposons que

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E}(X_i^2) = \sigma^2,$$

alors

$$T_n \xrightarrow{D} N(0, \sigma^2).$$

# Conclusion et perspectives

Dans ce travail, nous nous sommes intéressés aux aspects théoriques et pratiques de l'estimation non paramétrique des paramètres fonctionnels conditionnels comme la fonction de répartition conditionnelle, la densité conditionnelle et ses dérivées ainsi que le mode conditionnel ; quand la variable explicative est fonctionnelle en utilisant la méthode locale linéaire. Notons que la modélisation locale linéaire est une approche alternative à l'estimation de Nadaraya-Watson (NW), qui a plus d'avantages sur cette dernière (voir [31]). En particulier, le plus grand avantage de la méthode des polynômes locaux sur la méthode du noyau est la réduction considérable du biais de l'estimateur.

Comme résultat asymptotique, nous établissons la normalité asymptotique de ces estimateurs, dans le cas où les observations sont indépendantes et identiquement distribuées ainsi que le cas dépendant. Cela a été réalisé grâce aux hypothèses qui sont détaillées dans les Chapitres 2, 3, 4 et l'usage de quelques outils probabilistes et les théorèmes central limite cités dans l'annexe 4.4. Notons que cette propriété asymptotique est un sujet très important en statistique. Elle est utilisée pour la construction des intervalles de confiance ou de faire des tests statistiques comme les tests de normalité.

Il est à noter que l'originalité de cette étude est aussi dans la liste des perspectives qu'elle offre. À cet objectif nous présentons ci-dessous plusieurs projets de recherche qui se situent dans la continuité des études effectuées dans le cadre de cette thèse.

## Estimation des constantes

Dans notre contribution nous avons estimé les constantes  $M_1$  et  $M_2$  qui interviennent dans les termes dominants du biais et de variance par la méthode plug-in. Il est possible d'utiliser une autre méthode par exemple les techniques de bootstrap.

## Méthode d'estimation par polynômes locaux

Comme il est mentionné par Fan et Gigbels [17], si on augmente le degré du polynôme on réduit le biais de l'estimateur, cette réduction peut être plus importante si l'on considère une approximation par un polynôme des degrés supérieurs à un, notamment entre deux polynômes de degré consécutifs, celui à degré impair constituera un meilleur estimateur, au sens du biais.

## Choix du paramètre de lissage

Le paramètre de lissage est un paramètre crucial en estimation non paramétrique car ce paramètre intervient dans toutes les propriétés asymptotiques qu'on a étudié. La sélection des paramètres de lissage est un autre sujet de recherche. Nos résultats asymptotiques constituent une étape préliminaire indispensable permettant d'envisager cette perspective de recherche, dans les deux cas déterministe et le cas de l'uniformité de la largeur de fenêtre.

## Choix de la semi-métrique

Il est clair que le choix de la semi-métrique joue un rôle déterminant notamment dans l'amélioration des vitesses de convergence en augmentant la concentration de la mesure. À notre connaissance il n'existe aucune méthode automatique qui offre un choix optimal de la semi-métrique.

## Normalité asymptotique

Il est bien possible de généraliser nos résultats au cas de données incomplètes (tronquées, censurées), ergodiques ainsi qu'à des données spatiales.



---

## Bibliographie générale

Attouch, M. Laksaci, A. and Ould-Saïd E. (2010). Asymptotic normality of a robust estimator of the regression function for functional time series data. *J. of the Korean Statist. Soc.*, **39**, 489-500.

Baïllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102-111.

Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally Modelled Regression and Functional Data. *J. of Nonparametric Stat.*, **22**, 617-632.

Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data. *Computational Statistics*, **22**, 353-369.

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

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

Charlot, F. and Rachdi, M. (2008). On the statistical properties of a stationary process sampled by a stationary point process. *Stat. Probab. Lett.*, **78**, 456-462.

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

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

Dabo-Niang, S. and Laksaci, A. (2008). Conditional mode regression : Application to functional time series prediction. *Electron. J. Statist.*, ISSN : 1935-7524.

Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

Dabo-Niang, S., Rachdi, M. and Yao, A.-F. (2010). Spatial kernel regression estimation

- and prediction for functional random fields. *Far East J. Theor. Stat.*, **37**, 77-113.
- Dabo-Niang, S. and Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Stat. Probab. Lett.*, **80**, 1447-1458.
- Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Commun. in Statist.-Theory and Meth.*, **41**, 1254-1268.
- Dabo-Niang, S., Kaid, Z. and Laksaci, A. (2012) On spatial conditional mode estimation for a functional regressor. *Stat. Probab. Lett.*, **82**, 1413-1421.
- Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional time series analysis. *Statistics*, **43**, 13-33.
- Demongeot, J., Laksaci, A., Madani, F. et Rachdi, M. (2011). Estimation locale linéaire de la densité conditionnelle pour des données fonctionnelles. *C. R. Acad. Sci. Paris, Ser. I*, **348**, 931-934.
- Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics*, **47**, 26-44.
- Ezzahrioui, M. and Ould-Saïd, E. (2008(a)). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. of Nonparametric Stat.*, **20**, 3-18.
- Ezzahrioui, M. and Ould-Saïd, E. (2008b). Asymptotic normality of the kernel estimator of conditional quantiles in a normed space. *Far East J. Theoretical Statist.*, **25**, 15-38.
- Ezzahrioui, M. and Ould-Saïd, E. (2008c). Asymptotic results of a nonparametric conditional quantile estimator for functional time series *Commun. in Statist.-Theory and Meth.*, **37**, 2735-2759.
- Fan, J. (1992). Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998-1004.
- Fan, J., Yao, Q. and Tong, H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika*, **83**, 189-206.

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

Fan, J. and Yim, T.-H. (200). A cross-validation method for estimating conditional densities. *Biometrika*, **91**, 819-834.

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

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

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

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

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Stat. Plann. and Inf.*, **140**, 335-352.

Ferraty, F., Van Keilegom, I. and Vieu, P. (2010). On the Validity of the Bootstrap in Non-Parametric Functional Regression. *Scandinavian Journal of Statistics*, **37**, 286-306.

Gasser, T., Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. *J. R. Statist. Soc. B*, **60**, 681-691.

Gheriballah, A., Laksaci, A. and Rouane, R. (2010). Robust nonparametric estimation for spatial regression. *J. Stat. Plann. and Inf.*, **140**, 1656-1670.

Hallin, M., Lu, Z. and Tran, L.T. (2004). Local linear spatial regression. *Ann. of Statist.*, **32**, 2469-2500.

Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression. *Bernoulli*, **15**, 659-686.

Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and semi-parametric models : An introduction*. Springer Series in Statistics. New York.

Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Pub. Inst. Stat. Univ. Paris*, **3**, 69-80.

Laksaci, A. et Maref, F. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *C. R. Acad. Sci. Paris, Sér. I*, **347**, 1075-1080.

Laksaci, A. et Mechab, M. (2010). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle cas des données spatiales. *Rev. Roumaine, Math Pures Appl.*, **55**, 35-51.

Laksaci, A., Madani, F. and Rachdi, M. (2013). Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Commun in Statist.-Theory and Meth.*, **42**, 3544-3570.

Laksaci, A., Rachdi, M. and Rahmani, S. (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spatial Statistics*, **6**, 1-23.

Li, J. and Tran, L.T. (2009). Nonparametric estimation of conditional expectation. *J. Stat. Plann and Inf.*, **139**, 164-175.

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

Lu, Z. and Chen, X. (2004). Spatial kernel regression : weak consistency. *Stat. Probab. Lett.*, **68**, 125-136.

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

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

Ould-Saïd, E. (1993). Estimation nonparamétrique du mode conditionnel. Application à la prévision. *C. R. Acad. Sci. Paris., Série I.*, **316**, 943-947.

Ould-Saïd, E. (1997). A note on ergodic processes prediction via estimation of the conditional mode function. *Scandinavian Journal of Statistics*, **24**, 231-239.

Ould-Saïd, E. and Cai, Z. (2005). Strong uniform consistency of nonparametric estimation of the censored conditional mode function. *J. of Nonparametric Stat.*, **17**, 797-806.

Quintela-del-Río, A. and Vieu, P. (1997). A nonparametric conditionnal mode estimate. *J. of Nonparametric Stat.*, **8**, 253-266.

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

Ramsay, J.O. and Silverman, B.W. (2002). *Functional data analysis*. Springer Series in Statistics. New York.

Samanta, M. and Thavaneswaran, A. (1990). Nonparametric estimation of conditional mode. *Commun. in Statist.-Theory and Meth.*, **16**, 4515-4524.

Sarda, P. and Vieu, P. (2000). *Kernel Regression*. Wiley, New York. 43-70

Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*. Springer-Verlag, Berlin.

Stone, C.J. (1977). Consistent nonparametric regression. *Ann. of Statist.*, **5**, 595-620.

Tran, L.T. (1990). Kernel density estimation on random fields. *J. of Multivariate Anal.*, **34**, 37-53.

Xu, R. and Wang, J.(2008).  $L_1$ -estimation for spatial nonparametric regression. *J. of Nonparametric Stat.*, **20**, 523-537.

Yu, K. and Jones, M.C. (1997). A comparison of local constant and local linear regression quantile estimators. *Computational Statistics and Data Analysis*, **25**, 159-166.

Volkonskii, Yu.A. Rozanov. (1959). Some limit theorems for random functions, *Theory Probab. Appl.*, **4**, 178-197.

Zhou, Z. and Lin, Z.-Y. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *Nonparametric Stat.* **28**, 116-131.

## ملخص

المشكلة التي نتناولها في هذه الأطروحة هي الحالة الطبيعية المتقاربة لنماذج شرطية ذات متغير عشوائي دالي و متغير الاستجابة حقيقياً، باستعمال الطريقة الخطية المحلية. لاحظ أن فائدة النتيجة لهذه العملية الخاصة بالمقاربة هي بناء فترات التنبؤ أو التحقق من الاختبارات الإحصائية. في البداية نعتبر متتالية من العينات المستقلة و تتبع نفس القانون. وضعنا الهدف المتمثل في إنشاء الحالة الطبيعية للمقدر محلي خطي لدالة التوزيع الشرطي والكثافة الشرطية ومشتقاتها وكذلك الوضع الشرطي. نقدم أيضاً تطبيق النتائج التي تم الحصول عليها على المحاكاة البيانات و البيانات الحقيقية. في المرحلة الثانية نعمم نتائجنا في حالة سلسلة من الملاحظات المرتبطة بنوع المعامل  $\alpha$ . نقدر وظيفة التوزيع الشرطي بالطريقة المحلية الخطية. بعد ذلك ندرس الحالة الطبيعية المتقاربة لهذا المقدر في ظل افتراضات تركيز القانون المشترك للملاحظات والفرضيات على معاملات المعامل  $\alpha$ . يمكن استخدام هذه النتيجة لمشكلة التنبؤ بالسلاسل الزمنية. في الجزء الثالث من هذه الأطروحة ، نعتبر نفس النوع من الملاحظات كما في الحالة السابقة و نهتم على تقدير الوضع الشرطي بالطريقة المحلية الخطية و قدمنا في ظل ظروف معينة ، خاصة. إحصائية تتعلق الحالة الطبيعية المتقاربة لهذا المقدر. كلمات مفتاحية : النماذج الشرطية ، والتركييب متعدد الحدود المحلي ، تحليل البيانات الوظيفية (FDA) ، التقدير اللامعلي ، احتمال الكرة الصغيرة.

## Résumé

La problématique abordée dans cette thèse concerne la loi limite des estimateurs localement linéaires de certains modèles conditionnels, lorsque la variable explicative est de type fonctionnel et la variable réponse est réelle. Notons que la conséquence pratique de cette propriété asymptotique est la construction des intervalles de confiance où la réalisation des tests statistiques.

Nous supposons pour commencer que l'échantillon que nous étudions est constitué de variables indépendantes et identiquement distribuées. Nous avons fixé comme objectif l'établissement de la normalité asymptotique des estimateurs localement linéaires de la fonction de répartition conditionnelle, la densité conditionnelle et ses dérivées ainsi que le mode conditionnel. Les résultats obtenus sont également appliqués sur des données simulées puis réelles.

Dans la deuxième partie de ce manuscrit, nous généralisons nos résultats dans le cas où les observations sont fortement mélangées. Nous estimons la fonction de répartition conditionnelle par la méthode locale linéaire. Ensuite, nous étudions la normalité asymptotique de cet estimateur sous des hypothèses de concentration de la loi conjointe des observations et des hypothèses sur les coefficients de mélange. Ce résultat peut être utilisé pour le problème de la prévision en série chronologique.

Dans la troisième partie de cette thèse, nous considérons le même type de dépendance des observations que le cas précédent, en s'intéressant à l'estimation du mode conditionnel par la méthode locale linéaire, pour lequel nous avons établi sous certaines conditions, une propriété statistique qui concerne la loi limite de cet estimateur.

Mots clés: Modèles conditionnels, polynômes locaux, analyse de données fonctionnelles (FDA), estimation non paramétrique, probabilité de petite boules.

## Abstract

The problem addressed in this thesis concerns the limit law of locally linear estimators of some conditional models, when the explanatory variable is of functional nature and the response variable is real. Note that the practical consequence of this asymptotic property is the construction of confidence intervals or statistical tests. First, we consider a sequence of independent and identically distributed observations. We establish the asymptotic normality of the conditional distribution function estimators and the conditional density and its derivatives as well as the conditional mode. The established results were applied on both simulated and real data.

In the second part of this manuscript, we generalize our results in cases where the observations are strongly mixed. We estimate the conditional distribution function by the linear local method. Next, we study the asymptotic normality of this estimator under concentration assumptions of the joint law of observations  $(X_i, X_j)$  and assumptions about the mixing coefficients. This result can be used for the time series prediction problem.

In the third and last part of this thesis, we consider the same type of dependence of the observations as the previous case, by focusing on the estimation of the conditional mode by the linear local method, for which we have established under certain conditions, the limit law of this estimator.

Key words: Conditional models, local polynomial fitting, functional data analysis (FDA), nonparametric estimation, small ball probability.

