

UNIVERSITÉ KASDI MERBAH  
OUARGLA



Faculté des Mathématiques et des Sciences  
de la Matière

Département de Mathématiques



**THÈSE DE DOCTORAT 3<sup>ème</sup> Cycle Année : 2022/2023**

Présenté en vue de l'obtention du diplôme de  
Doctorat en Mathématiques

**Estimation non paramétrique de quelques fonctions robustes  
par la méthode des  $k$  plus proches voisins**

**Spécialité :**

Probabilités et Statistique

**Présenté par :**

Belatreche Nadjat

**Soutenue le 30 Mai 2024 Devant le jury :**

PRÉSIDENT :	Meflah Mabrouk	Prof.	UKM Ouargla, Algérie
DIRECTEUR DE THÈSE :	Attouch Mohammed Kadi	Prof.	U Djillali Liabès, SBA, Algérie
CO-DIRECTEUR :	Bouabsa Wahiba	MCA	U Djillali Liabès, SBA, Algérie
EXAMINATEURS :	Chikhi Mohamed	Prof.	UKM Ouargla, Algérie
	Ghazal Abderrazak	Prof.	UKM Ouargla, Algérie
	Saouli Mostapha Abdelouahab	MCA.	UKM Ouargla, Alger, Algérie
INVITÉ :	Keddi Abdelmalik	MCB.	U.A.E.H.M, Tamanrasset, Algérie

## *Dédicace*

*À ma mère et mon père, à qui je dédie ce travail, qui ont consacré leurs efforts pour que je sois persévérant et sérieux dans mes études et mon travail, leur désir est de me voir à des niveaux élevés dans ce monde et dans l'au-delà.*

*À celui sur qui je compte dans chaque situation, grande ou petite, mon mari et mes frères et sœurs.*

*À mon petit ange, ma fille " Malak Alrahman"*

*À mes amis et connaissances que j'admire et respecte.*

*À ceux qui aiment la science et l'apprentissage.*

# Remerciements

Avant tout, nous voulons exprimer notre gratitude à **ALLAH**, le Tout-Puissant et le Plus Miséricordieux, qui nous a fourni la persévérance et la force nécessaires pour accomplir cette humble tâche.

Cette thèse est dédiée au professeur **BAHADI Aissa** (qu'Allah lui fasse miséricorde), il a tous les éloges pour ce que nous avons atteint aujourd'hui.

À mon superviseur, je voudrais dire merci beaucoup, le professeur **Attouch Mohamed Kadi**, pour ses conseils et son soutien constant, et pour m'avoir donné de nombreuses idées fructueuses sur mon sujet.

Je tiens également à remercier profondément mon co-encadrant, le Dr. **Bouabsa Wahiba**, pour tous les conseils que j'ai eus pendant mon doctorat. Je la remercie également pour sa patience et son attention, étant toujours disponible pour discuter et répondre à mes questions.

Honnêtement, j'ai beaucoup appris d'eux. Qu'Allah le récompense abondamment.

Je suis très reconnaissant au président du jury, Prof. **Mabrouk Meflah** et aux examinateurs : Profs. **Chikhi Mohamed**, **Ghazal Abderrazak**, **Saouli Mostapha Abdelouahab**, **Keddi Abdelmalik** pour avoir accepté de rendre compte de cette thèse et pour leurs précieux commentaires et suggestions. Remerciements particuliers au Dr. **Keddi Abdelmalik**, qui m'a aidé à choisir le superviseur. Qu'Allah les récompense tous abondamment. Je remercie du fond du cœur mes professeurs qui m'ont consacré leurs efforts durant mes études.

Enfin, mon affectueuse gratitude et ma reconnaissance vont à mes **Parents** pour leurs encouragements constants, et leur patience tout au long de mes années scolaires. Qu'Allah les récompense tous abondamment.



---

## Liste des travaux

### I. Publications dans des revues internationales à comité de lecture :

- 1- Belatreche N., Bouabssa W., Attouch M. K., and Fetitah O. M-Regression Estimation with the  $k$  Nearest Neighbors Smoothing under Quasi-associated Data in Functional Statistics. Rev. Applications and Applied Mathematics : An International Journal (AAM). Vol. 17, Issue 2 (December 2022), 333 – 365. Lien de l'article : <https://digitalcommons.pvamu.edu/aam/vol17/iss2/4>
- 2- Belatreche N., Bouabssa W., Attouch M. K. The converge rate of the hazard function with function explanatory variable : case of spacial data with the  $k$  Nearest Neighbor method. Rev. Turkish Journal of Computer and Mathematics Education. Vol. 14, Issue 2 (8 April 2023), 180 – 202. Lien de l'article : <https://turcomat.org/index.php/turkbilmomat/article/view/13636/9783>
- 3- Belatreche N., Bouabssa W., Attouch M. K., and Fetitah O.  $k$ NN Asymptotic normality of a robust kernel estimator of the regression function for functional ergodic data in the case of unknown scale parameter. Soumis pour publication.
- 4- Belatreche N., Bouabssa W., Attouch M. K.  $k$ NN asymptotic normality of functional nonparametric kernel estimator of  $\sigma_z$ . Soumis pour publication.

### II. Communications :

#### 1- Communications internationales

- i. Non parametric estimation with  $k$  nearest neighbors method. The first International Conference on Pure and Applied Mathematics, IC-PAM'21, May 26-27, 2021, Ouargla, Algeria.
- ii. The M-Regression in functional statistics with the  $k$  nearest neighbor method. National Symposium on Mathematics And Applications Department Of Mathematics Indian Institute Of Technology Madras, 22nd December 2021. IIT Madras, India.

- 
- iii. Estimation data-driven using the  $k$ -nearest neighbors method in nonparametric function data analysis. International E-Conference On Pure And Applied Mathematical Sciences (ICPAMS-2022),04-07 May, 2022. Arabie Saoudite et Tunisie.
  - iv. Non parametric relative error estimation through functional regressor under truncation random data with the  $k$  nearest neighbors method. The 2nd International Conference on Applied Mathematics, Informatics, and Computing Sciences (AMICS 2023) December 05-07, 2022. Bosen Academic Research Institute, China.

## 2- Communications nationales

- i. Nonparametric estimation using  $k$  Nearest Neighbors method for missing data at random. National Conference on Applied Mathematics and Didactics NCAMD 2021, juin 26, 2021. Ecole Normale Supérieure El Katiba Assia Djebar Constantine.
- ii. The uniform in bandwidth consistency for different kernel estimators that include functional data. 1<sup>ère</sup> édition du Séminaire National Mathématiques et Informatique appliquée aux Sciences 11-12 Mai 2022, Tamanrasset, Algérie.
- iii. Nonparametric relative error estimation using a functional regressor by the  $k$  Nearest Neighbors smoothing. National Conference on Mathematics and Applications NCMA2022, 29-30 November 2022, Abdelhafid Boussouf University, Mila.
- iv. Nonparametric relative error estimation under truncation random data with the  $k$  Nearest Neighbors method. La conférence Nationale Nouvelles Tendances en Mathématiques Théoriques et Computationnelles. 08-09 Novembre 2022, Tamanrasset, Algérie.

## Abréviations et Symboles

Les notations suivantes seront utilisées dans tous les différents chapitres de cette thèse.

### Abréviations

$NFDA$	Nonparametric Functional Data Analysis,
$i.i.d.$	Indépendantes et identiquement distribuées,
$v.a.$	Variable aléatoire,
$(\Omega, \mathcal{A}, \mathbb{P})$	Espace probabilisé,
$NW$	Nadaraya- Watson,
$kNN$	$k$ plus proches voisins ( $k$ Nearest Neighbors),
$MSE$	Erreur quadratique moyenne (Mean Squared Error),
$p.s$	Prèsque sûre,
$p.co$	Prèsque complète,
$s.o$	Statistique d'ordre,
$\xrightarrow{l}$	Converge en loi,
$\xrightarrow{D}$	Converge en distribution,
$\xrightarrow{p}$	Converge en probabilité,
$\xrightarrow{p.s}$	Converge presque sûre,
$O_P(\cdot)$	Être borné en probabilité .
$o_P(\cdot)$	Converge vers 0 en probabilité,

---

## Symboles

$X$	Variable aléatoire fonctionnelle,
$Y$	Variable aléatoire réelle,
$X_{1,n}, \dots, X_{n,n}$	Statistique d'ordre associées à $X_1, \dots, X_n$ ,
$H$	Espace de Hilbert,
$L^p$	Les espaces de fonctions de puissance $p$ -ième intégrable, avec $1 \leq p < \infty$ ,
$\mathbb{N}$	Ensemble des entiers naturels,
$\mathbb{R}$	Ensemble des nombres réels,
$\mathbb{B}_{\mathbb{E}}$	La tribu borélienne,
$N(0, 1)$	La distribution de la loi Normale de moyenne 0 et d'écart-type 1,
$\mathbb{I}_A$	Fonction indicatrice de l'ensemble $A$ qui vaut 1 sur l'ensemble $A$ et 0 ailleurs,
$\mathbb{P}(A)$	Probabilité d'un évènement $A$ ,
$B(x, h)$	La boule de centre $x$ et de rayon $h$ ,
$\langle \cdot, \cdot \rangle$	Produit scalaire définie dans $H$ ,
$d(\cdot, \cdot)$	Semi-métrique,
$a \vee b$ resp. $a \wedge b$	Désignent le maximum resp. minimum des nombres $a, b \in \mathbb{R}$ ,
$\mathbb{E}(X)$	Espérance mathématique de la variable $X$ ,
$\text{Var}(X)$	Variance de la variable $X$ ,
$\text{Cov}(X, Y)$	Covariance des variables $X$ et $Y$ ,
$:=$	La définition d'une quantité,
$a_n = o(b_n), n \rightarrow \infty$	$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 0$ avec $a_n$ et $b_n$ deux suites réelles,
$a_n = O(b_n), n \rightarrow \infty$	$\lim_n \sup \left  \frac{a_n}{b_n} \right  < \infty$ avec $a_n$ et $b_n$ deux suites réelles,
$a_n \sim b_n, n \rightarrow \infty$	$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = c$ avec $a_n$ et $b_n$ deux suites réelles et $c$ une constante,
$\lim_{n \rightarrow \infty} X_n = X$ p.	$\Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{P}( X_n - X  > \epsilon) = 0$ ,
$\lim_{n \rightarrow \infty} X_n = X$ p.s.	$\Leftrightarrow \mathbb{P}(\lim_{n \rightarrow \infty} X_n = X) = 1$ ,
$\lim_{n \rightarrow \infty} X_n = X$ p.co.	$\Leftrightarrow \forall \epsilon > 0, \sum_{n \in \mathbb{N}} \mathbb{P}( X_n - X  > \epsilon) < \infty$ ,
$X_n - X = O_{p.co.}(u_n)$	$\Leftrightarrow \exists \epsilon_0 > 0, \sum_{n \in \mathbb{N}} \mathbb{P}( X_n - X  > \epsilon_0 u_n) < \infty$ .

# TABLE DES MATIÈRES

Liste des travaux . . . . .	ii
Abréviations et Symboles . . . . .	iv
<b>1 Introduction</b>	<b>1</b>
1.1 Contexte bibliographique sur statistique non paramétrique fonctionnelle .	1
1.2 Les données fonctionnelles . . . . .	2
1.2.1 Contexte bibliographique sur les données fonctionnelles . . . . .	2
1.2.2 Variables et données fonctionnelles . . . . .	4
1.2.3 Domaines d'applications des données fonctionnelles . . . . .	6
1.2.4 Quelques exemples de données fonctionnelles . . . . .	8
1.3 Données spatiale . . . . .	9
1.3.1 Présentation des données spatiales . . . . .	9
1.3.2 Données fonctionnelles spatiales . . . . .	10
1.4 Contexte bibliographique sur l'estimation par la méthode des k plus proches voisins . . . . .	10
1.5 Robustesse en statistique non paramétrique . . . . .	12
1.6 Probabilités de petites boules et semi-métriques . . . . .	13
1.7 Estimation de la fonction de hasard . . . . .	14
1.8 Les méthodes d'estimation . . . . .	16
1.8.1 Le régressogramme . . . . .	16
1.8.2 L'estimateur de la fenêtre mobile . . . . .	16
1.8.3 L'estimateur à noyau . . . . .	16
1.8.4 L'estimateur localement linéaire . . . . .	17
1.8.5 L'estimation des $k$ plus proche voisins ( $k$ NN) . . . . .	18
1.9 Description et contribution de la thèse . . . . .	19

<b>2</b>	<b>Estimation de M-Régression avec la méthode du k plus proche voisin sous des données quasi-associées dans les statistiques fonctionnelles</b>	<b>24</b>
2.1	Introduction	25
2.2	Le modèle M-estimateur fonctionnel	27
2.2.1	Propriétés de l'estimateur $k$ NN M-asymptotique	28
2.3	Principales hypothèses et notations	29
2.4	Résultats	31
2.4.1	La consistance	31
2.4.2	Normalité asymptotique	45
2.5	Étude de simulation	53
2.6	Conclusion	58
<b>3</b>	<b>Le taux de convergence de la fonction de hasard avec la variable explicative fonctionnelle : cas des données spatiales avec la méthode des k plus proches voisins</b>	<b>59</b>
3.1	Introduction	60
3.2	Le modèle et son estimateur avec la méthode $k$ NN	63
3.3	Propriétés asymptotiques	66
3.4	Résultat et preuve	67
3.5	Application	80
3.5.1	Estimation du point de risque	80
3.5.2	Mise en place du paramètre de lissage uniforme pour le choix des résultats du paramètre de lissage	82
3.5.3	Sélection du paramètre de lissage dans la fonction de distribution de hasard	82
3.5.4	Application de données simulées	85
3.5.5	Application de données réelles	88
3.6	Conclusion	90
	<b>Bibliographie</b>	<b>91</b>
	<b>Abstract</b>	<b>105</b>
	<b>Résumé</b>	<b>106</b>

## TABLE DES FIGURES

1.1	Les données fonctionnelles. . . . .	5
1.2	Courbes de marche par mesurant l'angle de flexion au niveau du genou . . . . .	6
1.3	Courbes de température de surface autour du courant océanique El Niño (57 courbes) par incréments de 12 mois depuis 1950. . . . .	7
1.4	Les courbes spectrométriques des 215 échantillons de viande. . . . .	8
1.5	Données obtenues en observant la dureté de la farine 115 pendant 480 secondes. . . . .	9
2.1	Les courbes $A_i(t), t \in [0, \pi/2]$ pour $i = 1, \dots, 150$ et $m = 1, 5, 10$ . . . . .	54
2.2	Prédiction de l'estimateur classique et robuste avec les méthodes CV et $k$ NN, respectivement. . . . .	56
2.3	Les boxplots des $MSE$ des valeurs de prédiction par les quatre méthodes pour les différents degrés de dépendance sans valeurs aberrantes . . . . .	57
3.1	Le MSE des deux modèles . . . . .	87
3.2	MSE en présence de valeurs aberrantes . . . . .	88
3.3	Températures mensuelles les plus élevées sur 125 sites météorologiques aux États-Unis . . . . .	89
3.4	Comparaison des résultats de prédiction entre les deux méthodes. . . . .	90

## LISTE DES TABLEAUX

2.1	Comparaison entre les quatre méthodes en présence de valeurs aberrantes avec différents degrés de dépendance . . . . .	58
-----	---	----

## 1.1 Contexte bibliographique sur statistique non paramétrique fonctionnelle

La statistique fonctionnelle est un domaine d'exploration scientifique en pleine expansion. Il se caractérise par une grande diversité, et couvrant différents domaines et aspects fondamentaux. Elle représente une importance considérable dans la statistique moderne, à la fois sur le plan conceptuel et dans ses applications pratiques. La statistique fonctionnelle se divise en deux catégories : la statistique paramétrique, qui suppose que la distribution suit un modèle décrit par un nombre fini de paramètres, et la statistique non paramétrique, qui ne fait aucune hypothèse sur la distribution. Ce champ d'étude suscite un intérêt croissant parmi de nombreux chercheurs et dans divers domaines. Comme en témoignent les nombreuses publications scientifiques dans ce domaine. Les travaux pionniers dans ce champ remontent aux les cinquantaines années, avec les études de Tucker [188] et Rao [169]. Ils ont examiné l'estimation de maximum de vraisemblance pour la distribution multinomiale. Depuis cette époque, plusieurs chercheurs ont apporté leur contribution au développement de cette approche.

L'un des problèmes les plus courants en statistique est la prédiction. Le modèle de régression est l'outil principal pour répondre à cette question en statistique non paramétrique. Ce modèle a été largement étudié et développé par de nombreux auteurs pour les variables explicatives, quelle que soit la dimension des variables explicatives, uni-, multi- ou infiniment dimensionnelles.

À l'aide d'un modèle de régression, vous pouvez déterminer comment deux ou plusieurs variables aléatoires sont liées. De nombreux auteurs et dans de nombreux contextes ont

exploré les caractéristiques et les résultats de l'estimateur de la fonction de régression. Par exemple, lorsque des variables sont indépendantes et identiquement distribuées (i.i.d), Nadaraya en 1964 a examiné comment estimer la fonction de régression. La même année, Watson a joué un rôle essentiel dans l'établissement la convergence systématique de cet estimateur, puis la convergence régulière presque certaine par Devroye [61], et la normalité asymptotique par Roussas [175], ou encore les résultats asymptotiques sur des processus  $\alpha$ -mélangeants pour de l'estimateur de la fonction de régression par Györfi. Dans 1991, les termes asymptotiques précis de l'erreur quadratique de l'estimateur du noyau de la fonction de régression ont été fournis par Vieu. Le concept de dépendance de quasi-association a été introduit par Douge [66] pour des variables aléatoires qui prennent ses valeurs dans un espace de Hilbert, et a établi quelques théorèmes limites pour ce type de variables.

Cette idée d'association peut être considérée comme un exemple particulier de la condition de dépendance faible, permettant le traitement de la corrélation mélange et de l'association à travers une seule méthode. Les résultats les plus récents portent sur la régression relative non paramétrique pour les variables aléatoires associées, une technique basée sur la minimisation de l'erreur relative quadratique moyenne, qui a été développée par Mechab et Laksaci [151].

Enfin, la normalité asymptotique de la fonction de régression quand la variable explicative prend ses valeurs dans un espace de Hilbert et sous des données quasi-associées a été étudiée par Douge [67].

## 1.2 Les données fonctionnelles

### 1.2.1 Contexte bibliographique sur les données fonctionnelles

La modélisation statistique des données fonctionnelles est liée à l'étude d'ensembles des données dans lesquels les observations peuvent être assimilées à des surfaces ou des courbes, comme la météorologie, la médecine, et dans de nombreux champs. A conséquence au l'avancement des technologies informatiques et de l'extension des capacités de stockage, d'énormes capacités ont été développées pour traiter ce type de données provient de divers champs. En addition, ce thème englobe toutes les préoccupations de la communauté statistique, des plus pratiques aux plus théoriques, sans privilégier l'une par rapport à l'autre.

Dans l'histoire de ce domaine, ont été obtenus les conclusions préliminaires concernant l'étude d'observations sous forme de trajectoires par Holmstrom [109], Obukhov [155] en climatologie, Deville [57] en démographie, puis en génétique, par Molenaar et Boomsma

[152], Kirkpatrick et Heckman [116] .

Il existe de nombreux exemples de types de données, Altman [4] et Presnell [158] ont utilisé des données biométriques. De plus, des données sur les ondes radar peuvent également être trouvées dans Poskitt et Presnell [157]. Les données les plus récentes utilisées, données spectrométriques utilisées en chimie par Abraham et al. [1]. Vous pouvez trouver plus d'exemples dans les travaux de Ramsay et Silverman [167] et Dabo et Ferraty [50], et autres.

Récemment, les progrès technologiques ont contribué à l'utilisation fréquente des données fonctionnelles, ce qui a incité et encouragé au l'évolution de la théorie des statistiques fonctionnelles. Actuellement, il existe un nombre croissant de situations dans divers domaines des sciences appliquées dans lesquels les données collectées fonctionnent. Il existe de nombreux exemples illustrant à la fois le grand nombre et la variété des données de nature fonctionnelle auxquelles un statisticien peut être confronté sont courbes spectrales, séries temporelles, images satellites et enregistrements audio. D'un point de vue théorique, de nombreux problèmes statistiques distincts peuvent impliquer un échantillon de données fonctionnelles, comme l'analyse en composantes principales (ACP), la classification, la régression, la prédiction, et les études longitudinales. Par conséquent, un nouveau domaine a été créé pour l'étude des données fonctionnelles en statistique. Où il suscitait beaucoup d'intérêt au début des années 1980, en particulier, des travaux de Dauxois *et al.* [52], et Grenander [97], et Ramsay [163]. Il a été publié par Ramsay et Silverman [164], puis, à travers les nombreux travaux de Ramsay et Silverman ([167], [165]), Bosque [29], et Ferraty et Vieu [80].

Ferraty et Vieu font une étude qui résume la plupart de leurs contributions à l'estimation non paramétrique dans le cas où les données sont de nature fonctionnelle. Ferraty et Vieu [86] ont présenté les premiers résultats dans le cas fonctionnel. Puis ces résultats ont été élargis et développés dans 2002 par Ferraty *et al.* [83]. En abordant la question de la prédiction sur les opérations en temps continu. Ces résultats ont été généralisés par plusieurs auteurs. La normalité asymptotique de l'estimateur de la fonction de régression sous la condition de dépendance a été examinée par Masri [147], la convergence quadratique moyenne du même estimateur a été montrée par Laksaci [124].

Prenez également note du fait que d'autres auteurs ont abordé des cas supplémentaires de variables lors de l'estimation de la fonction de régression. Comme l'étude de Laïb et Louani ([120], [121]), où ils ont étudié l'estimation de la fonction de régression fonctionnelle.

Dans leur contribution, Ramsay et Silverman [168], [167] proposent un travail très efficace sur la modélisation statistique des données fonctionnelles. D'autre part, Bosq [29] s'est intéressé aux modèles fonctionnels linéaires, comme le processus d'auto-régression,

puis a développé ses recherches vers la prédiction dans le domaine des modèles de grande dimension. Et les fondations en Cardot *et al.* [38] ont utilisé l'analyse en composantes principales de la variable fonctionnelle estimée par le modèle de régression linéaire Hilbertien. Concernant les fonctions conditionnelles en situation fonctionnelle, il existe de nombreuses publications parmi elles Ramsay et Selvarman ([167] et [165]), Ferraty et Vieu [79] qui ont présenté l'estimation non paramétrique fonctionnelle à noyau de la fonction de répartition conditionnelle, la fonction de densité conditionnelle, le mode et le quantile conditionnel avec des données indépendantes et identiquement distribuées et  $\alpha$  mélangées dont ils ont établi la convergence presque complète avec le taux de convergence. La normalité asymptotique de ces estimateurs est donnée par Ezzahraoui et Ould Saïd [72]. Lorsque la variable d'intérêt est fonctionnelle, Crambes *et al.* [46] ont récemment proposé la recherche de la prédiction dans le modèle linéaire fonctionnel.

Au cours des dernières décennies, la régression non paramétrique a considérablement progressé, Ferraty et Vieu [86] ont produit les premiers résultats et obtenu une convergence presque complète d'un estimateur à noyau de ce modèle non paramétrique dans le cas où les variables sont i.i.d. Ferraty *et al.* [87] ont élargi ces résultats pour inclure des problèmes de régression non standard tels que la prévision dans le cadre de séries chronologiques. En considérant l'estimateur de Ferraty et Vieu [77] et employant la propriété de concentration de la variable explicative sur les petites boules. Demongeot *et al.* [54] ont récemment publié une nouvelle approche pour l'estimation locale linéaire de la densité conditionnelle de données fonctionnelles dans un espace semi-métrique ; en utilisant cette méthode, ils ont pu obtenir des vitesses de convergence et uniforme presque complète et une convergence ponctuelle. Aussi, Ferraty *et al.* [85], Ferraty *et al.* [79], Ferraty et Vieu [80] et Ezzahrioui [72], ont également suggéré et étudié des techniques pour obtenir des quantiles conditionnels en utilisant l'estimation par noyau de la fonction de répartition conditionnelle. Pour les variables fonctionnelles spatialement dépendantes, Laksaci et Meref [125] ont démontré que l'estimateur de quantiles conditionnels à noyau converge presque complètement, avec aussi le taux. De plus, il y a la convergence en moyenne quadratique, que Ferraty *et al.* [82] ont développé. Ils ont expliqué en détail le terme asymptotique précis de l'erreur quadratique.

### 1.2.2 Variables et données fonctionnelles

Le développement technologique a contribué à l'évolution de la branche statistique, consacrée à l'analyse fonctionnelle, ils l'ont développé et élargi pour traiter les observations comme des éléments aléatoires fonctionnels. Dans ce contexte, pour fixer le vocabulaire utilisé nous présentons quelques définitions.

**Définition 1.2.1.** (Ferraty et Vieu [79]) Une variable aléatoire est dite fonctionnelle si

ses valeurs sont dans un espace de dimension infinie. Une observation d'une variable fonctionnelle est appelée donnée fonctionnelle.

$$X = \{X_{t,t \in \mathcal{T}}, X_t : \Omega \rightarrow \mathbb{R}\}.$$

$-\mathcal{T} \subseteq \mathbb{R}$  : courbe  $-\mathcal{T} \subseteq \mathbb{R}^2$  : image.

En pratique, une donnée fonctionnelle est observée dans un nombre fini d'instantants :

$$0 = t_0 < t_1 < \dots < t_k = T : \\ \{(t_0, X_{t_0}), (t_1, X_{t_1}), \dots, (t_k, X_{t_k})\} \text{ ( Figure (1) )}$$

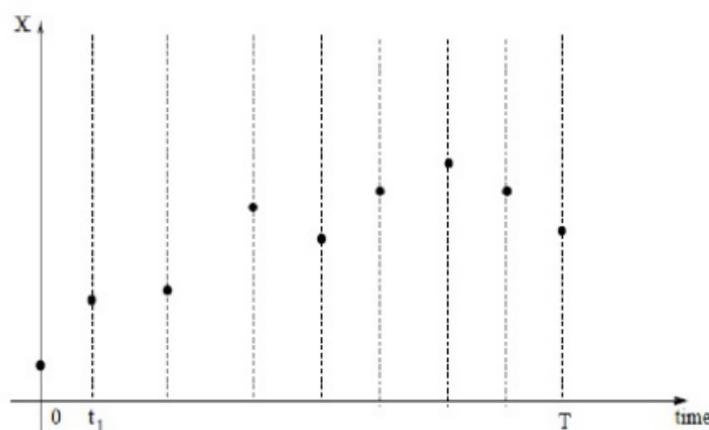


FIGURE 1.1 – Les données fonctionnelles.

**Définition 1.2.2.** (Ferraty et Vieu [79]) On appelle modèle fonctionnel, tout modèle prenant en compte au moins une variable aléatoire fonctionnelle.

Soit  $X$  une variable aléatoire à valeurs dans un espace de dimension infini  $\mathcal{A}$  et soit une application  $\phi$  définie dans  $\mathcal{A}$  et dépendant de la distribution de  $X$ . Un modèle d'estimation de  $\phi$  consiste à introduire une certaine contrainte de la forme  $\phi \in \mathcal{C}$ , d'où la définition suivante :

**Définition 1.2.3.** (Ferraty et Vieu [79]) Un modèle fonctionnel est dit paramétrique réel (ou vectoriel) si la classe de fonctions  $\mathcal{C}$  est indexable par un nombre fini de paramètres réels (ou vectoriels). Dans le cas contraire, il est dit non-paramétrique, i.e. que la classe  $\mathcal{C}$  est indexable par un nombre infini de paramètres d'un espace de dimension infinie (espace fonctionnel).

### 1.2.3 Domaines d'applications des données fonctionnelles

Notons que l'étude de ces données fonctionnelles est en plein développement tant sur le plan d'articles de recherche et la variété des phénomènes examinés, nous présentons dans cette partie quelques domaines variés dans lesquels apparaissent les données fonctionnelles (Économétrie, Chimométrie et Climatologie), et pour illustrer les types de problèmes auxquels les statistiques fonctionnelles pourraient répondre.

- En science de la biologie, on retrouve principalement le travail pionnier de Rao [169] liés à l'étude des courbes de croissance. Ces derniers temps, Ramsay et Silverman [167], ont étudié les variations de l'angle du genou au cours de la marche, pour chaque patient souffrant de la maladie de Parkinson, où les courbes de marche sont utilisées comme indicateur pour comparer l'efficacité des deux types de traitement (Figure 1.2). Concernant la biologie animale, des études de la ponte de mouches méditerranéennes ont été effectuées par plusieurs auteurs, Chiou *et al.* [40], Chiou *et al.* [41], Cardot [38], les données consistent en des courbes donnant pour chaque mouche la quantité d'oeufs pondus en fonction du temps.

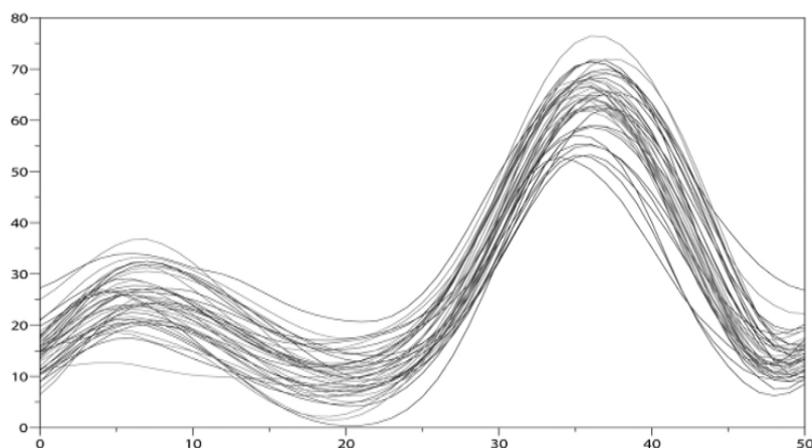


FIGURE 1.2 – Courbes de marche par mesurant l'angle de flexion au niveau du genou

- La chimométrie c'est une partie des champs qui utilisation de méthodes de la statistique fonctionnelle, il y a des ensemble de travaux consacrés à ce sujet, on peut mentionner Frank et Friedman [92], ainsi que Hastie et Mallows [107], et en trouvent les travaux de Ferraty et Vieu ([87], [77]) qui ont étudié et analysé la teneur en matières grasses de la viande compte tenu des longueurs d'ondes infrarouges qu'absorbent ces morceaux de viande (morceau de viande comme une variable d'intérêt, courbe d'absorption comme une variable explicative).

- Dans le domaine de l'environnement, plusieurs auteurs se sont préoccupés sur la ques-

tion de prévision de pollution en utilisant des méthodes de la statistique fonctionnelle. Comme Aneiros-Perez *et al.* [5], Febrero *et al.* [76], ces données sont des mesures des pics quotidiens de pollution à l’ozone (variable d’intérêt) basées sur les courbes des polluants et les courbes météorologiques de la veille (variables explicatives)..

-El Niño est phénomène climatique, les mécanismes qui régulent ce phénomène ils inclut les doubles interactions entre l’atmosphère et l’océan dans les zones équatoriales. Les mesures mensuelles de la température de surface de l’océan prises depuis le début des années 1950. Constituent l’ensemble de données dont nous disposons. Il est possible d’évaluer le caractère fonctionnel des données en prenant en compte l’évolution temporelle des températures. (Pour plus de références voir Bosq [28], Ramsay et Silverman [165], Ferraty *et al.* [85], Hall et Vial [101]).

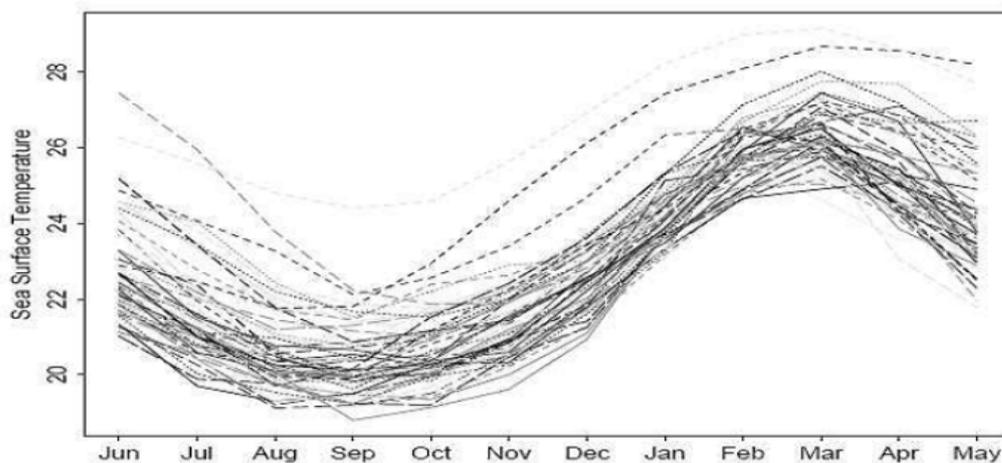


FIGURE 1.3 – Courbes de température de surface autour du courant océanique El Niño (57 courbes) par incréments de 12 mois depuis 1950.

-En linguistique, plusieurs auteurs ont étudié des applications liées à ce domaine, par exemple Berlinet *et al.* [22], Hastie *et al.* [108], et aussi Ferraty et Vieu [88]. Ils ont travaillé sur un type de données qui sont les enregistrements de phonèmes prononcés par une variété de personnes sous forme de courbes. Où ils ont associé l’étiquette à chaque son et l’ont représenté comme une variable d’intérêt, et l’objectif est de créer une classification de ces courbes en utilisant la courbe enregistrée qui a été représentée comme une variable explicative.

Les statistiques fonctionnelles ont été employées dans de plusieurs domaines différents, par exemple, en l’économie ont été étudiées par Benko *et al.* [21], et par Kneip et Utikal [118], en graphologie, par Hastie *et al.* [108] et Ramsay [166], l’étude des audiences télévisuelles par Cardot *et al.* [37], des profils de croissance par Sauder *et al.* [180], l’étude

de l'évolution des taux d'intérêts au cours du temps par Laurini [131].

### 1.2.4 Quelques exemples de données fonctionnelles

1. Cet exemple d'application de données fonctionnelles se concentre sur la problématique du contrôle qualité de la viande hachée dans le secteur alimentaire, (voir Thèse Doctorat de Laksaci [123]). Le but de cette recherche est de prédire la teneur en gras d'un morceau de viande ( $Z$ ) compte tenu de sa courbe spectrométrique ( $A$ ). La variable fonctionnelle est présentée par la courbe d'absorption lumineuse en fonction de la longueur d'onde (ces courbes appelées spectrométriques), qui est une courbe obtenue à l'aide d'une méthode traditionnelle de spectrométrie proche infra-rouge. Les courbes spectrométriques  $A_i$  correspondant à 215 morceaux de viande sont présentées ci-dessous (voir Figure 1.4).

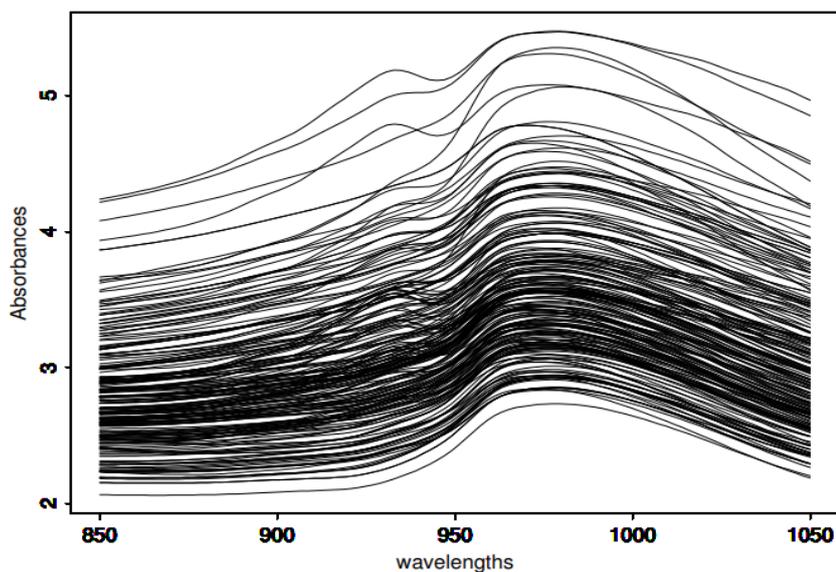


FIGURE 1.4 – Les courbes spectrométriques des 215 échantillons de viande.

2. Un autre domaine est celui de l'industriel, elle a été menée au centre de recherche Danone Vitapol, Paris, son intérêt pour la prédiction de la qualité des biscuits à travers la courbe issue du pétrissage qui liée à la qualité de la farine utilisée dans la fabrication des biscuits, où la courbe de pétrissage représente une mesure de la résistance de la pâte durant 480s dans l'étape de pétrissage, et l'objectif est de mesurer la capacité à prédire la qualité des biscuits grâce à des courbes de pétrissage et à prédire la qualité avant 480s. (voir Figure 1.5).

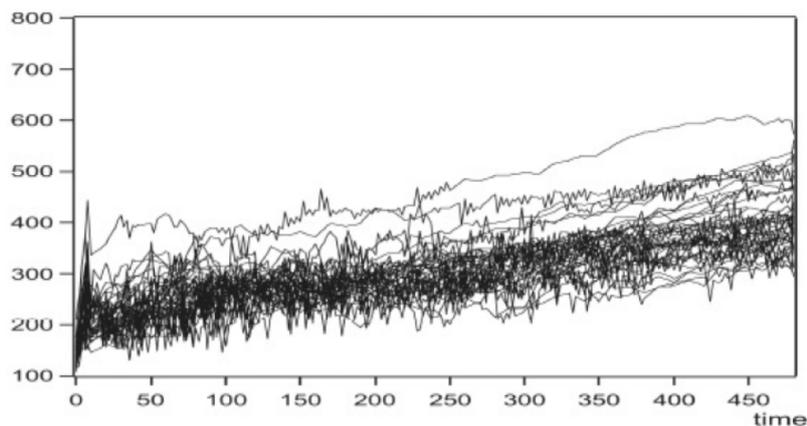


FIGURE 1.5 – Données obtenues en observant la dureté de la farine 115 pendant 480 secondes.

## 1.3 Données spatiale

### 1.3.1 Présentation des données spatiales

Le terme « analyse spatiale » a un pedigree en géographie qui remonte au moins aux années 1950. L'analyse spatiale est un terme largement utilisé dans les littératures sur les systèmes d'information géographique (SIG) et la science de l'information géographique (SIGc). Une définition de l'analyse spatiale est qu'elle représente un ensemble de techniques et de modèles qui utilisent explicitement le référencement spatial associé à chaque valeur de données ou objet spécifié dans le système à l'étude. Les méthodes d'analyse spatiale doivent formuler des hypothèses ou s'appuyer sur des données décrivant les relations spatiales ou les interactions spatiales entre les cas. Pour plus de précisions, dans le discipline de la météorologie, lors de l'étude des précipitations cumulées enregistrées sur un certain nombre de stations météorologiques. Le pourcentage de personnes atteintes du Covid-19 dans les wilayas d'Algérie en est un autre exemple en épidémiologie lorsqu'il s'agit d'analyser une variable d'intérêt. Ces données sont actuellement simples à obtenir, mais elles ne peuvent être examinées par les méthodes classiques de la statistique classique, dont l'un des principes fondamentaux est l'indépendance entre les observations. Il est nécessaire d'utiliser de nouvelles techniques d'analyse qui prennent en considération la dépendance spatiale des observations en conséquence. Selon le type de données, différentes approches doivent être utilisées, et afin de modéliser de telles actions, il est nécessaire d'identifier un type de corrélation entre certaines variables aléatoires dans un site et d'autres variables aléatoires dans des sites voisins. Il s'agit d'une qualité primordiale de l'analyse des données spatiales. De ce fait, les techniques de statistiques spatiales offriront la possibilité, entre autres, d'analyse exploratoire des données, l'étude des cor-

rélations spatiales, et leur modélisation jusqu'à ce que la prévision d'un événement dans des sites non observées.

Lorsque nous étudions des phénomènes en statistiques spatiales, nous sommes guidés pour regarder un ensemble spatial  $\mathcal{D}$  de sites. Ces observations sont les réalisations d'un processus appelé champ aléatoire (en mathématiques)  $\mathcal{X} = \{\mathcal{X}_i, i \in \mathcal{D}\}$ , indexé par  $\mathcal{D}$  (ensemble spatial), et  $\mathcal{X}_i$  sont des variables aléatoires appartenant à un espace d'états  $T$ . L'ensemble des indices est un range  $\mathcal{D}$  de  $\mathbb{R}^N$ ,  $N > 1$  ayant un rectangle de volume strictement positif, et l'indice  $\mathbf{i}$  'la localisation', se déplace donc continûment dans cet espace et est soit fixe et prévisible, soit aléatoire. En revanche, les observations de terrain sont réalisées en réalité à un nombre limité d'emplacements déterministes à l'intérieur du sous-ensemble bidimensionnel  $\mathcal{D}$  ( $\mathcal{D} \in \mathbb{R}^2$ ). On peut aussi se référer aux différentes dimensions de  $\mathcal{D}$ ,  $\mathcal{D}$ -unidimensionnel (chromatographie) ou ( $\mathcal{D} \in \mathbb{R}^3$ ) (prospection minière, imagerie 3D),  $\mathcal{D}$  de dimension  $d > 3$  (la planification des expériences numériques ou la statistique bayésienne).

### 1.3.2 Données fonctionnelles spatiales

Pour introduire la fonctionnelle spatiale, supposons que l'on note par  $\mathcal{Z}_i = (\mathcal{X}_i, \mathcal{Y}_i)$  le processus mesurable et la technique strictement stationnaire, avec  $\mathbf{i} \in \mathbb{Z}^N$  et  $N \geq 1$  choisis parmi  $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$  variables aléatoires, qui est valorisé dans  $(F \times \mathbb{R})$ , où l'on définit un espace semi-métrique par  $(F, d)$ . On note  $x$  une courbe fixe dans  $F$ , et on désigne au voisinage de  $x$  par  $\mathcal{N}_x$ , et à un sous-ensemble fixe compact de  $\mathbb{R}$  par  $\delta$ . Ensuite, nous supposons que les  $\mathcal{Z}_i$  sont (i.d.), c'est-à-dire identiquement distribués à  $\mathcal{Z} = (\mathcal{X}, \mathcal{Y})$  et que la probabilité conditionnelle de  $\mathcal{Y}$  étant donné  $\mathcal{X}$  a une forme régulière. De plus, supposons que  $\mathbf{i} = (i_1, \dots, i_N) \in \mathbb{Z}^N$  soit un point appelé site. Supposons que le processus spatial est accessible sur un rectangle régional

$$\mathcal{I}_{\mathbf{n}} = \{ \mathbf{i} = (i_1, \dots, i_{\omega}, \dots, i_N) \in \mathbb{Z}^N, 1 \leq i_{\omega} \leq n_{\omega}, \omega = 1, \dots, N \}, \text{ où } \mathbf{n} = (n_1, \dots, n_N) \in \mathbb{Z}^N.$$

## 1.4 Contexte bibliographique sur l'estimation par la méthode des k plus proches voisins

Pendant ces dernières années, un large éventail de littérature dans le domaine de l'estimation par la méthode des  $k$  plus proches voisins ( $k$ NN) est fourni par les revues bibliographiques à cause de nombreux avantages qu'elle présente. Le premier intérêt de cette méthode vient de la nature du paramètre de lissage. En effet, dans la méthode à noyau traditionnelle le paramètre de lissage est la fenêtre  $h$  qui est un nombre réel positif. Ce-

pendant dans notre méthode, on remplace ce dernier par une variable aléatoire réelle ( $H_n$ ) pour la variable explicative fonctionnelle ( $X$ ). D'autre plusieurs aspects de cette méthode ; dans le cadre fonctionnel, elle respecte la structure locale des données, ce qui est primordial en dimension infini. Elle est fréquemment employée en pratique, comme en Ferraty et Vieu [79] et s'avère facile à manipuler parce que l'utilisateur a un seul paramètre à contrôler, qui est représenté par le nombre  $k$  de plus proche voisin, ce paramètre  $k$  prend ses valeurs dans un ensemble fini. Cette méthode permet de créer un voisinage adapté aux données à chaque point.

La bibliographie de l'estimation par la méthode  $k$ NN existe depuis Royall [178] et Stone [181] qui ont commencé par estimer la fonction de régression. D'autre auteurs avaient aussi l'intérêt d'étudier cet estimateur, trouvons par exemple Collomb [44] qui a montré les différents types de convergence (probabilité et presque complète), Mack [144] a étudié la convergence en L2 et la distribution asymptotique, la convergence uniforme est donnée par Devroye [61]. Pour le cas des données fonctionnelles, Ferraty et Vieu [78] ont commencé par une introduction sur l'estimation des  $k$ NN.

Dans une publication récente, Burba *et al.* [32] utilisent la méthode des  $k$  plus proches voisins dans le contexte de l'estimation à noyau de l'opérateur de régression, ils ont réussi à obtenir la convergence presque complète de l'estimateur élaboré. La normalité asymptotique de l'estimateur non paramétrique robuste de la fonction de régression par la méthode des  $k$  plus proches voisins a été établie par Attouch et Benchik [11] dans le même contexte lorsque les variables sont fonctionnelles, et cela sous les propriétés de concentration sur les petites boules de la mesure de probabilité. La propriété de robustesse a été présentée comme une alternative aux méthodes d'estimation. La méthodologie robuste employée dans cette étude relève en fait de la catégorie des M-estimations présenté par Huber [112].

On retrouve également plusieurs études très récentes basées sur la méthode d'estimation des  $k$ -plus proches voisins, comme l'étude présentée par Almanjahi *et al.* [2], estimation de la densité conditionnelle dans la structure de régression scalaire sur fonction : approche linéaire locale  $k$ NN, où un nouvel estimateur est proposé en combinant la procédure des  $k$  plus proches voisins ( $k$ NN) avec l'approche linéaire locale. Ensuite, la consistance uniforme dans le nombre de voisins (UNN) de l'estimateur proposé est établie. Le travail qui soumis par Bouabsa [30], estimation d'erreur relative non paramétrique via un régresseur fonctionnel par le méthode des  $k$  plus proches voisins sous troncature aléatoire, Attouch et Bouabsa [10] où ils ont étudié l'estimation des  $k$  plus proches voisins du mode conditionnel pour les données fonctionnelles sous dépendance et Attouch *et al.* [8] ont étudié l'estimation des  $k$  plus proches voisins du mode conditionnel pour les données fonctionnelles. Le document de recherche pour Kara *et al.* [115] où il a présenté des estimateurs

du noyau le plus proche voisin ( $k$ NN) pour l'analyse non paramétrique d'échantillons statistiques impliquant des données fonctionnelles, et le point principal est d'examiner les méthodes basées sur les données pour sélectionner le nombre de voisins. Almanjahi *et al.* [3] ils ont traité le problème de l'analyse non paramétrique au moyen de la régression d'erreur relative lorsque l'interprétation de la variable est de dimension infinie. Sur la base de la procédure des  $k$  plus proches voisins ( $k$ NN), ils ont construit un estimateur et établi ses propriétés asymptotiques.

### 1.5 Robustesse en statistique non paramétrique

En statistique, le sujet de la robustesse de la procédure statistique usuelle (estimation, test) est crucial. Elle autorise de contrôler la stabilité de cette procédure par rapport à la déviation du modèle et/ou des observations. À la fin du XIXe siècle, ce problème a fait l'objet de longs débats, et de nombreux chercheurs avaient une idée relativement évidente de cette idée de robustesse. En effet, en 1818, les premiers travaux mathématiques sur le thème de l'estimation robuste sont présentés, ajouté à cela, des travaux de Laplace (P. S.) de la théorie analytique des probabilités dans son deuxième supplément. Plus précisément, G. E. P. Box a inventé le terme "robuste" en 1954. Cependant, jusqu'au milieu des années 1960, cette idée n'a pas été reconnue comme un sujet de recherche. Notamment avec les travaux de Huber [112], Hampel [103], une théorie cohérente des statistiques robustes est développée sur la base de critères de type min-max et emploie principalement des arguments de convexité.

En revanche, des méthodes d'ajustement automatisées robustes ont été développées par Huber ([110] et [111]), Andrews [15] et Krasker et Welsh [117], plus efficaces que la méthode des moindres carrés, mais en présence d'outliers ou lorsque la distribution des erreurs dans le modèle suit une distribution à queue lourde. Dans la forme de distribution lourde, ils sont plus efficaces. Le thème d'estimation robuste ayant eu un grand intérêt en statistique non paramétrique, surtout l'estimation robuste de la fonction de régression. Au début des années 1960, les premiers résultats dans ce domaine ont été générés par Huber [112], à partir duquel il a pu déterminer, pour cette fonction, la normalité asymptotique et la consistance d'une classe d'estimateurs.

Härdle [106], et Robinson [171] ont prouvé la normalité asymptotique d'une famille d'estimateurs pondérés issue de la méthode du noyau pour la fonction de régression sous des conditions de mélange. L'estimateur de Robinson [172] a été utilisé par Boente et Fraiman ([25], [26]) en même temps pour examiner simultanément les deux paramètres d'échelle et de position. La cohérence des estimateurs créés a été développée dans des contextes dépendants et indépendants, et dans des conditions généraux. Collomb et Härdle [43],

en tenant compte d'observations  $\phi$ -mélangeantes, ont déduit la convergence uniforme de l'estimateur robuste de la fonction de régression. Fan *et al.* [75] ont suggéré une méthode différente pour une estimation robuste de la fonction de régression. Avec l'aide de cette méthode, il est possible d'englober de nombreux modèles non paramétrique et de robustifier la régression classique. L'estimateur de Collomb et Härdle [43] ont été adaptés par Laïb et Ould-Saïd [122] pour le modèle d'autorégression d'un processus ergodique stationnaire. Même lorsque la fonction objective n'est pas bornée, ils ont réussi à obtenir la convergence uniforme de cet estimateur. La fonction de régression a été estimée par Cai et Ould-Saïd [34] à l'aide d'une version robuste de l'estimation polynomiale locale. Ils ont prouvé que la convergence presque sûre et la normalité asymptotique de ces estimateurs sous des conditions standards et quand les observations sont  $\alpha$ -mélangeantes.

## 1.6 Probabilités de petites boules et semi-métriques

La régression non paramétrique est bien connue pour son problème de dimension lorsque les variables sont multivariées (voir Stone [182]), résulte de cela, les méthodes non paramétriques dans les modèles de régression fonctionnelle peuvent une vitesse de convergence très lente. À cet égard, lorsque la variable explicative est multivariée (i.e. à valeurs dans un espace de dimension finie  $k$ ), les vitesses de convergence de l'estimateur à noyau dépendent d'un terme exprimé par  $h_n^k$ , qui représente la probabilité que la variable explicative appartienne à la boule de centre  $x$  et de rayon  $h_n$ . D'autre part, pour la variable explicative fonctionnelles, ce qui signifie, a une valeur dans un espace semi-métrique de dimension infinie  $(A, d)$  où  $d$  est une semi-métrique, nous utilisons des concepts de concentration plus généraux appelés petites boules de probabilité et traduisons les résultats asymptotiques en fonction de ces quantités. Les petites boules de probabilités sont définies comme suit :

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

La manière diminuent jusqu'à zéro a un impact significatif sur le taux de convergence de l'estimation en utilisant la méthode  $k$ NN. Il existe de nombreux articles sur la probabilité des équivalents asymptotiques pour les petites boules de probabilité lorsque  $d$  est une norme, pour plus de précisions, voir Dereich [55], Li et Shao [138], Gao et Li [94], Lifshits *et al.* [140] ou une semi-norme spécifique voir Aurzada et Simon [16], Lifshits et Simon [140]. Pour les processus non lisses tels que le mouvement brownien et les processus d'Ornstein-Ullenberg, ces petites boules de probabilité prennent une forme exponentielle concernant  $h_n$ , ce qui entraîne un taux de convergence qui est une puissance de  $\log(n)$  (Ferraty et Vieu [79]). La semi-métrie choisie a un impact direct sur la topologie

et la petite boule de probabilités. Différentes semi-métriques sont disponibles dans différentes situations pour trouver une topologie qui donne une idée pertinente de la proximité entre les courbes. Le choix d'une projection semi-métrique (basée sur les composantes principales fonctionnelles, la décomposition de Fourier ou les splines) peut être une alternative efficace au problème de la dimension. Ferraty et Vieu [80] démontrent lorsque  $E$  est un espace de Hilbert séparable, l'existence d'une projection semi-métrique spécifique qui conduit à des petites boules fractale de probabilités. Par ailleurs, dans certaines situations, par exemple lorsque l'ensemble de données est constitué de courbes lisses comme les courbes spectrométriques, l'utilisation de la semi-métrique basée sur la dérivée peut être avantageuse. Cependant, il peut être difficile de décider quelle semi-métrique utiliser en pratique. Pour résoudre ce problème, Ferraty *et al.* [87] ont proposé une méthode qui consiste à sélectionner une famille de semi-métrique de l'information arrive sur les données et à déterminer la semi-métrique la plus appropriée parmi cette famille en utilisant la Cross-Validation. Bien que la justification théorique de l'utilisation d'une semi-métrique spécifique reste un problème ouvert.

### 1.7 Estimation de la fonction de hasard

Pour l'étude de la fiabilité des statistiques, la fonction de hasard (appelée aussi le taux de hasard ou taux de survie ou taux de défaillance) est largement utilisée. Elle mesure la probabilité instantanée qu'à une date donnée un événement se produise, sachant que juste avant cette date, il ne s'est pas encore produit. Remarque que l'utilisation de ce modèle est devenue populaire en économétrie, notamment pour l'analyse des transitions (trajectoires individuelles sur le marché du travail). Donc, on peut essayer de mesurer, pour un chômeur, l'évolution dans le temps de sa propension à trouver un emploi. Vous pouvez regarder par exemple chez Lancaster [130], Florens *et al.* [91], et autres. Ils ont trouvé la fonction de risque sous deux formes, la première est la forme continue et la seconde est la forme discrète. Dans le cas de la forme continue, quand on n'a pas d'idée sur la forme initiale de la fonction de risque, l'estimation fonctionnelle non paramétrique s'impose (ou quand on fait des hypothèses d'appartenance à un ensemble de lois particulier). Dans le cas de la forme discrète, on estime les taux de hasard (décès, panne) comme étant des paramètres. Les débuts de l'analyse non paramétrique étaient Watson *et al.* [190] qui ont fait une proposition sur un estimateur, pour lequel ils construisaient la propriété de convergence asymptotiquement sans biais.

Les statisticiens sont généralement intéressés dans l'étude des caractéristiques numériques inconnues d'une population. L'analyse de survie comprend une large éventail de méthodes statistiques données de survie. Pour arriver à une (conjecture) étroite et signifi-

cative pour les paramètres inconnus, nous devons peut-être envisager certaines méthodes de calcul dont les propriétés théoriques peuvent être justifiées. Le processus d'acquisition d'un tel une (supposition) significative s'appelle une estimation.

Dans l'analyse de survie, il est souvent impossible de simplifier la tâche d'estimation à quelques paramètres de distribution, nous sommes intéressés à déterminer l'ensemble de la distribution du temps de survie, donc l'objectif n'est pas quelques paramètres, et nous devons présenter l'estimation fonctionnelle de l'ensemble courbe de distribution de survie à tous les points dans le temps. La méthode d'estimation pour ce type de problème est appelée estimation non paramétrique, car l'estimateur n'est pas un paramètre classique à dimensions finies. Une fois que l'estimation a été faite, un effort d'inférence supplémentaire est requis. La fonction de survie présentée par :  $P(Y > y)$ , qui mesure la probabilité d'observer un temps de survie plus long qu'une valeur fixe  $y$ .

Supposons  $Y$  une variables aléatoire à valeur dans  $\mathbb{R}^+$  associée à une durée de vie. La fonction de hasard est définie au point  $y$  comme étant la probabilité instantanée que cette durée de vie se termine à l'instant  $y$ . Plus précis, si  $h$  est la fonction de hasard, nous aurons :

$$h(y) = \lim_{\Delta y \rightarrow 0} \frac{\mathbb{P}(Y \leq y + \Delta y / Y \geq y)}{\Delta y}; \quad y > 0. \quad (1.1)$$

Pour que la variable  $Y$  possède une densité  $f$  par rapport à la mesure de Lebesgue, l'égalité (1.1) s'écrit :

$$h(y) = \frac{f(y)}{1 - F(y)},$$

avec :  $F$  est la fonction de répartition de  $Y$ .

De la même façon, la fonction de hasard conditionnée à une variable explicative  $X$  est définie comme suite :

$$h^X(y) = \lim_{\Delta y \rightarrow 0} \frac{\mathbb{P}(Y \leq y + \Delta y / Y \geq y, X)}{\Delta y}; \quad y > 0, \quad (1.2)$$

à partir de la fonction de densité conditionnelle  $f^X(y)$  et de la fonction de répartition conditionnelle  $F^X(y)$ , cette dernière s'écrit sous la forme :

$$h^X(y) = \frac{f^X(y)}{1 - F^X(y)}.$$

## 1.8 Les méthodes d'estimation

### 1.8.1 Le régressogramme

L'étude porte sur l'estimation de la fonction de régression. L'estimateur qui présente par Tukey [189] pour la régression a l'avantage d'être facile à mettre en œuvre et économique en opérations de calcul. La définition de cet estimateur est la suivante : considérons une partition  $(\pi_t)_{t=1,\dots,T}$  du support de  $X$ , fixée à priori. Pour la classe  $\pi_t$  où se trouve le point  $x$ , on effectue la moyenne des  $Y_i$  correspondants aux  $X_i$  de cette même classe  $\pi_t$ .

On a :

$$\forall x \in \pi_t : \hat{r}_{reg}(x) = \frac{\sum_{i=1}^n Y_i \mathbb{I}_{\{X_i \in \pi_t\}}}{\sum_{i=1}^n \mathbb{I}_{\{X_i \in \pi_t\}}}.$$

Le problème avec ce genre d'estimation lié au choix du positionnement des bornes des intervalles  $\pi_t$ .

### 1.8.2 L'estimateur de la fenêtre mobile

Pour éviter le problème classique d'estimation par le régressogramme, un nouvel estimateur dit de la "fenêtre mobile" est construit en remplaçant la partition en intervalles  $\pi_t$  par un seul intervalle qui varie de manière continue. Et donc on a :

$$\forall x \in \mathbb{R} : \hat{r}_{FM}(x) = \frac{\sum_{i=1}^n Y_i \mathbb{I}_{\{X_i \in [x-h; x+h]\}}}{\sum_{i=1}^n \mathbb{I}_{\{X_i \in [x-h; x+h]\}}}, \text{ avec } h \in \mathbb{R}_+^*,$$

où  $h$  un réel strictement positif appelé "paramètre de lissage".

Cet estimateur présente encore l'inconvénient d'être de nature discontinue.

### 1.8.3 L'estimateur à noyau

L'estimateur à noyau a été proposée en (1964) par Nadaraya et Watson (indépendamment et simultanément), également appelé aussi estimateur de Nadaraya-Watson(N-W), et il est construit de manière similaire à l'estimateur à noyau de la densité introduit par Parzen-Rosenblatt (voir Parzen [156] et Rosenblatt [173]), et il a depuis lors été largement étudié. Cet estimateur est défini comme suit :

$$\forall x \in \mathbb{R} : \hat{r}_{NW}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x - X_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)}, \text{ avec } h \in \mathbb{R}_+^*,$$

où  $K$  est une "fonction noyau", et  $h$  est "le paramètre de lissage".

### 1.8.4 L'estimateur localement linéaire

L'estimation de localement linéaire est fondée sur une simple généralisation de l'estimateur de Nadarya-Watson. Le principe de l'estimation localement linéaire consiste en l'ajustement local d'un polynôme de degré 1 :

$$r(\cdot) \approx \alpha + \beta(\cdot - x)$$

aux données  $\{(X_i, Y_i)_{i=1, \dots, n}\}$  par la méthode des moindres carrés, où  $\alpha$  et  $\beta$  sont des paramètres à estimé.

Nous essayons de résoudre le problème suivante de minimisation par rapport à  $(\alpha, \beta)$  :

$$\arg \min_{\alpha, \beta} \sum_{i=1}^n \{Y_i - \alpha - \beta(X_i - x)\}^2 K\left(\frac{X_i - x}{h_n}\right).$$

L'estimateur solution du problème de minimisation situé au-dessus est appelé estimateur localement linéaire (LL) et est défini par :

$$\hat{r}_{LL}(x) = \frac{\sum_{1 \leq i, j \leq n} Y_i w_{i,j}(x)}{\sum_{1 \leq i, j \leq n} w_{i,j}(x)}$$

où

$$w_{i,j}(x) = (X_i - x) ((X_i - x) - (X_j - x)) K\left(\frac{X_i - x}{h_n}\right) K\left(\frac{X_j - x}{h_n}\right).$$

Une autre forme de l'estimateur LL est possible et est donnée par :

$$\hat{r}_{LL}(x) := \frac{\hat{f}_0(x)\hat{g}_2(x) - \hat{f}_1(x)\hat{g}_1(x)}{\hat{g}_0(x)\hat{g}_2(x) - \hat{g}_1^2(x)},$$

où

$$\begin{aligned} \hat{f}_\gamma(x) &:= \frac{1}{nh_n} \sum_{i=1}^n Y_i (X_i - x)^\gamma K\left(\frac{X_i - x}{h_n}\right), \quad \gamma = 0, 1, \\ \hat{g}_\gamma(x) &= \frac{1}{nh_n} \sum_{1 \leq i \leq n} (X_i - x)^\gamma K\left(\frac{X_i - x}{h_n}\right), \quad \text{pour } \gamma = 0, 1, 2. \end{aligned}$$

L'estimateur localement linéaire est préférable à l'estimateur à noyau de Nadaraya-Watson dans la mesure où il réduit les effets de bord. Selon Fan [73], l'estimateur localement linéaire présente à la fois un biais plus faible que celui de l'estimateur de Nadaraya-Watson. Pour des informations plus détaillées sur les propriétés de l'estimateur localement linéaire et des applications sur des données réelles, nous recommandons au lecteur de consulter le Livre de Fan et Gijbels [74].

### 1.8.5 L'estimation des $k$ plus proche voisins ( $k$ NN)

Il y a eu des recherches approfondies sur l'estimation des  $k$  plus proches voisins, des nombreuses études ont été faites à ce sujet, lorsque la variable explicative appartient à un espace de dimension finie (voir les travaux de Györfi et al [100]). Lorsque la variable fonctionnelle est estimée avec une réponse réelles, deux méthodes distinctes existent pour l'estimation par  $k$ NN de la fonction de régression. Laloë [128] a présenté une méthode d'estimation  $k$ NN non paramétrique pour une variable fonctionnelle qui est un élément d'un espace de Hilbert séparable. Bien que Laloë [128] ait établi un résultat de convergence faible, son approche n'est pas entièrement fonctionnelle.

La technique de Laloë [128] consiste à réduire la dimensionnalité de la variable d'entrée en la projetant sur un sous-espace de dimension finie et en utilisant une technique multivariée sur les données projetées. Les travaux de Burba *et al.* [33], quant à eux, ont adopté une approche fonctionnelle pure et ont considéré le problème sur un espace fonctionnel semi-métrique. Ils ont obtenu une convergence presque complète et un taux de convergence pour les données indépendantes. Burba *et al.* [33] ont développé le Lemme de Collomb [44], qui sera utilisé dans les preuves.

En plus, Cérou et Guyader [39] ont abordé l'estimation du noyau par  $k$ NN pour la classification dimensionnelle infinie, tandis que Lian [139] a constaté que l'estimation de régression par  $k$ NN converge lorsque la réponse est un élément d'un espace de Hilbert. Lorsque la variable explicative de dimension finie, Tran [187] et Lu et Cheng [142] discutent tous deux de l'estimation par noyau des  $k$ NN pour des variables aléatoires  $\alpha$ -mélangeant.

Le concept de base derrière cette méthode est basé sur l'estimation d'un nombre fixe d'observations  $k$  qui est le plus proche du point souhaité, supposons  $X \in \mathbb{R}^q$  et nous avons un échantillon  $\{X_1; \dots; X_n\}$ . Pour n'importe quel point fixe  $X \in \mathbb{R}^q$ , en utilisant la distance euclidienne  $\|x\| = (x'x)^{\frac{1}{2}}$ , nous pouvons déterminer la distance entre chaque observation  $X_i$  et  $x$ .

$$S_i = \|x - X_i\| = (x - X_i)'(x - X_i)^{\frac{1}{2}}$$

Il s'agit d'une simple opération arithmétique dans l'ensemble de données. On a la statistique d'ordre pour les distances  $S_i$  sont  $0 \leq S_1 \leq S_2 \leq \dots \leq S_n$ .

Les observations qui correspondent à cette statistique d'ordre sont appelées les "plus proches voisins" de  $x$ . L'observation la plus proche de  $x$  est le premier voisin le plus proche, le prochain voisin le plus proche est la deuxième observation la plus proche, et ainsi de suite. Les observations triées par distances, où les plus proches voisins sont  $\{X(1); X(2); X(3); \dots; X(n)\}$ , le  $k$ -ième plus proche voisin de  $x$  est  $X(k)$ .

pour un  $k$  donné; spécifier :

$$T_x = \|X_{(k)} - x\| = S_{(k)}$$

$T_x$  est le  $k$ -ème ordre statistique sur les distances  $S_i$ .

La recherche scientifique se repose sur le développement de l'idée émise dans les premiers travaux de Royall [178] et Stone [181], où l'idée est de ne considérer que  $Y_i$  correspondant aux  $k$  points  $X_i$  les plus proches de  $x$  ( $k \in \mathbb{N}$ ), pour estimer la régression. Pour appliquer cela, nous apportons un changement dans la méthode du noyau où nous remplaçons un paramètre  $h$  par une variable aléatoire positive notée  $H_{n,k}$ , où il exprime la plus grande valeur des distances  $|x - X_i|$ .

L'avantage initial de l'utilisation de la méthode  $k$ NN est que son paramètre de lissage  $k$  prend des valeurs à partir d'un ensemble discret  $R = \{1, 2, \dots, n\}$ , alors que le paramètre de lissage  $h$  dans la méthode des noyaux est un nombre réel positif. La capacité de la méthode à construire un voisinage qui s'adapte à la structure locale des données est le deuxième et le plus crucial avantage. La méthode traditionnelle du noyau n'est pas très performante pour estimer la fonction de régression lorsque les données sont hétérogènes, c'est-à-dire de natures diverses. Alors que la méthode  $k$ NN évite ce problème en sélectionnant le nombre de voisins. En raison du paramètre adaptatif  $k$ , la méthode  $k$ NN est nettement plus efficace lorsque les données sont fonctionnelles. La méthode  $k$ NN pose cependant un défi technologique important. Spécifiquement que le choix des voisins les plus proches produit une variable aléatoire. De plus, comme nous ne pouvons pas utiliser les mêmes outils que dans la méthode du noyau standard, cette randomisation rend les démonstrations techniquement difficiles.

## 1.9 Description et contribution de la thèse

Le large éventail d'applications offertes par l'estimation fonctionnelle, ainsi que son outil théorique et sa capacité à décrire ces données en les considérant comme des réalisations d'une variable aléatoire dans un espace abstrait, expliquent l'enthousiasme de nombreux auteurs dans des domaines variés. Ce domaine est d'actualité et d'une grande importance scientifique sur la base de les différentes conclusions ou découvertes issues de la recherche et qui ont été validées et diffusées dans des revues scientifiques reconnues.

Nous présentons dans cette section un bref résumé des principales contributions de cette thèse. Le cadre de recherche exposé dans cette thèse est de l'estimation de quelques fonctions robustes dans le cas où les observations sont de dimension éventuellement infinie en utilisant la méthode des  $k$  plus proches voisins, dans l'objectif par la suite est d'étudier les propriétés asymptotiques. Cette étude est principalement fourni en trois chapitres, dont le premier est présenté au contexte bibliographique et représentative, dans lequel nous avons

contribué au domaine de l'estimation fonctionnelle non paramétrique. Nous avons discuté de les données fonctionnelles et ses champs d'application, robustesse en statistique non paramétrique, fonction de hasard, et nous avons donné une idée de l'estimation par la méthode  $k$ NN.

Le chapitre 2, l'objectif principal de l'étude est de construire un estimateur de la fonction de régression en reliant l'approche de M-estimation avec le cadre de quasi-association à l'aide de la méthode  $k$ NN, et nous étudions les propriétés asymptotiques de cet estimateur. Ce chapitre est présenté comme suit :

**Chapitre 2 : Estimation de M-Régression avec la méthode du  $k$  plus proche voisin sous des données quasi-associées dans les statistiques fonctionnelles** (Belatreche *et al.* [19])

Nous considérons  $(W_i)_{i=1,\dots,n} := (A_i, B_i)_{i=1,\dots,n}$  une série de données dans l'espace de Hilbert séparable des variables aléatoires stationnaires quasi-associées et identiquement distribuées  $\varepsilon := \mathcal{G} \times \mathbb{R}$ , où  $\mathcal{G}$  est un espace de Hilbert réel séparable avec les  $\|\cdot\|$  norme créée par le produit interne  $\langle \cdot, \cdot \rangle$ .

Le modèle non paramétrique, noté  $\vartheta_a$ , est implicitement défini comme un zéro par rapport à  $\phi$  dans l'équation

$$\Lambda(a, \phi) := \mathbf{E}[\rho(B_1, \phi) \mid A_1 = a] = 0, \quad (1.3)$$

L'estimateur naturel de  $\vartheta_a$  indiqué par  $\widehat{\vartheta}_a$  est un zéro par rapport à  $\phi$  de l'équation

$$\widehat{\rho}(a, \phi) = 0, \quad (1.4)$$

alors, on a que

$$\widehat{\Lambda}(a, \phi) := \frac{\sum_{i=1}^n L(h_L^{-1}(a - A_i)) \rho(B_i, \phi)}{\sum_{i=1}^n L(h_L^{-1}(a - A_i))},$$

avec  $L$  est une fonction de noyau et  $h_{L,n} = h_n$  est une suite de nombres réels positifs qui tend vers zéro lorsque  $n$  tend vers l'infini.

Le M-estimateur de  $k$ NN :

$$\widetilde{\Lambda}(a, \phi) := \frac{\sum_{i=1}^n L(E_k^{-1}(a - A_i)) \rho(B_i, \phi)}{\sum_{i=1}^n L(E_k^{-1}(a - A_i))}.$$

avec,

$$E_k = \min \left\{ h_L \in \mathbb{R}^+ \text{ tel que } \sum_{i=1}^n \mathbb{I}_{B(a, h_L)}(A_i) = k \right\}.$$

Les principaux résultats obtenus dans ce chapitre sont :

1. la consistance presque complète de  $\hat{\vartheta}_a$ , pour un  $a \in S$ .

**Théorème 1.1.** (Belatreche et al. [19]) *Sous des hypothèses de régularité de la fonction robuste ainsi que des hypothèses de concentration de la mesure de probabilité de la variable fonctionnelle, l'estimateur  $\hat{\vartheta}_a$  existe et est unique. De plus, on a, quand  $n \rightarrow \infty$  avec  $\varpi = \min(\varpi_1, \varpi_2)$ , que*

$$\sup_{a \in S} \sup_{k_n^1 \leq k \leq k_n^2} |\hat{\vartheta}_a - \vartheta_a| = O\left(\varphi_a^{-1} \left(\frac{k_n^2}{n}\right)^\varpi\right) + O\left(\left(\frac{\log n}{n^{-\gamma} (k_n^1)^d}\right)^{1/2}\right), \quad p.co.$$

2. La normalité asymptotique de  $\hat{\vartheta}_a$  pour un fixe  $a \in S$ .

**Théorème 1.2.** (Belatreche et al. [19]) *Sous certaines conditions et hypothèses, on a que*

$$\left(\frac{n\varphi_a(x_n)^d}{\sigma^2(a, \vartheta_a)}\right)^{1/2} (\hat{\vartheta}_a - \vartheta_a) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty,$$

où,

$$\sigma^2(a, \vartheta_a) = \frac{\mathbb{E}[\Lambda_a^2(B\vartheta_a) \mid A = a]}{\left(\frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a)\right)^2} \int_{\mathbb{R}'} L^2(z) dz,$$

$$\mathcal{A} = \left\{ a \in S, \mathbb{E}[\Lambda_a^2(B, \vartheta_a) \mid A = a] \frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a) \neq 0 \right\},$$

et  $\xrightarrow{\mathcal{D}}$  désigne la convergence en distribution.

Ci-dessous, au chapitre 3, dans les cas où les co-variables sont de nature fonctionnelle, nous avons développé un nouvel estimateur de hasard. Cet estimateur combine des données fonctionnelles spatiales avec la méthode des  $k$  plus proches voisins. Sous certaines hypothèses générales, nous exposons la convergence presque complète, aussi le taux de convergence.

**Chapitre 3 : Le taux de convergence de la fonction de hasard avec la variable explicative fonctionnelle : cas des données spatiales avec la méthode des  $k$  plus proches voisins.** (Belatreche et al. [20])

Soit  $M$  un entier naturel dans  $\mathbb{N}^*$ . On considère le champ aléatoire  $W_i = (A_i, B_i)$ ,  $i \in \mathbb{N}^M$  avec des valeurs dans  $\mathcal{G} \times \mathbb{R}$ , où  $(\mathcal{G}, d)$  est un espace semi-métrique de dimension éventuellement infinie. Dans ce contexte,  $(A_i)_{i \in \mathbb{N}^M}$  peut être une variable aléatoire fonctionnelle.

On fixe un point  $a$  dans  $\mathcal{G}$  (respectivement, un compact  $Q \in \mathbb{R}$ ), on suppose que les observations spatiales  $(A_i, B_i)_{i \in \mathbb{N}^M}$  ont la même distribution que  $W = (A, B)$  et que la version régulière de la probabilité conditionnelle de  $B$  sachant  $A = a$  existe et admet une densité

bornée par rapport à la mesure de Lebesgue sur  $\mathbb{R}$ , notée  $f_k^a$ . Avec la méthode  $k$ NN le paramètre fonctionnel étudié dans ce chapitre, noté  $h_k^a$ , est défini, pour tout  $b \in \mathbb{R}$  tel que  $F_k^a(b) < 1$ , par

$$h_k^a(b) = \frac{f_k^a(b)}{1 - F_k^a(b)},$$

où,  $F_k^a$  est la fonction de distribution conditionnelle de  $B$  sachant  $A = a$ , avec

$$F_k^a(\cdot) = \mathbb{P}(B \leq \cdot \mid A = a).$$

Nous estimons avec la méthode  $k$ NN la fonction de distribution conditionnelle par

$$\widehat{F}_k^a(b) = \frac{\sum_{i \in I_n} L(P_k^{-1}d(a, A_i)) J(h_J^{-1}(b - B_i))}{\sum_{i \in I_n} L(P_k^{-1}d(a, A_i))}, \quad \forall b \in \mathbb{R},$$

avec,  $L$  est un noyau et  $J$  est une fonction de distribution conditionnelle (cdf), définie par

$$J(\cdot) = \frac{\int_0^\cdot L(s)ds}{\int_0^{1/2} L(s)ds},$$

De  $\widehat{F}_k^a$ , on déduit un estimateur de la densité conditionnelle, noté  $\widehat{f}_k^a$ , où

$$f_k^a(\cdot) = (F_k^a)'(\cdot).$$

Donc

$$\widehat{f}_k^a(b) = \frac{h_J^{-1} \sum_{i \in I_n} L(P_k^{-1}d(a, A_i)) J'(h_J^{-1}(b - B_i))}{\sum_{i \in I_n} L(P_k^{-1}d(a, A_i))}, \quad \forall b \in \mathbb{R},$$

où,  $J'$  est la dérivée de  $J$ . L'estimateur  $k$ NN de la fonction de risque conditionnelle est noté par

$$\widehat{h}_k^a(b) = \frac{\widehat{f}_k^a(b)}{1 - \widehat{F}_k^a(b)}, \quad \forall b \in \mathbb{R}.$$

L'objectif de ce travail est d'étudier la convergence presque complète de l'estimateur  $\widehat{h}_k^a$  vers  $h_k^a$ , lorsque le champ aléatoire fonctionnel  $(W_i)_{i \in \mathbb{N}^M}$  satisfait la condition de mélange

suivante.

$$\left\{ \begin{array}{l} \text{Il existe une fonction } \Phi(s) \downarrow 0 \text{ quand } s \rightarrow \infty, \text{ telle que} \\ \forall Z, Z' \text{ sous-ensemble de } \mathbb{N}^M \text{ cardinal fini} \\ \alpha(\mathcal{B}(Z), \mathcal{B}(Z')) = \sup_{Y \in \mathcal{B}(Z), E \in \mathcal{B}(Z')} |\mathbb{P}(Y \cap E) - \mathbb{P}(Y)\mathbb{P}(E)| \\ \leq \Psi(\text{Card}(Z), \text{Card}(Z')) \Phi(\text{dist}(Z, Z')), \end{array} \right.$$

**Théorème 1.3.** (Belatreche et al. [20]) *Sous certaine conditions et hypothèses, on a que*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h, J \leq s_n} |\hat{h}_k^a(b) - h_k^a(b)| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}} \right).$$

**Théorème 1.4.** (Belatreche et al. [20]) *Sous certaine hypothèses, on a que*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h, J \leq s_n} |\hat{F}_k^a(b) - F_k^a(b)| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log n}{n \varphi_a(x_n)}} \right).$$

**Théorème 1.5.** (Belatreche et al. [20]) *Sous les hypothèses (K1), (K7) et, (K8-ii), (K9-ii) on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h, J \leq s_n} |\hat{f}_k^a(b) - f_k^a(b)| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}} \right).$$

## CHAPITRE 2

# ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

## Estimation de M-Régression avec la méthode du k plus proche voisin sous des données quasi-associées dans les statistiques fonctionnelles

L'objectif principal de ce chapitre est d'étudier la M-estimation non paramétrique sous séquence quasi-associée avec la méthode k Nearest Neighbor's short ( $k$ NN). Nous construisons un estimateur de cette fonction non paramétrique et nous étudions ses propriétés asymptotiques. En outre, une étude comparative basée sur des données simulées est également fournie pour illustrer la grande sensibilité de l'approche  $k$ NN à la présence même d'une petite proportion de valeurs aberrantes dans les données.

### 2.1 Introduction

Il est très bien reconnu que la régression robuste dans les statistiques est une méthode de recherche attrayante. Il est utilisé pour surmonter certaines des faiblesses de la régression classique, notamment lorsque les valeurs aberrantes contiennent des données hétéroscédastiques.

Dans la littérature statistique, plusieurs articles ont été consacrés à l'étude des propriétés des M-estimateurs non paramétriques. Les premiers résultats concernant ce sujet pour la normalité asymptotique dans les données dépendantes et indépendantes sont Györfi [99], Collomb [43], Huber [112], Hardle, [105], Robinson [171], Boente *et al.* ([26], [25]), pour les résultats antérieurs, Boente *et al.* [27], Laib *et al.* [122], Attouch *et al.* [12], Attouch *et al.* [13]. L'estimation par régression robuste non paramétrique a d'abord été introduite par Azzedine *et al.* [17]. Ils ont obtenu la convergence presque complète avec les taux dans le cas indépendant et identiquement distribué (i.i.d.), Crambes [46] a examiné le problème similaire pour une covariable fonctionnelle, puis Cai [34] a étudié ses propriétés asymptotiques sous les  $\alpha$ -hypothèse de mélange dans la norme  $L^p$ . Dans le cas des données ergodiques fonctionnelles et stationnaires, Gheriballah [96] a trouvé une convergence presque complète avec le taux, et pour les recherches récentes, nous pouvons nous référer à Derrar [56] et aux références qu'il contient.

Pour les variables aléatoires quasi-associées, la littérature sur M-estimation pour l'analyse des données fonctionnelles non paramétriques n'est pas encore bien documentée, notez que, Bulinski et Suquet [31] a introduit ce type de structure de dépendance comme une généralisation des variables associées positivement donnée par Esary *et al.* [71] et les variables aléatoires liées négativement considérées par Jong *et al.* [113] pour les champs aléatoires à valeurs réelles. Les deux types d'association ont une grande importance dans

divers domaines appliqués (voir le Livre de Barlow [18] pour une discussion plus approfondie sur ce sujet). L'estimation non paramétrique impliquant (positivement et négativement) des variables aléatoires associées a été largement étudiée. Nous citons, par exemple, Matula [149], Roussas [177], Masry [146], Roussas [176], Mebsout [150] et leur référence. Nous renvoyons le lecteur à Dedecker [53] ou Doukhan [65] pour d'autres structures de dépendance faibles et leurs applications.

L'objectif principal de l'étude est de créer un estimateur de la fonction de régression en rétablissant le lien entre l'approche de M-estimation et le cadre de la quasi-association à l'aide de la méthode du  $k$ NN. Cela est motivé par le fait que l'estimateur par régression robuste présente plusieurs avantages par rapport à l'estimateur par régression à noyau classique. Le principal avantage de l'utilisation d'une régression robuste est qu'elle permet de réduire l'effet des données aberrantes.

Cette célèbre méthode d'estimation  $k$ NN a suscité beaucoup d'intérêt dans la littérature statistique pour l'évaluation de données multivariées en raison de sa flexibilité et de son efficacité. Poussée par ses caractéristiques attrayantes, l'approche de lissage fonctionnel  $k$ NN a suscité un intérêt croissant au cours des dernières années. Le Livre de Györfi [100] est une analyse approfondie des estimateurs  $k$ NN dans le contexte de dimension finie. Les travaux dans ce domaine ont été lancés par Cover [45], et un grand nombre d'articles sont maintenant disponibles dans divers contextes d'estimation, parmi lesquels la régression, la discrimination, l'estimation de la densité et du mode et l'analyse de clustering, nous faisons référence à Collomb [42], Devroye et Wager ([59], [60]), Moore *et al.* [153], Devroye *et al.* [58], Beirlant [23], Li [134], Laloë [129], Burba *et al.* [33], Tran [185], Lian [133], Attouch *et al.* ([10], [8]), Kudraszow [119] et Kara *et al.* [115], Almanjahie *et al.* [3], Bouabssa [30] pour les dernières avancées et références. Notez qu'une telle étude a un grand impact sur la pratique. D'une part, la régression robuste est un modèle de régression alternatif essentiel qui permet de surmonter de nombreux inconvénients de la régression classique, tels que la sensibilité aux valeurs aberrantes ou les phénomènes d'hétéroscédasticité. De plus, il est bien connu que la méthode de  $k$ NN est meilleure que la méthode de noyau classique. Cependant, la difficulté de la méthode de  $k$ NN réside dans le fait que le paramètre de lissage est une variable aléatoire, contrairement à la régression classique dans laquelle le paramètre de lissage est un scalaire déterministe. Ainsi, l'étude des propriétés asymptotiques de notre estimateur proposé est compliquée et nécessite des outils et des techniques supplémentaires.

In NFDA, la méthode  $k$ NN M-estimation avec des données quasi-associées est nouvelle. L'objectif principal de cette recherche est de fournir des généralisations, au cas du  $k$  plus

proche voisin, des résultats obtenus par Attaoui *et al.* [6] dans le cas de la dépendance quasi-associée. Plus précisément, on établit uniformément la convergence presque complète sur le nombre de voisins avec les taux d'un estimateur construit en combinant les notions de robustesse avec celles de régression lissée. Nous soulignons que la caractéristique principale de notre approche est de développer un modèle de prédiction alternatif à la régression classique qui n'est pas sensible aux données aberrantes ou hétéroscédastiques, en tenant compte de la structure locale des données.

L'article est organisé de la manière suivante. La Section 2.2 est consacrée à la présentation de notre estimation puis les notations fixes et les hypothèses sont données dans la Section 2.3. Nous énonçons le résultat et leurs preuves dans la Section 2.4, où la convergence uniforme presque complète avec les taux est donnée dans la Sous-Section 2.4.1 et la Sous-Section 2.4.2 est consacrée à l'étude des normalité asymptotique. La Section 2.5 est consacrée à l'étude par simulation pour prouver l'efficacité de notre étude.

## 2.2 Le modèle M-estimateur fonctionnel

Soit  $(W_i)_{i=1,\dots,n} := (A_i, B_i)_{i=1,\dots,n}$  une série de données dans l'espace de Hilbert séparable des variables aléatoires stationnaires quasi-associées et identiquement distribuées  $\varepsilon := \mathcal{G} \times \mathbb{R}$ , où  $\mathcal{G}$  est un espace de Hilbert réel séparable avec les  $\|\cdot\|$  norme créée par le produit interne  $\langle \cdot, \cdot \rangle$ . Ce modèle non paramétrique, noté  $\vartheta_a$ , est implicitement défini comme un zéro par rapport à  $\phi$  dans l'équation

$$\Lambda(a, \phi) := \mathbf{E}[\rho(B_1, \phi) \mid A_1 = a] = 0, \quad (2.1)$$

dans laquelle  $\rho$  est une fonction borélienne à valeurs réelles qui satisfait certaines exigences de régularité décrites ci-dessus. Nous supposons que  $\vartheta_a$  existe et est unique pour tout  $a \in S$  où  $S$  est un sous-ensemble compact fixe de  $\mathbb{R}^d$  (voir, par exemple, Boente et Fraiman [25]). L'estimateur naturel de  $\vartheta_a$  indiqué par  $\widehat{\vartheta}_a$  est un zéro par rapport à  $\phi$  de l'équation

$$\widehat{\rho}(a, \phi) = 0, \quad (2.2)$$

alors, on a que

$$\widehat{\Lambda}(a, \phi) := \frac{\sum_{i=1}^n L(h_L^{-1}(a - A_i)) \rho(B_i, \phi)}{\sum_{i=1}^n L(h_L^{-1}(a - A_i))},$$

avec  $L$  est une fonction de noyau et  $h_{L,n} = h_n$  est une suite de nombres réels positifs qui tend vers zéro lorsque  $n$  tend vers l'infini.

### 2.2.1 Propriétés de l'estimateur $k$ NN M-asymptotique

En fait, cette étude s'est concentrée sur les propriétés asymptotiques de  $k$ NN M-estimateur, pour lequel le paramètre de lissage scalaire  $h_L$  est remplacé par un paramètre aléatoire défini par

$$E_k = \min \left\{ h_L \in \mathbb{R}^+ \text{ tel que } \sum_{i=1}^n \mathbb{I}_{B(a, h_L)}(A_i) = k \right\}.$$

Principalement, le M-estimateur de  $k$ NN est représenté par

$$\tilde{\Lambda}(a, \phi) := \frac{\sum_{i=1}^n L(E_k^{-1}(a - A_i)) \rho(B_i, \phi)}{\sum_{i=1}^n L(E_k^{-1}(a - A_i))}.$$

Effectivement, l'utilisation de  $E_k$  plutôt que  $h_L$  augmente la précision du cas précédent car le paramètre de lissage  $E_k$  dépend fortement des données contrairement au premier cas où  $h_L$  est choisi arbitrairement indépendamment de l'échantillon. De plus, la méthode  $k$ NN facilite la sélection du paramètre de lissage car ce choix se réduirait à un problème de choix d'un entier  $k$  entre 1 et la taille d'échantillon  $n$ . Cependant, établir les propriétés asymptotiques de l'estimateur proposé devient plus compliqué que dans le cas classique car le paramètre de lissage  $E_k$  est une variable aléatoire. Ainsi, le traitement de cette situation nécessite des outils mathématiques supplémentaires et des arguments analytiques spécifiques.

En réalité, l'estimateur  $\tilde{\Lambda}(a, \phi)$  est plus adapté que  $\hat{\Lambda}(a, \phi)$ , puisque son paramètre de lissage est choisi de manière sophistiquée, alors que  $\hat{\Lambda}(a, \phi)$  lorsque le paramètre de lissage est sélectionné arbitrairement indépendamment des résultats. Le réglage des propriétés asymptotiques de  $\tilde{\Lambda}(a, \phi)$ , est en effet plus complexe que l'estimateur  $\hat{\Lambda}(a, \phi)$  puisque son paramètre de lissage est aléatoire variable.

## 2.3 Principales hypothèses et notations

Jusqu'à préciser notre résultat clé, dans le nombre de voisins  $k \in (k_n^1, k_n^2)$ , la consistance presque complète de  $\hat{\Lambda}(a, \phi)$  est défini uniformément et on suppose que les deux premiers moments inverses conditionnels de  $B$  étant donnés  $A$ , sont finis et strictement positifs.

Dans tout chapitre on suppose que la suite  $(A_i, B_i)_{i=1, \dots, n}$  est un processus stationnaire quasi associé,  $S$  est un compact fixe sous-ensemble de  $\mathbb{R}^d$  et  $\Upsilon$  (respectivement  $\Upsilon_{i,j}$ ) la densité de  $A$  (respectivement la densité conjointe de  $(A_i, A_j)$ ). De plus, on fixe par  $C$  ou  $C'$  des constantes génériques positives et par

$$\Delta_\ell := \sup_{s > \ell} \sum_{|i-j| \geq s} \Delta_{i,j},$$

où

$$\begin{aligned} \Delta_{i,j} = & \sum_{\ell=1}^d \sum_{l=1}^d |\text{Cov}(A_i^\ell, A_j^l)| + \sum_{\ell=1}^d |\text{Cov}(A_i^\ell, B_j)| \\ & + \sum_{l=1}^d |\text{Cov}(B_i, A_j^l)| + |\text{Cov}(B_i, B_j)|. \end{aligned}$$

Dans la suite, nous désignerons par  $L_1 = L(d(a, A_1)/h_L)$ .

Maintenant, nous allons énoncer les hypothèses suivantes qui sont nécessaires pour montrer notre résultat principal.

(H1) Pour tous  $p \in (0, 1)$  et pour tous  $t > 0$ ,  $\mathbb{P}(A \in \mathcal{B}(a, t)) = \varphi_a(t) > 0$ , nous avons que

$$\lim_{t \rightarrow 0} \frac{\varphi_a(pt)}{\varphi_a(t)} = \chi_a(p) < \infty.$$

(H2) Le noyau  $L$  est supporté dans  $(0, 1/2)$  et a une première dérivée continue sur  $(0, 1/2)$  qui est telle que

$$0 < C \mathbb{I}_{(0,1/2)}(\cdot) \leq L(\cdot) \leq C' \mathbb{I}_{(0,1/2)}(\cdot) \text{ et } L(1/2) - \int_0^{1/2} L'(p) \chi_a(p) dp > 0,$$

où  $\mathbb{I}_A$  est la fonction indicatrice de l'ensemble  $A$ .

(H3) La classe de fonctions  $\kappa = \{\mapsto L(\psi^{-1}d(a, \cdot)), \psi > 0\}$  est une classe qui peut être évaluée ponctuellement de telle sorte que

$$\sup_G \int_0^1 \sqrt{1 + \log \mathcal{N}(\varepsilon \|\Theta\|_{G,2}, L, d_G)} d\varepsilon < \infty.$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

où le supremum est pris sur toutes les mesures de probabilité  $G$  sur l'espace  $\mathcal{G}$  avec  $G(\Theta^2) < \infty$  et où  $\Theta$  est l'enveloppe fonction de l'ensemble  $\kappa$ . Ici,  $d_G$  est la  $L_2(G)$ -métrique et  $\mathcal{N}(\varepsilon, \kappa, d_G)$  est le nombre minimal de boules ouvertes (par rapport à  $L_2(G)$ -métrique) de rayon  $\varepsilon$  qui sont nécessaires pour couvrir la classe de fonctions  $\kappa$ . On notera  $\|\cdot\|_{G,2}$  la norme  $L_2(G)$ .

(H4)  $k_n^1$  et  $k_n^2$  séquences vérifiées

$$\varphi_a^{-1}\left(\frac{k_n^2}{n}\right) \rightarrow 0, \quad \text{et} \quad \frac{\log n}{\min\left(n\varphi_a^{-1}\left(\frac{k_n^1}{n}\right), k_n^1\right)} \rightarrow 0.$$

(H5) La densité  $\Upsilon$  est de classe  $\mathcal{C}^1(\mathbb{R}^d)$ , tel que  $\inf_{a \in S} \Upsilon(a) > C > 0$  et la densité des joints  $\Upsilon_{i,j}$  satisfait  $\sup_{|i-j| \geq 1} \|\Upsilon_{(A_i, A_j)}\|_\infty < \infty$ , où  $\|\cdot\|_\infty$  est la norme supremum.

(H6) Il existe  $\beta_0 > 0$  tel que

$$\sup_{a \in S} |\vartheta_a| \leq \beta_0.$$

(H7) Le processus  $\{(A_i, B_i), i \in \mathbb{N}\}$  est quasi-associé au coefficient de covariance  $\Delta_\ell, \ell \in \mathbb{N}$  satisfaisant

$$\exists a > 0 \text{ tel que } \Delta_\ell \leq C e^{-a\ell}.$$

(H8)  $\left\{ \begin{array}{l} \text{(H8.a) La fonction } \Lambda(\dots) \text{ est de classe } \mathcal{C}^1 \text{ sur } S \times [-\beta_0, +\beta_0], \text{ tel que} \\ \inf_{a \in S, \phi \in [-\beta_0, \beta_0]} \frac{\partial \Lambda}{\partial \phi}(a, \phi) > C > 0. \\ \text{(H8.b) Pour chaque fixe } \phi \in [-\beta_0, +\beta_0], \text{ la fonction } \Lambda(\cdot, \phi) \text{ est continue} \\ \text{sur } S. \end{array} \right.$

(H9) La fonction  $\rho$  est strictement monotone par rapport à la seconde composante, Lipschitz et telle que,  $\forall \phi \in [-\beta_0, +\beta_0], \mathbb{E}(\exp(|\rho(B, \phi)|)) \leq C$ , et  $\forall i \neq j, \mathbb{E}(|\rho(B_i, \phi) \rho(B_j, \phi)| | A_i, A_j) \leq C'$ .

(H10) Les moments inverses de la variable de réponse vérifient

$$\text{pour tout } m \geq 2, \mathbb{E}[B^{-m} | A = a] < C < \infty.$$

**Remarque 1.** Notre travail est le lien entre le travail de Kara et al. [115], Attaoui et al. [6] et Bouabssa [30]. Donc, ces différentes hypothèses sont les mêmes considérées dans toutes ces recherches.

## 2.4 Résultats

### 2.4.1 La consistance

Nous étudions maintenant dans cette section la consistance presque complète de  $\hat{\vartheta}_a$ , pour un  $a \in S$ .

**Théorème 2.1.** *Sous les hypothèses (H1)-(H4) et (H5) l'estimateur  $\hat{\vartheta}_a$  existe et est unique. De plus, on a, quand  $n \rightarrow \infty$  avec  $\varpi = \min(\varpi_1, \varpi_2)$ , que*

$$\sup_{a \in S} \sup_{k_n^1 \leq k \leq k_n^2} |\hat{\vartheta}_a - \vartheta_a| = O\left(\varphi_a^{-1}\left(\frac{k_n^2}{n}\right)^\varpi\right) + O\left(\left(\frac{\log n}{n^{-\gamma}(k_n^1)^d}\right)^{1/2}\right), \quad p.co.$$

*Démonstration.* Sous une modification de Kara *et al.* [115], il existe  $\xi \in ]0,1[$ , alors on a que

$$\sum_n \sum_{k=k_n^1}^{k_n^2} \mathbb{P}\left(E_k \leq \varphi_a^{-1}\left(\frac{\xi k_n^1}{n}\right)\right) < \infty, \quad \text{et} \quad \sum_n \sum_{k=k_n^1}^{k_n^2} \mathbb{P}\left(E_k \geq \varphi_a^{-1}\left(\frac{k_n^2}{n\xi}\right)\right) < \infty.$$

Ainsi, on écrit, pour tout  $\varepsilon > 0$ ,

$$\begin{aligned} & \mathbb{P}\left\{\left[\left(\varphi_a^{-1}\left(\frac{k_n^2}{n}\right)\right)^\varpi + \sqrt{\frac{\log n}{n^{-\gamma}(k_n^1)^d}}\right] \sup_{a \in S} \sup_{k_n^1 \leq k \leq k_n^2} |\tilde{\vartheta}_a - \vartheta_a| \geq \varepsilon\right\} \\ & \leq \mathbb{P}\left\{\left[\left(\varphi_a^{-1}\left(\frac{k_n^2}{n}\right)\right)^\varpi + \sqrt{\frac{\log n}{n^{-\gamma}(k_n^1)^d}}\right] \sup_{a \in S} \sup_{k_n^1 \leq k \leq k_n^2} |\tilde{\vartheta}_a - \vartheta_a| \right. \\ & \quad \left. \times \mathbb{I}\left\{\left(\varphi_a^{-1}\left(\frac{\xi k_n^1}{n}\right)\right) \leq h_L \leq \varphi_a^{-1}\left(\frac{k_n^2}{\xi a}\right)\right\} \geq \frac{\varepsilon}{2}\right\} \\ & \quad + \mathbb{P}\left\{h_L \notin \left(\varphi_a^{-1}\left(\frac{\xi k_n^1}{n}\right), \varphi_a^{-1}\left(\frac{k_n^2}{n\xi}\right)\right)\right\}. \end{aligned}$$

Ainsi, il ne reste plus qu'à démontrer les résultats asymptotiques suivants. □

**Proposition 2.4.1.** *Semblable aux conditions du Théorème 2.1, nous avons*

$$\sup_{a \in S} \sup_{x_n \leq h_L \leq y_n} |\hat{\vartheta}_a - \vartheta_a| = O(y_n^\varpi) + O\left(\sqrt{\frac{\log n}{n^{1-\gamma}\varphi_a(x_n)^d}}\right) \quad p.co.$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

où,  $x_n = \varphi_a^{-1} \left( \frac{\xi k_n^1}{n} \right)$  et  $y_n = \varphi_a^{-1} \left( \frac{k_n^2}{n\xi} \right)$ .

*Démonstration.* La preuve repose sur le fait que la deuxième variable de  $\rho$  est purement monotone. Après tout, nous allons juste donner la preuve pour le cas croissant par souci de simplicité. Sous cette hypothèse, on écrit

$$\begin{aligned} \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\widehat{\vartheta}_a - \vartheta_a| &= \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\widehat{\vartheta}_a - \vartheta_a| \mathbb{I}_{\left\{ \sup_{a \in S} |\widehat{\vartheta}_a| \leq \beta_0 \right\}} \\ &+ \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\widehat{\vartheta}_a - \vartheta_a| \mathbb{I}_{\left\{ \sup_{a \in S} |\widehat{\vartheta}_a| > \beta_0 \right\}}. \end{aligned} \quad (2.3)$$

Donc, pour démontrer le résultat, il faut montrer que

$$\sum_n \mathbb{P} \left( \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \widehat{\vartheta}_a < -\beta_0 \right) < \infty, \quad \sum_n \mathbb{P} \left( \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \widehat{\vartheta}_a > \beta_0 \right) < \infty, \quad (2.4)$$

et

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\widehat{\vartheta}_a - \vartheta_a| \mathbb{I}_{\left\{ \sup_{a \in S} |\widehat{\vartheta}_a| \leq \beta_0 \right\}} = O_{p.co.} \left( (y_n^{\overline{\omega}}) + \left( \frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d} \right)^{\frac{1}{2}} \right). \quad (2.5)$$

Puisque  $\widehat{\Lambda}(a, \cdot)$  est croissant pour chaque  $a \in S$ , nous devons démontrer que

$$\sum_n \mathbb{P} \left( \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \widehat{\Lambda}(a, -\beta_0) > 0 \right) < \infty, \quad \text{et} \quad \sum_n \mathbb{P} \left( \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \widehat{\Lambda}(a, \beta_0) < 0 \right) < \infty.$$

L'hypothèse (H6) implique que

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \Lambda(a, -\beta_0) < 0, \quad \text{et} \quad \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \Lambda(a, \beta_0) > 0.$$

En supposant qu'on puisse vérifier

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \widehat{\Lambda}(a, -\beta_0) \longrightarrow \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \Lambda(a, -\beta_0),$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

on obtient

$$\sum_n \mathbb{P} \left( \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \hat{\Lambda}(a, -\beta_0) > 0 \right) \leq \sum_n \mathbb{P} \left( \left| \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \hat{\Lambda}(a, -\beta_0) - \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \Lambda(a, -\beta_0) \right| \geq \epsilon_1 \right) < \infty,$$

et

$$\sum_n \mathbb{P} \left( \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \hat{\Lambda}(a, \beta_0) < 0 \right) \leq \sum_n \mathbb{P} \left( \left| \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \hat{\Lambda}(a, \beta_0) - \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \Lambda(a, \beta_0) \right| \geq \epsilon_2 \right) < \infty,$$

avec,  $\epsilon_1 = - \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \Lambda(a, -\beta_0)$  et  $\epsilon_2 = \inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} \Lambda(a, \beta_0)$ .

De plus, on écrit sous (H8.a) que

$$(\widehat{\vartheta}_a - \vartheta_a) \mathbb{I}_{\{|\widehat{\vartheta}_a - \vartheta_a| \leq \beta\}} = \frac{\Lambda(a, \widehat{\vartheta}_a) - \hat{\Lambda}(a, \widehat{\vartheta}_a)}{\frac{\partial \Lambda}{\partial t}(a, \alpha_n)} \mathbb{I}_{\{|\widehat{\vartheta}_a - \vartheta_a| \leq \beta\}},$$

où,  $\alpha_n$  est entre  $\widehat{\vartheta}_a$  et  $\vartheta_a$ . En conséquence, la seule chose qui reste à révéler est le taux de convergence de

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}(a, \phi) - \Lambda(a, \phi)|. \quad (2.6)$$

La vérification de (2.6) se concentre sur la décomposition ci-dessous.

$$\begin{aligned} & \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}(a, \phi) - \Lambda(a, \phi)| \leq \\ & \frac{1}{\inf_{x_n \leq h_L \leq y_n} \inf_{a \in S} |\hat{\Lambda}_D(a)|} \left\{ \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N(a, \phi) - \mathbb{E}[\hat{\Lambda}_N(a, \phi)] \right. \\ & + \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E}[\hat{\Lambda}_N(a, \phi)] - H(a, \phi)| \\ & + \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\Lambda(a, \phi) (\Upsilon(a) - \mathbb{E}[\hat{\Lambda}_D(a)])| \\ & \left. + \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\Lambda(a, \phi) (\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a))| \right\}, \end{aligned} \quad (2.7)$$

où,

$$\begin{aligned}\hat{\Lambda}_N(a, \phi) &:= \frac{1}{n\mathbb{E}[L_1]^d} \sum_{i=1}^n L(h_L^{-1}(a - A_i)) \rho_a(B_i, \phi), \\ \hat{\Lambda}_D(a) &:= \frac{1}{n\mathbb{E}[L_1]^d} \sum_{i=1}^n L(h_L^{-1}(a - A_i)), \\ \text{et } H(a, \phi) &:= \Lambda(a, \phi)\Upsilon(a).\end{aligned}$$

Les lemmes et le corollaire qui suivent sont utilisés pour démontrer la Proposition 2.4.1.  $\square$

**Lemme 2.1.** *Sous les hypothèses (H2), (H4)-(H5) et (H8), quand  $n \rightarrow \infty$ , on a que*

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E}[\hat{\Lambda}_N(a, \phi)] - H(a, \phi)| = O(y_n^{\overline{\omega}}).$$

*Démonstration.* On a,

$$\begin{aligned}\mathbb{E}[\hat{\Lambda}_N(a, \phi)] &= \frac{1}{\mathbb{E}[L_1]^d} \int_{\mathbb{R}^d} \mathbb{E}[\rho(B, \phi) | (A = u)] L\left(\frac{a - u}{h_L}\right) \Upsilon(u) du \\ &= \frac{1}{\mathbb{E}[L_1]^d} \int_{\mathbb{R}^d} \Lambda(u, \phi) L\left(\frac{a - u}{h_L}\right) \Upsilon(u) du \\ &= \int_{\mathbb{R}^d} H(a - h_L z, \phi) L(z) dz.\end{aligned}$$

Nous utilisons le fait que  $\mathbb{E}[L_1] \leq C\varphi_a(h_L)$ , alors puisque  $\Upsilon$  et  $\Lambda$  sont de classe  $\mathcal{C}^1$ , un développement de Taylor de  $H(a - h_L z, \phi)$  avec (H2), permet d'écrire

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E}[\hat{\Lambda}_N(a, \phi)] - H(a, \phi)| = O(y_n^{\overline{\omega}}).$$

$\square$

**Lemme 2.2.** *Sous les hypothèses (H1) (H4), (H5)-(H7), on a*

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N(a, \phi) - \mathbb{E}[\hat{\Lambda}_N(a, \phi)]| = O\left(\sqrt{\frac{\log n}{n^{1-\gamma}\varphi_a(x_n)^d}}\right) \quad p.co.$$

*Démonstration.* Nous utilisons une approche de troncature en introduisant la variable aléatoire suivante, car  $\rho$  peut ne pas être borné,

$$\hat{\Lambda}_N^*(a, \phi) = \frac{1}{n\mathbb{E}[L_1]^d} \sum_{i=1}^n L(h_L^{-1}(a - A_i)) \rho_a(B_i, \phi) \mathbb{I}_{|\rho_a(B_i, \phi)| < \theta_n} \quad \text{avec } \theta_n = n^{\gamma/6}.$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Le résultat défini est alors une conséquence des résultats intermédiaires qui suivent,

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E} [\hat{\Lambda}_N^*(a, \phi)] - \mathbb{E} [\hat{\Lambda}_N(a, \phi)]| = O_{\text{a.co}} \left( \sqrt{\frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d}} \right), \quad (2.8)$$

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a, \phi) - \hat{\Lambda}_N(a, \phi)| = O_{\text{a.co}} \left( \sqrt{\frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d}} \right), \quad (2.9)$$

et

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a, \phi) - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| = O_{\text{p.co.}} \left( \sqrt{\frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d}} \right). \quad (2.10)$$

On commence par prouver (2.10). Comme  $S$  est compact on écrit

$$S \subset \bigcup_{j=1}^{\wp_n} B(a_\ell, \tau_n),$$

avec  $\wp_n = O(n^\beta)$  et  $\tau_n = O(\wp_n^{-1})$ , où  $t = \frac{\delta(d+2)}{2} + \frac{1}{2} + \frac{\gamma}{6}$  et  $\delta \leq (1 - \gamma - \xi_2)/d$ .

Considérons maintenant tout  $a \in S$ ,

$$\ell(a) = \arg \min_{\ell \in \{1, \dots, \wp_n\}} \|a - a_\ell\|,$$

et on considère la décomposition suivante,

$$\begin{aligned} & \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a, \phi) - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| \\ & \leq \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a, \phi) - \hat{\Lambda}_N^*(a_{\ell(a)}, \phi)|}_{Z_1} \\ & + \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a_{\ell(a)}, \phi) - \mathbb{E} [\hat{\Lambda}_N^*(a_{\ell(a)}, \phi)]|}_{Z_2} \\ & + \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E} [\hat{\Lambda}_N^*(a_{\ell(a)}, \phi)] - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]|}_{Z_3}. \end{aligned} \quad (2.11)$$

— Premièrement, nous utilisons la compacité de  $[-\beta_0, \beta_0]$ , pour  $Z_2$  et on écrit

$$[-\beta_0, \beta_0] \subset \bigcup_{j=1}^{z_n} (\phi_j - g_n, \phi_j + g_n), \quad (2.12)$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

avec  $g_n = n^{-1/2}$  et  $z_n = O(n^{1/2})$ . Définir

$$\mathcal{N}_n = \{\phi_j - g_n, \phi_j + g_n, 1 \leq j \leq z_n\}. \quad (2.13)$$

Pour  $1 \leq j \leq z_n$ , par  $\mathbb{E}[\hat{\Lambda}_N^*(a, \cdot)]$  et  $\hat{\Lambda}_N^*(a, \cdot)$  monotonie, on obtient

$$\mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi_j - g_n)] \leq \sup_{\phi \in (\phi_j - g_n, \phi_j + g_n)} \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi)] \leq \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi_j + g_n)] \quad (2.14)$$

$$\hat{\Lambda}_N^*(a_{\ell(a)}, \phi_j - g_n) \leq \sup_{\phi \in (\phi_j - g_n, \phi_j + g_n)} \hat{\Lambda}_N^*(a_{\ell(a)}, \phi) \leq \hat{\Lambda}_N^*(a_{\ell(a)}, \phi_j + g_n). \quad (2.15)$$

De plus, par hypothèse (H9), pour tout  $\phi_1, \phi_2 \in [-\beta_0, \beta_0]$ , on a que

$$|\mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi_1)] - \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi_2)]| \leq C|\phi_1 - \phi_2|. \quad (2.16)$$

Par conséquent, nous pouvons déduire de (2.12) et (2.16) que

$$\begin{aligned} & \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a_{\ell(a)}, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell(a)}, \phi)]| \\ & \leq \max_{1 \leq \ell \leq \rho_n} \max_{1 \leq j \leq z_n} \max_{\phi \in \{\phi_j - g_n, \phi_j + g_n\}} |\hat{\Lambda}_N^*(x_{\ell}, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell}, \phi)]| + 2Cg_n \end{aligned} \quad (2.17)$$

Nous le découvrirons à l'aide d'une simple équation algébrique

$$g_n = o\left(\sqrt{\frac{\log n}{n\varphi_a(x_n)^d}}\right). \quad (2.18)$$

Alors, pour un réel positif  $\eta$  suffisamment grand, il suffit de démontrer que

$$\sup_{x_n \leq h_L \leq y_n} \max_{1 \leq \ell \leq \rho_n} \max_{1 \leq j \leq z_n} \max_{\phi \in \{\phi_j - g_n, \phi_j + g_n\}} |\hat{\Lambda}_N^*(a_{\ell}, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell}, \phi)]| = O_{a.co} \left( \sqrt{\frac{\log n}{n^{1-\gamma}\varphi_a(x_n)^d}} \right). \quad (2.19)$$

Pour ce faire, nous utilisons une inégalité de type Bernstein pour les variables aléatoires dépendantes, après tout, nous avons écrit

$$\hat{\Lambda}_N^*(a_{\ell}, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a_{\ell}, \phi)] = \sum_{i=1}^n \Omega_i,$$

où

$$\Omega_i = \frac{1}{n\mathbb{E}[L_1]} \Psi(A_i, B_i),$$

avec

$$\begin{aligned} \Psi(u, v) = & \rho(v, \phi) L(h_L^{-1}(a_\ell - u)) \mathbb{I}_{(|\rho(v, \phi)| < \theta_n)} \\ & - \mathbb{E} \left[ \rho(B_1, \phi) L(h_L^{-1}(a_\ell - A_1)) \mathbb{I}_{(|\rho(B_1, \phi)| < \theta_n)} \right], u \in \mathbb{R}^d, v \in \mathbb{R}. \end{aligned}$$

Donc,

$$\begin{aligned} \|\Psi\|_\infty & \leq C\theta_n \|L\|_\infty \text{ et} \\ \text{Lip}(\Psi) & \leq (\|L\|_\infty \text{Lip}(\rho) + \theta_n h_L^{-1} \text{Lip}(L)) \leq C\theta_n h_L^{-1} \text{Lip}(L). \end{aligned}$$

L'inégalité de Newmann et Kallabis est centrée sur l'analyse asymptotique de  $\text{Var} \left( \sum_{i=1}^n \Omega_i \right)$  et  $\text{Cov}(\Omega_{m_1} \dots \Omega_{m_u}, \Omega_{\phi_1} \dots \Omega_{\phi_v})$ , pour tout  $(m_1, \dots, m_u) \in \mathbb{N}^u$  et  $(\phi_1, \dots, \phi_v) \in \mathbb{N}^v$ . On commence par étudier le terme de variance,

$$\text{Var} \left( \sum_{i=1}^n \Omega_i \right) = n \text{Var}(\Omega_1) + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov}(\Omega_i, \Omega_j). \quad (2.20)$$

Sous (H9), on a

$$\begin{aligned} \text{Var}(\Omega_1) & \leq \frac{1}{n^2 \mathbb{E}[L_1]^{2d}} \mathbb{E} [ |\rho(B_1, \phi) L_1(a_\ell)|^2 ] \leq C' \frac{1}{n^2 \mathbb{E}[L_1]^{2d}} \mathbb{E} [ |L_1(a_\ell)|^2 ] \\ & \leq C' n^{-2} \mathbb{E}[L_1]^{-d} \int_{\mathbb{R}^d} L^2(u) \Upsilon(a_\ell - h_L u) du. \end{aligned} \quad (2.21)$$

Pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , on obtient que

$$\text{Var}(\Omega_1) = O \left( n^{-2} \varphi_a(x_n)^{-d} \right). \quad (2.22)$$

Maintenant, évaluons le comportement asymptotique de la somme dans le côté-droit de (2.20). Pour cela nous utilisons la technique développée par Masry [146]. En effet, nous avons besoin de la décomposition suivante

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov}(\Omega_i, \Omega_j) = \sum_{i=1}^n \sum_{\substack{j=1 \\ 0 < |i-j| \leq \varrho_n}}^n \text{Cov}(\Omega_i, \Omega_j) + \sum_{i=1}^n \sum_{\substack{j=1 \\ |i-j| > \varrho_n}}^n \text{Cov}(\Omega_i, \Omega_j),$$

avec  $(\varrho_n)$  est une suite d'entiers positifs qui tend vers l'infini lorsque  $n$  tend vers l'infini.

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Pour  $|i - j| \leq \varrho_n$ , on utilise (H2) – (H5) et (H9), pour obtenir

$$\begin{aligned}
 \mathbb{E} [|\Omega_i \Omega_j|] &\leq C \frac{1}{n^2 \mathbb{E} [L_1]^{2d}} \left( \mathbb{E} [|\rho(B_i, \phi) L_i(a_\ell) \rho(B_j, z) L_j(a_\ell)|] \right. \\
 &\quad \left. + (\mathbb{E} [|\rho(B_1, \phi) L_1(a_\ell)|])^2 \right), \\
 &\leq C \frac{1}{n^2 \mathbb{E} [L_1]^{2d}} \left( \mathbb{E} [L_i(a_\ell) L_j(a_\ell)] + (\mathbb{E} [L_1(a_\ell)])^2 \right), \\
 &\leq \frac{1}{n^2} \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} L(u) L(v) \Upsilon_{(A_i, A_j)}(a_\ell - h_L u, a_\ell - h_L v) dudv \right) \\
 &\quad + \left( \int_{\mathbb{R}^d} L(u) \Upsilon(a_\ell - h_L u) du \right)^2 \\
 &= O(n^{-2}).
 \end{aligned}$$

On obtient ainsi

$$\begin{aligned}
 \sum_{i=1}^n \sum_{\substack{j=1 \\ 0 < |i-j| \leq \varrho_n}}^n \text{Cov}(\Omega_i, \Omega_j) &\leq n \varrho_n (\mathbb{E} [|\Omega_i \Omega_j|]), \\
 &\leq C n^{-1} \varrho_n.
 \end{aligned}$$

Par contre, pour  $|i - j| > \varrho_n$ , on utilise (H7) et la quasi-association de la suite  $(A_i, B_i)$  pour écrire

$$\begin{aligned}
 \sum_{i=1}^n \sum_{\substack{j=1 \\ |i-j| > \varrho_n}}^n \text{Cov}(\Omega_i, \Omega_j) &\leq \theta_n^2 n^{-2} \mathbb{E} [L_1]^{-2(d+1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ |i-j| > \varrho_n}}^n \Delta_{i,j} \\
 &\leq \theta_n^2 n^{-1} \mathbb{E} [L_1]^{-2(d+1)} \Delta_{\varrho_n} \\
 &\leq \theta_n^2 n^{-1} \mathbb{E} [L_1]^{-2(d+1)} e^{-a \varrho_n}.
 \end{aligned}$$

Donc

$$\sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \text{Cov}(\Omega_i, \Omega_j) \leq C \left( n^{-1} \varrho_n + \theta_n^2 n^{-1} \mathbb{E} [L_1]^{-2(d+1)} e^{-a \varrho_n} \right).$$

Prenez  $\varrho_n = \frac{1}{a} \log \left( a \theta_n^2 \mathbb{E} [L_1]^{-2(d+1)} \right)$ . Alors, par (H4), on obtient

$$n \mathbb{E} [L_1]^d \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\Omega_i, \Omega_j) \rightarrow 0. \tag{2.23}$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Enfin, en mélangeant (2.22) et (2.23), on obtient pour tout  $\sup_{x_n \leq h_L \leq y_n}$  que

$$\text{Var} \left( \sum_{i=1}^n \Omega_i \right) = O \left( \frac{1}{n \varphi_a(x_n)^d} \right).$$

Maintenant, pour tout  $(m_1, \dots, m_u) \in \mathbb{N}^u$  et  $(\phi_1, \dots, \phi_v) \in \mathbb{N}^v$ , nous traitons avec le terme de covariance dans (2.20). Pour cela, on considère les cas suivants,

— Si  $\phi_1 = m_u$ , on obtient

$$\begin{aligned} |\text{Cov}(\Omega_{m_1} \dots \Omega_{m_u}, \Omega_{\phi_1} \dots \Omega_{\phi_v})| &\leq \left( \frac{C\theta_n \|L\|_\infty}{n \mathbb{E}[L_1]^d} \right)^{u+v} \mathbb{E}[L_1^2(a_\ell)] \\ &\leq \mathbb{E}[L_1]^d \left( \frac{C\theta_n}{n \mathbb{E}[L_1]^d} \right)^{u+v}. \end{aligned}$$

— Si  $\phi_1 > m_u$ , on utilise la condition de quasi-association, et on obtient

$$\begin{aligned} &|\text{Cov}(\Omega_{m_1} \dots \Omega_{m_u}, \Omega_{\phi_1} \dots \Omega_{\phi_v})| \\ &\leq \left( \frac{\theta_n \mathbb{E}[L_1]^{-1} \text{Lip}(L)}{n \mathbb{E}[L_1]^d} \right)^2 \left( \frac{2\theta_n \|L\|_\infty}{n \mathbb{E}[L_1]^d} \right)^{u+v-2} \sum_{i=1}^u \sum_{j=1}^v \Delta_{m_i, \phi_j} \\ &\leq \mathbb{E}[L_1]^{-2} \left( \frac{C\theta_n}{n \mathbb{E}[L_1]^d} \right)^{u+v} v \Delta_{\phi_1 - m_u} \\ &\leq \mathbb{E}[L_1]^{-2} \left( \frac{C\theta_n}{n \mathbb{E}[L_1]^d} \right)^{u+v} v e^{-a(\phi_1 - m_u)}. \end{aligned} \tag{2.24}$$

D'une autre part, on a

$$\begin{aligned} |\text{Cov}(\Omega_{m_1} \dots \Omega_{m_u}, \Omega_{\phi_1} \dots \Omega_{\phi_v})| &\leq \left( \frac{2C\theta_n \|L\|_\infty}{n \mathbb{E}[L_1]^d} \right)^{u+v-2} (|\mathbb{E}\Omega_{m_u} \Omega_{\phi_1}|) \\ &\leq \left( \frac{C\theta_n}{n \mathbb{E}[L_1]^d} \right)^{u+v} \mathbb{E}[L_1]^{2d}. \end{aligned} \tag{2.25}$$

Ensuite, nous prenons la  $\frac{d}{2d+2}$ -puissance de (2.24) et la  $\frac{d+2}{2d+2}$ -puissance de (2.25)

$$|\text{Cov}(\Omega_{m_1}, \dots, \Omega_{m_u}, \Omega_{\phi_1}, \dots, \Omega_{\phi_v})| \leq \mathbb{E}[L_1]^d \left( \frac{C\theta_n}{n \mathbb{E}[L_1]^d} \right)^{u+v} v e^{-\frac{ad}{2d+2}(\phi_1 - m_u)}.$$

On applique le Théorème (2.1) de Kallabis et Neumann [114] pour les va-

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

riables  $\Omega_i, i = 1, \dots, n$ , avec  $\sup_{x_n \leq h_L \leq y_n}$ ,

$$L_n = \frac{C\theta_n}{n\sqrt{h^d}}, M_n = \frac{C\theta_n}{n\mathbb{E}[L_1]^d} \text{ et } A_n = \text{Var}\left(\sum_{i=1}^n \Omega_i\right) = O\left(\frac{1}{n\varphi_a(x_n)^d}\right).$$

Maintenant, nous appliquons l'inégalité de Bernstein pour les processus empiriques,

$$\mathbb{P}\left\{\sup_{a_n \leq h_L \leq b_0} \sqrt{\frac{n^{1-\gamma}\varphi_a(h)}{\log n}} |\hat{\Lambda}_N^*(a_\ell, t) - \mathbb{E}[\hat{\Lambda}_N^*(a_\ell, t)]| \geq \eta_0\right\} \leq \log(n)n^{-C'\eta_0^2}.$$

Ensuite, en utilisant le fait que  $\varphi_n z_n \leq n^\zeta$ , où  $\zeta = \gamma + \frac{1}{2}$ , pour des valeurs appropriées de  $\gamma$  et  $\eta_0$ , nous aurons pour chaque  $h_L \in (a_n, b_0)$ ,

$$\mathbb{P}\left(\max_{1 \leq k \leq \varphi_n} \max_{1 \leq j \leq z_n} \max_{\phi \in \{\phi_j - g_n, \phi_j + g_n\}} |\hat{\Lambda}_N^*(a, z) - \mathbb{E}[\hat{\Lambda}_N^*(a, z)]| = O\sqrt{\frac{\log n}{n^{1-\gamma}\varphi(x_n)^d}}\right).$$

La preuve de (2.10) est ainsi achevée.

— Deuxièmement, en ce qui concerne  $Z_1$  et  $Z_3$ , la condition de Lipschitz sur le noyau  $L$  dans (H2) permet d'écrire directement, pour tout  $a \in S$ , et  $\forall \phi \in [-\beta_0, \beta_0]$  que

$$\begin{aligned} |\hat{\Lambda}_N^*(a, t) - \hat{\Lambda}_N^*(a_{\ell(a)}, t)| &= \frac{\theta_n}{n\mathbb{E}[L_1]^d} \left| \sum_{i=1}^n L_i(a) - \sum_{i=1}^n L_i(a_{\ell(a)}) \right| \\ &\leq \frac{\theta_n}{\mathbb{E}[L_1]^{d+1}} \|a - a_{\ell(a)}\| \\ &\leq \frac{C\theta_n\tau_n}{\mathbb{E}[L_1]^{d+1}}. \end{aligned}$$

Alors, on en choisissant  $\tau_n = O(n^{-\gamma})$ , on obtient

$$\frac{\tau_n}{\mathbb{E}[L_1]^{d+1}} = O_{a.co} \left( \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right). \quad (2.26)$$

Ainsi,

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\hat{\Lambda}_N^*(a, \phi) - \hat{\Lambda}_N^*(a_{\ell(a)}, \phi)| = O_{a.co} \left( \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right) \quad (2.27)$$

et

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E} [\hat{\Lambda}_N^*(a_{\ell(a)}, \phi)] - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| = O \left( \sqrt{\frac{\log n}{n \varphi_a(x_n)^d}} \right) \quad (2.28)$$

Nous démontrons maintenant (2.8). Pour tout  $a \in S$  et tout  $\phi \in [-\beta_0, \beta_0]$ , avec

$\sup_{x_n \leq h_L \leq y_n}$ , nous pouvons écrire

$$\begin{aligned} |\mathbb{E} [\hat{\Lambda}_N(a, \phi)] - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| &= \frac{1}{n \mathbb{E} [L_1]^d} \left| \mathbb{E} \left[ \sum_{i=1}^n \rho(B_i, \phi) \mathbb{I}_{\{|\rho(B_i, \phi)| > \theta_n\}} L_i \right] \right| \\ &\leq \mathbb{E} [L_1]^{-d} \mathbb{E} [|\rho(B_1, \phi)| \mathbb{I}_{\{|\rho(B_1, \phi)| > \theta_n\}} L_1] \\ &\leq \mathbb{E} [L_1]^{-d} \mathbb{E} [\exp(|\rho(B_1, \phi)|/4) \mathbb{I}_{\{|\rho(B_1, \phi)| > \theta_n\}} L_1]. \end{aligned}$$

On utilise l'inégalité de Holder pour obtenir

$$\begin{aligned} &\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E} [\hat{\Lambda}_N(a, \phi)] - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| \\ &\leq \mathbb{E} [L_1]^{-d} \left( \mathbb{E} [\exp(|\rho(B_1, \phi)|/2) \mathbb{I}_{\{|\rho(B_1, \phi)| > \theta_n\}}] \right)^{\frac{1}{2}} \left( \mathbb{E} (L_1^2) \right)^{\frac{1}{2}} \quad (2.29) \\ &\leq \mathbb{E} [L_1]^{-d} \exp(-\theta_n/4) \left( \mathbb{E} [\exp(|\rho(B_1, \phi)|)] \right)^{\frac{1}{2}} \left( \mathbb{E} (L_1^2) \right)^{\frac{1}{2}} \\ &\leq C \varphi_a(x_n)^{-\frac{d}{2}} \exp(-\theta_n/4). \end{aligned}$$

Alors, puisque  $\theta_n = n^{\gamma/6}$ , on peut écrire

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} \sup_{\phi \in [-\beta_0, \beta_0]} |\mathbb{E} [\hat{\Lambda}_N(a, \phi)] - \mathbb{E} [\hat{\Lambda}_N^*(a, \phi)]| = O \left( \left( \frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d} \right)^{1/2} \right).$$

L'inégalité de Markov est utilisée pour démontrer le dernier résultat énoncé (2.9).

De plus, pour tout  $\ell, \phi \in \mathcal{N}_n$ ,  $\epsilon > 0$ , et pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , on peut observer que

$$\begin{aligned} &\sum \mathbb{P} \left( \sup_{x_n \leq h_L \leq y_n} \sup_{n \geq 1} \sup_{a \in S} |\hat{\Lambda}_N(a, \phi) - \hat{\Lambda}_N^*(a, \phi)| > \epsilon_0 \left( \sqrt{\frac{\log n}{n^{1-\gamma} \varphi_a(x_n)^d}} \right) \right) \quad (2.30) \\ &\leq C \sum_{n \geq 1} n \exp(-\theta_n). \end{aligned}$$

La preuve de (2.9) est complétée en utilisant la définition de  $\theta_n$ , qui à son tour complète la démonstration de ce Lemme.

□

**Lemme 2.3.** *Sous les hypothèses du Lemme 2.2, on a*

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a)| = O\left(\sqrt{\frac{\log n}{n\varphi_a(x_n)^d}}\right).$$

*Démonstration.* On utilise alors la compacité de  $S$ , de plus des notations utilisées dans le Lemme 2.2, on a

$$\begin{aligned} & \sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\hat{\Lambda}_D(a) - \mathbb{E}[\hat{\Lambda}_D(a)]| \leq \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\hat{\Lambda}_D(a) - \hat{\Lambda}_D(a_\ell)|}_{Z'_1} \\ & + \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\hat{\Lambda}_D(a_\ell) - \mathbb{E}[\hat{\Lambda}_D(a_\ell)]|}_{Z'_2} + \underbrace{\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a_\ell)] - \mathbb{E}[\hat{\Lambda}_D(a)]|}_{Z'_3}. \end{aligned}$$

— Pour  $Z'_1$  et  $Z'_3$ , pour tout  $a \in S$  et pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , la condition de Lipschitz sur le noyau  $L$  permet d'écrire,

$$\begin{aligned} |\hat{\Lambda}_D(a) - \hat{\Lambda}_D(a_\ell)| &= \frac{1}{n\mathbb{E}[L_1]^d} \left| \sum_{i=1}^n L_i(a) - \sum_{i=1}^n L_i(a_\ell) \right| \\ &\leq \frac{C}{\varphi_a(x_n)^{d+1}} \|a - a_\ell\| \\ &\leq \frac{C\tau_n}{\varphi_a(x_n)^{d+1}}. \end{aligned}$$

Du fait que  $\frac{\tau_n}{\varphi_a(x_n)^{d+1}} = O\left(\sqrt{\frac{\log n}{n\varphi_a(x_n)^d}}\right)$ , il en résulte,

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\hat{\Lambda}_D(a) - \hat{\Lambda}_D(a_\ell)| = O\left(\sqrt{\frac{\log n}{n\varphi_a(x_n)^d}}\right), \quad \text{p.co.}, \quad (2.31)$$

et

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a_\ell)] - \mathbb{E}[\hat{\Lambda}_D(a)]| = O\left(\sqrt{\frac{\log n}{n\varphi_a(x_n)^d}}\right), \quad \text{p.co.} \quad (2.32)$$

— Pour tout réel  $\eta > 0$ , nous avons pour  $Z'_2$ , que

$$\begin{aligned}
 & \mathbb{P} \left( Z'_2 > \eta \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right) \\
 & \leq \mathbb{P} \left( \sup_{x_n \leq h_L \leq y_n} \max_{\ell \in \{1, \dots, \rho_n\}} |\hat{\Lambda}_D(a_\ell) - \mathbb{E}[\hat{\Lambda}_D(a_\ell)]| > \eta \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right) \quad (2.33) \\
 & \leq \varphi_n \sup_{x_n \leq h_L \leq y_n} \max_{\ell \in \{1, \dots, \varphi_n\}} \mathbb{P} \left( |\hat{\Lambda}_D(a_\ell) - \mathbb{E}[\hat{\Lambda}_D(a_\ell)]| > \eta \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right).
 \end{aligned}$$

En utilisant la même procédure que dans le Lemme 2.2, et en changeant  $\Lambda$  par 1, il vient que

$$\sup_{x_n \leq h_L \leq y_n} \mathbb{P} \left( \sup_{a \in S} |\hat{\Lambda}_D(a) - \mathbb{E}[\hat{\Lambda}_D(a)]| > \eta \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right) \leq n^{\gamma - C\eta^2}.$$

Avec un choix judicieux de  $\eta$ , on peut obtenir

$$\sup_{x_n \leq h_L \leq y_n} \sum_{n=1}^{\infty} \mathbb{P} \left( \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a)| > \eta \sqrt{\frac{\log n}{n\varphi_a(x_n)^d}} \right) < \infty.$$

□

**Lemme 2.4.** *Sous les hypothèses du Lemme 2.1 on a*

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a) - \Upsilon(a)]| = O(y_n^{\varpi}).$$

*Démonstration.* Il suffit juste d'utiliser le même raisonnement que dans le Lemme 2.1,

$$\begin{aligned}
 \mathbb{E}[\hat{\Lambda}_D(a)] &= \frac{1}{\mathbb{E}[L_1]^d} \int_{\mathbb{R}^d} L\left(\frac{a-u}{h_L}\right) \Upsilon(u) du \\
 &= \frac{1}{\mathbb{E}[L_1]} \mathbb{E} \left[ L\left(\frac{d(a, A_1)}{h_L}\right) \mathbb{I}_{B(a, h_L)} \Upsilon(A) \right].
 \end{aligned}$$

Ensuite, en utilisant un raisonnement analytique, nous arrivons à la conclusion suivante,

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a)] - \Upsilon(a)| = \frac{1}{\mathbb{E}[L_1]} \mathbb{E} \left[ L\left(\frac{d(a, A_1)}{h_L}\right) \mathbb{I}_{B(a, h_L)} |\Upsilon(A) - \Upsilon(a)| \right].$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Par la condition de Lipschitz (H10), il vient

$$\mathbb{I}_{\mathcal{B}(a, h_L/2)} |\Upsilon(A) - \Upsilon(a)| \leq C(h_L/2)^\varpi.$$

Nous obtenons le résultat demandé en opérant le changement suivant

$$\sup_{x_n \leq h_L \leq y_n} \sup_{a \in S} |\mathbb{E}[\hat{\Lambda}_D(a)] - \Upsilon(a)| \leq C y_n^\varpi.$$

Le Lemme est ainsi démontré. □

**Corollaire 2.1.1.** *Sous les hypothèses du Lemme 2.3, on a*

$$\exists C > 0 \sum_{n=1}^{\infty} \mathbb{P} \left( \inf_{a \in S} \inf_{x_n \leq h_L \leq a_n} \hat{\Lambda}_2(a) < C \right) < \infty.$$

*Démonstration.* Sous la condition (H10), on a

$$\mathbb{E}[\hat{\Lambda}_D(a)] \rightarrow \Lambda_D(a) > 0.$$

Par conséquent, il existe une constante  $C > 0$ , pour  $n$  suffisamment grand,

$$\mathbb{E}[\hat{\Lambda}_D(a)] \geq C \text{ pour tout } h_L \in (x_n, y_n).$$

Alors,

$$\inf_{a \in S} \inf_{h_L \in (x_n, y_n)} \hat{\Lambda}_1(a) \leq \frac{C}{2} \Rightarrow \exists h_L \in (x_n, y_n) \text{ tel que } |\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a)| \geq \frac{C}{2},$$

ce qui nous permet d'écrire

$$\sup_{a \in S} \sup_{h_L \in (a_n, b_n)} |\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a)| \geq \frac{C}{2},$$

$$\mathbb{P} \left( \inf_{a \in S} \inf_{h_L \in (x_n, y_n)} \hat{\Lambda}_D(a) \leq \frac{C}{2} \right) \leq \mathbb{P} \left( \sup_{a \in S} \sup_{h_L \in (x_n, y_n)} |\mathbb{E}[\hat{\Lambda}_D(a)] - \hat{\Lambda}_D(a)| \geq \frac{C}{2} \right).$$

A partir des résultats des Lemmes 2.3 et 2.4, nous obtenons le résultat souhaité. □

**Lemme 2.5.** *Sous les hypothèses du Théorème 2.1,  $\hat{\vartheta}_a$  existe et il est unique presque sûrement pour  $n$  assez grand.*

*Démonstration.* La monotonie stricte de  $\Lambda$  implique :

$$\sup_{k_n^1 \leq k \leq k_n^2} \Lambda(a, \vartheta_a - \epsilon) \leq \sup_{k_n^1 \leq k \leq k_n^2} \Lambda(a, \vartheta_a) \leq \sup_{k_n^1 \leq k \leq k_n^2} \Lambda(a, \vartheta_a + \epsilon).$$

Les Lemmes 2.1 et 2.3 et le Corollaire 2.1.1 montrent que pour tout réel  $\phi$  fixé, on a

$$\sup_{k_n^1 \leq k \leq k_n^2} \hat{\Lambda}(a, \phi) - \sup_{k_n^1 \leq k \leq k_n^2} \Lambda(a, \phi) \rightarrow 0, \quad \text{p.co.}$$

On obtient donc, pour un  $n$  suffisamment grand que

$$\sup_{k_n^1 \leq k \leq k_n^2} \hat{\Lambda}(a, \vartheta_a - \epsilon) \leq 0 \leq \sup_{k_n^1 \leq k \leq k_n^2} \hat{\Lambda}(a, \vartheta_a + \epsilon), \quad \text{p.co.}$$

Notons que  $\hat{\Lambda}(i, \phi)$  est continue pour tout  $\phi$  puisque  $\rho$  et  $L$  sont des fonctions continues.

Il existe alors  $\phi_0 = \hat{\vartheta}_a$  dans l'intervalle  $[\vartheta_a - \epsilon, \vartheta_a + \epsilon]$  pour tout  $\sup_{k_n^1 \leq k \leq k_n^2} \hat{\Lambda}(a, \hat{\vartheta}_a) = 0$ .

Enfin, l'unicité de  $\hat{\vartheta}_a$  est le résultat direct de la stricte monotonie de  $\Lambda$  et de la positivité de  $L$ . □

## 2.4.2 Normalité asymptotique

La consistance presque complète de  $\hat{\vartheta}_a$ , pour un  $a \in S$  étant établi, nous étudions dans cette partie la normalité asymptotique de  $\hat{\vartheta}_a$  pour un fixe  $a \in \mathcal{A}$ .

**Théorème 2.2.** *Supposons que (H4) – (H5) soit vérifiée, alors on a*

$$\left( \frac{n\varphi_a(x_n)^d}{\sigma^2(a, \vartheta_a)} \right)^{1/2} (\hat{\vartheta}_a - \vartheta_a) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty,$$

où,

$$\sigma^2(a, \vartheta_a) = \frac{\mathbb{E}[\Lambda_a^2(B\vartheta_a) | A = a]}{\left( \frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a) \right)^2} \int_{\mathbb{R}'} L^2(z) dz,$$

$$\mathcal{A} = \left\{ a \in S, \mathbb{E}[\Lambda_a^2(B, \vartheta_a) | A = a] \frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a) \neq 0 \right\},$$

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

*Démonstration.* Nous présentons la démonstration identique au Théorème 2.1, dans le cas où la fonction  $\rho$  est croissante le cas décroissant est obtenu pour  $-\rho$ .

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Nous définissons pour tout  $u \in \mathbb{R}$ ,

$$\phi = \vartheta_a + u \left[ \hat{n}\varphi_a(x_n)^d \right]^{-1/2} \sigma(a, \vartheta_a).$$

On pourra donc écrire, si  $\hat{\Lambda}_D(a) \neq 0$ , que

$$\begin{aligned} & \mathbb{P} \left\{ \left( \frac{n\varphi_a(x_n)^d}{\sigma^2(a, \vartheta_a)} \right)^{1/2} (\hat{\vartheta}_a - \vartheta_a) < u \right\} \\ &= \mathbb{P} \left\{ \hat{\vartheta}_a < \vartheta_a + u \left[ \hat{n}\varphi_a(x_n)^d \right]^{-1/2} \sigma(a, \vartheta_a) \right\} \\ &= \mathbb{P} \{ 0 < \hat{\Lambda}_N(a, \phi) \} \\ &= \mathbb{P} \{ \mathbb{E} [\hat{\Lambda}_N(a, \phi)] - \hat{\Lambda}_N(a, \phi) < \mathbb{E} [\hat{\Lambda}_N(a, \phi)] \}. \end{aligned}$$

Par conséquent, le Théorème 2.2 découle des résultats intermédiaires. □

**Lemme 2.6.** *Sous les hypothèses (H4), (H5), (H6) – (H9), on a*

$$\mathbb{P} \{ (\hat{\Lambda}_D(a) = 0) \} \longrightarrow 0 \quad \text{quand} \quad n \longrightarrow \infty.$$

*Démonstration.* On a pour tout  $\varepsilon < 1$ ,

$$\mathbb{P} \{ \hat{\Lambda}_D(a) = 0 \} \leq \mathbb{P} \{ \hat{\Lambda}_D(a) \leq 1 - \varepsilon \} \leq \mathbb{P} \{ |\hat{\Lambda}_D(a) - 1| \geq \varepsilon \}.$$

Le Lemme 2.3 et le Lemme 2.4 suffisent à démontrer que

$$\hat{\Lambda}_D(a) - 1 \rightarrow 0 \quad \text{en probabilité.} \tag{2.34}$$

□

**Lemme 2.7.** *Sous les hypothèses du Théorème 2.2, on a pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , que*

$$\left( \left( \frac{n\varphi_a(x_n)^d}{\left( \frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a) \right)^2 \sigma^2(a, \vartheta_a)} \right)^{1/2} (\hat{\Lambda}_N(a, \phi) - \mathbb{E} [\hat{\Lambda}_N(a, \phi)]) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad \text{quand } n \rightarrow \infty.$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

*Démonstration.* Pour un choix judicieux de  $a \in S$ , identique au Lemme 2.2, on a

$$\begin{aligned} \hat{\Lambda}_N(a, \phi) - \mathbb{E}[\hat{\Lambda}_N(a, \phi)] &= \hat{\Lambda}_N(a, \phi) - \hat{\Lambda}_N^*(a, \phi) \\ &\quad + \hat{\Lambda}_N^*(a, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a, \phi)] \\ &\quad + \mathbb{E}[\hat{\Lambda}_N^*(a, \phi)] - \mathbb{E}[\hat{\Lambda}_N(a, \phi)], \end{aligned}$$

où  $\hat{\Lambda}_N^*(a, \phi)$  est déjà donné dans le Lemme 2.2.

Une fois de plus nous utilisons les mêmes justifications que dans le Lemme 2.2, alors nous avons

$$\left( \frac{n\varphi_a(x_n)^d}{\left(\frac{\partial}{\partial\phi}\Lambda(a, \vartheta_a)\right)^2 \sigma^2(a, \vartheta_a)} \right)^{1/2} |\hat{\Lambda}_N(a, \phi) - \hat{\Lambda}_N^*(a, \phi)| = o_p(1),$$

et

$$\left( \frac{n\varphi_a(x_n)^d}{\left(\frac{\partial}{\partial\phi}\Lambda(a, \vartheta_a)\right)^2 \sigma^2(a, \vartheta_a)} \right)^{1/2} |\mathbb{E}[\hat{\Lambda}_N^*(a, \phi)] - \mathbb{E}[\hat{\Lambda}_N(a, \phi)]| = o(1).$$

Il reste alors à démontrer la normalité asymptotique de

$$\left( \frac{n\varphi_a(x_n)^d}{\left(\frac{\partial}{\partial\phi}\Lambda(a, \vartheta_a)\right)^2 \sigma^2(a, \vartheta_a)} \right)^{1/2} |\hat{\Lambda}_N^*(a, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a, \phi)]|.$$

Pour se faire, notons que

$$\begin{aligned} \rho^*(B_i, \phi) &= \rho_a(B_i, \phi) \mathbb{I}_{|\rho(B_i, \phi)| < \theta_n}, \quad \Omega_i = \frac{1}{n\mathbb{E}[L_1]^d} (L_i \rho^*(B_i, \phi) - \mathbb{E}[L_i \rho^*(B_i, \phi)]), \\ W_{ni} &= \sqrt{n\varphi_a(h_L)^d} \Omega_i, \quad \text{et} \quad Y_n = \sum_{i=1}^n W_{ni}, \end{aligned}$$

alors,

$$Y_n = \sqrt{n\varphi_a(h_L)^d} (\hat{\Lambda}_N^*(a, \phi) - \mathbb{E}[\hat{\Lambda}_N^*(a, \phi)]).$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

En conséquence, le résultat recherché devient

$$Y_n \rightarrow \mathcal{N}(0, \sigma_1(a)), \quad (2.35)$$

où,  $\sigma_1^2(a) = \left( \frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a) \right)^2 \sigma^2(a, \vartheta_a)$ .

Pour ce faire, nous utilisons la méthodologie simple de Doob [63]. Après tout, nous supposons deux séries de nombres naturels tendant vers  $\infty$ ,  $\zeta = \zeta_n$ , et  $\eta = \eta_n$ , telles que  $\zeta = o\left(n^{1/2}\theta_n^{-1}\varphi_a(x_n)^{d/2}\right)$  et  $\eta = O(\zeta^{1-\varsigma})$  pour une certaine  $\varsigma \in (0, 1)$  et on divise  $Y_n$  comme suit

$$Y_n = Z_n + Z'_n + D_\ell, \text{ avec } Z_n = \sum_{j=1}^{\ell} U_j, \text{ et } Z'_n = \sum_{j=1}^{\ell} \xi_j,$$

où

$$U_j := \sum_{i \in I_j} W_{ni}, \quad \xi_j := \sum_{i \in J_j} W_{ni}, \quad D_\ell := \sum_{i=\ell(\zeta+\eta)+1}^n W_{ni},$$

avec

$$I_j = \{(j-1)(\zeta + \eta) + 1, \dots, (j-1)(\zeta + \eta) + \zeta\},$$

$$J_j = \{(j-1)(\zeta + \eta) + \zeta + 1, \dots, j(\zeta + \eta)\}.$$

Remarquez que, pour  $\ell = \left\lfloor \frac{n}{\zeta + \eta} \right\rfloor$ , (où  $\lfloor \cdot \rfloor$  représente la partie entière), nous avons  $\frac{\ell\eta}{n} \rightarrow 0$ ,

et  $\frac{\ell\zeta}{n} \rightarrow 1$ ,  $\frac{\eta}{n} \rightarrow 0$ , ce qui implique que  $\frac{\zeta}{n} \rightarrow 0$  lorsque  $n \rightarrow \infty$ .

Maintenant, notre résultat asymptotique est basé sur

$$\mathbb{E}(Z'_n)^2 + \mathbb{E}(D_\ell)^2 \rightarrow 0, \quad (2.36)$$

$$Z_n \rightarrow \mathcal{N}(0, \sigma_1^2(a)). \quad (2.37)$$

Pour la preuve de (2.36), la stationnarité des variables est utilisée pour obtenir

$$\mathbb{E}(Z'_n)^2 = \ell \text{Var}(\xi_1) + 2 \sum_{1 \leq i < j \leq \ell} |\text{Cov}(\xi_i, \xi_j)|, \quad (2.38)$$

$$\ell \text{Var}(\xi_1) \leq \eta \ell \text{Var}(W_{n1}) + 2\ell \sum_{1 \leq i < j \leq \eta} |\text{Cov}(W_{ni}, W_{nj})|. \quad (2.39)$$

Une première expression en (2.39) en haut à droite peut être déduite à (2.22) et du fait que

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

$\frac{\ell\eta}{n} \rightarrow 0$ . En faite,

$$\eta\ell \text{Var}(W_{n1}) = \varphi_a(h_L)^d n\ell\eta \text{Var}(\Omega_1) = O\left(\frac{\ell\eta}{n}\right) \rightarrow 0, \text{ quand } n \rightarrow \infty. \quad (2.40)$$

Alors, le second terme est le suivant

$$\ell \sum_{1 \leq i < j \leq \eta} |\text{Cov}(W_{ni}, W_{nj})| = \ell n \varphi_a(h_L)^d \sum_{1 \leq i < j \leq \eta} |\text{Cov}(\Omega_i, \Omega_j)|,$$

après cela, d'après (2.23), on démontre que

$$\sup_{x_n \leq h_L \leq y_n} \sum_{1 \leq i < j \leq \eta} |\text{Cov}(\Omega_i, \Omega_j)| = o\left(\frac{\eta}{n^2 \varphi_a(x_n)^d}\right).$$

Donc,

$$\ell \sum_{1 \leq i < j \leq \eta} |\text{Cov}(W_{ni}, W_{nj})| = o\left(\frac{\ell\eta}{n}\right) \rightarrow 0, \text{ quand } n \rightarrow \infty. \quad (2.41)$$

Nous utilisons la stationnarité pour évaluer le second terme dans le côté droit de (2.38)

$$\begin{aligned} \sum_{1 \leq i < j \leq \ell} |\text{Cov}(\xi_i, \xi_j)| &= \sum_{l=1}^{\ell-1} (\ell-l) |\text{Cov}(\xi_1, \xi_{l+1})| \\ &\leq \ell \sum_{l=1}^{\ell-1} |\text{Cov}(\xi_1, \xi_{l+1})| \\ &\leq \ell \sum_{l=1}^{\ell-1} \sum_{(i,j) \in J_1 \times J_{l+1}} \text{Cov}(W_{ni}, W_{nj}). \end{aligned}$$

Donc, pour tout  $(i, j) \in J_1 \times J_j$ , on a  $|i - j| \geq \zeta + 1 > \zeta$ , et pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , on a

$$\begin{aligned} \sum_{1 \leq i < j \leq \ell} |\text{Cov}(\xi_i, \xi_j)| &\leq \ell \frac{C\theta_n^2}{n\varphi_a(x_n)^{d+2}} \sum_{i=1}^{\zeta} \sum_{j=2\zeta+\eta+1, |i-j|>\zeta}^{\ell(\zeta+\eta)} \Delta_{i,j} \\ &\leq \frac{C\ell\zeta\theta_n^2}{n\varphi_a(x_n)^{d+2}} \Delta_{\zeta} \\ &\leq \frac{C\ell\zeta\theta_n^2}{n\varphi_a(x_n)^{d+2}} e^{-a\zeta} \rightarrow 0. \end{aligned}$$

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

Enfin, lorsque ce dernier résultat est combiné avec (2.39) et (2.41), on peut écrire

$$\mathbb{E} (Z'_1)^2 \rightarrow 0 \text{ quand } n \rightarrow \infty.$$

Quand  $(n - \ell(\zeta + \eta)) \leq \zeta$ , il en découle que

$$\begin{aligned} \mathbb{E} (D_\ell)^2 &\leq (n - \ell(\zeta + \eta)) \text{Var} (W_{n1}) + 2 \sum_{1 \leq i < j \leq n} |\text{Cov} (W_{ni}, W_{nj})| \\ &\leq \zeta \text{Var} (W_{n1}) + 2 \sum_{1 \leq i < j \leq n} |\text{Cov} (W_{ni}, W_{nj})| \\ &\leq \zeta n \varphi_a (x_n)^d \text{Var} (\Omega_1) + n \varphi_a (x_n)^d \sum_{1 \leq i < j \leq n} |\text{Cov} (\Omega_i, \Omega_j)| \\ &\leq \frac{C\zeta}{n} + o(1). \end{aligned}$$

Alors,

$$\mathbb{E} (D_\ell)^2 \rightarrow 0, \text{ quand } n \rightarrow \infty.$$

Pour la démonstration de (2.37), remarquons que

$$\left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \prod_{j=1}^{\ell} \mathbb{E} \left( e^{i\phi U_j} \right) \right| \rightarrow 0, \quad (2.42)$$

et

$$\ell \text{Var} (U_1) \rightarrow \sigma_1^2(a), \quad \ell \mathbb{E} \left( U_1^2 \mathbb{I}_{\{U_1 > \epsilon \sigma_1(a)\}} \right) \rightarrow 0. \quad (2.43)$$

Pour la preuve de (2.42), nous avons

$$\begin{aligned} &\left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \prod_{j=1}^{\ell} \mathbb{E} \left( e^{i\phi U_j} \right) \right| \\ &\leq \left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell-1} U_j} \right) \mathbb{E} \left( e^{i\phi U_\ell} \right) \right| \\ &\quad + \left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell-1} U_j} \right) - \prod_{j=1}^{\ell-1} \mathbb{E} \left( e^{i\phi U_j} \right) \right|, \end{aligned} \quad (2.44)$$

$$\left| \text{Cov} \left( e^{i\phi \sum_{j=1}^{\ell-1} U_j}, e^{i\phi U_\ell} \right) \right| + \left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell-1} U_j} \right) - \prod_{j=1}^{\ell-1} \mathbb{E} \left( e^{i\phi U_j} \right) \right|, \quad (2.45)$$

il en résulte que

$$\begin{aligned} \left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \prod_{j=1}^{\ell} \mathbb{E} \left( e^{i\phi U_j} \right) \right| &\leq \left| \text{Cov} \left( e^{i\phi \sum_{j=1}^{\ell-1} U_j}, e^{i\phi U_{\ell}} \right) \right| \\ &+ \left| \text{Cov} \left( e^{i\phi \sum_{j=1}^{\ell-2} U_j}, e^{i\phi U_{\ell-1}} \right) \right| \\ &+ \dots + \left| \text{Cov} \left( e^{i\phi U_2}, e^{i\phi U_1} \right) \right|. \end{aligned} \quad (2.46)$$

A partir de la propriété de la quasi-association, on peut écrire

$$\left| \text{Cov} \left( e^{i\phi U_2}, e^{i\phi U_1} \right) \right| \leq \frac{C\phi^2\theta_n^2}{n\varphi_{\alpha}(h_L)^{d+2}} \sum_{i \in I_1} \sum_{j \in I_2} \Delta_{i,j},$$

en appliquant cette inégalité à chaque terme du côté droit de l'équation (2.46), on obtient

$$\begin{aligned} &\left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \prod_{j=1}^{\ell} \mathbb{E} \left( e^{i\phi U_j} \right) \right| \\ &\leq \frac{C\phi^2\theta_n^2}{n\varphi_{\alpha}(h_L)^{d+2}} \left[ \sum_{i \in I_1} \sum_{j \in I_2} \Delta_{i,j} + \sum_{i \in I_1 \cup I_2} \sum_{j \in I_3} \Delta_{i,j} + \dots + \sum_{i \in I_1 \cup \dots \cup I_{\ell-1}} \sum_{j \in I_{\ell}} \Delta_{i,j} \right]. \end{aligned}$$

Remarquons que pour chaque  $\ell - 1 \geq l \geq 2$ ,  $(i, j) \in I_l \times I_{l+1}$ , on a  $\eta < \eta + 1 \leq |i - j|$ , alors

$$\sum_{i \in I_1 \cup \dots \cup I_{l-1}} \sum_{j \in I_l} \Delta_{i,j} \leq p\Delta_{\eta}.$$

Pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , l'inégalité (2.45) devient

$$\begin{aligned} \left| \mathbb{E} \left( e^{i\phi \sum_{j=1}^{\ell} U_j} \right) - \prod_{j=1}^{\ell} \mathbb{E} \left( e^{i\phi U_j} \right) \right| &\leq \frac{C\phi^2\theta_n^2}{n\varphi_{\alpha}(x_n)^{d+2}} \ell\zeta\Delta_{\eta} \\ &\leq \frac{C\phi^2\theta_n^2}{n\varphi_{\alpha}(x_n)^{d+2}} \ell\zeta e^{-a\eta} \rightarrow 0. \end{aligned}$$

Ensuite, pour (2.43), nous utilisons des arguments similaires à ceux de (2.38), pour obtenir

$$\begin{aligned} \lim_{n \rightarrow \infty} \ell \text{Var} (U_1) &= \lim_{n \rightarrow \infty} \ell\zeta \text{Var} (W_{n1}) \\ &= \lim_{n \rightarrow \infty} \ell\zeta n\varphi_{\alpha}(h_L)^d \text{Var} (\Omega_1). \end{aligned}$$

D'autre part,

$$\begin{aligned} \text{Var}(\Omega_1) &= \frac{1}{n^2 \varphi_a(h_L)^{2d}} \left\{ \mathbb{E} \left[ L^2(h_L^{-1}(a - A_i)) \rho_a^2(B_i, \phi) \right] \right. \\ &\quad \left. - \mathbb{E} \left[ L^2(\varphi_a(h_L)^{-1}(a - A_i)) \rho_a^2(B_i, \phi) \mathbb{I}_{|\rho(B_i, \phi)| > \theta_n} \right] \right\} \\ &\quad - \frac{1}{n^2} \left( \frac{1}{\varphi_a(h_L)^d} \mathbb{E} \left[ L(\varphi_a(h_L)^{-1}(a - A_i)) \rho_a(B_i, \phi) \mathbb{I}_{|\rho(B_i, \phi)| < \theta_n} \right] \right)^2. \end{aligned}$$

En utilisant les même démarche pour obtenir (2.9) et le Lemme 2.1, pour tout  $\sup_{x_n \leq h_L \leq y_n}$ , il suffit de montrer que

$$\text{Var}(\Omega_1) = \frac{\sigma_1^2(a)}{n^2 \varphi_a(x_n)^d} + o\left(\frac{1}{n^2 \varphi_a(x_n)^d}\right).$$

Ainsi,

$$\ell \text{Var}(U_1) = \frac{\ell \zeta \sigma_1^2(a)}{n} + o\left(\frac{\ell \zeta}{n}\right) \rightarrow \sigma_1^2(a).$$

Nous pouvons maintenant utiliser l'inégalité de Tchebychev avec la deuxième partie de (2.43), et en utilisant le fait que  $|U_1| \leq C\zeta |W_{n1}| \leq \frac{C\theta_n \zeta}{\sqrt{n\varphi_a(h_L)^d}}$ , pour obtenir

$$\begin{aligned} \ell \mathbb{E} \left( U_1^2 \mathbb{I}_{\{U_1 > \epsilon \sigma_1(a)\}} \right) &\leq \frac{C\theta_n^2 \zeta^2 \ell}{nh^d} \mathbb{P}(U_1 > \epsilon \sigma_1(a)) \\ &\leq \frac{C\theta_n^2 \zeta^2 \ell}{n\varphi_a(h_L)^d} \frac{\text{Var}(U_1)}{\epsilon^2 \sigma_1^2(a)} \\ &= O\left(\frac{\theta_n^2 \zeta^2}{n\varphi_a(x_n)^d}\right). \end{aligned}$$

Cela conclut la preuve du lemme. □

**Lemme 2.8.** *Sous les hypothèses (H4), (H5) et (H9) et si le paramètre de lissage  $h_L$  satisfait quand  $\sup_{x_n \leq h_L \leq y_n}$ , on a*

$$\left( \frac{n\varphi_a(x_n)^d}{\left(\frac{\partial}{\partial \phi} \Lambda(a, \vartheta_a)\right)^2 \sigma^2(a, \vartheta_a)} \right)^{1/2} \mathbb{E}[\hat{\Lambda}_N(a, \phi)] = u + o(1), \text{ quand } n \rightarrow +\infty.$$

*Démonstration.* Par des arguments analytiques simples, on écrit

$$\mathbb{E}[\hat{\Lambda}_N(a, \phi)] = \int_{\mathbb{R}^d} H(a - h_L z, \phi) L(z) dz.$$

On utilise un développement de Taylor de  $H\left(a - h_L z, \vartheta_a + u \left[n\varphi_a(x_n)^d\right]^{-1/2} \sigma(a, \vartheta_a)\right)$ , avec  $\sup_{x_n \leq h_L \leq y_n}$  pour obtenir

$$\mathbb{E}[\hat{\Lambda}_N(a, \phi)] = u \left[n\varphi_a(x_n)^d\right]^{-1/2} \sigma(a, \vartheta_a) \Lambda'(a, \vartheta(a)) + o(y_n).$$

Le résultat est alors une conséquence de l'hypothèse (H4).  $\square$

## 2.5 Étude de simulation

Dans cette section, nous évaluons le comportement de nos résultats sur des données d'échantillon simulées. Plus précisément, notre objectif principal est de montrer l'implémentation facile de l'estimateur  $k$ NN M-Regression par rapport à l'estimateur du noyau avec la méthode de crose validation (CV) donné par Ferraty [79] et d'examiner l'influence du degré de dépendance sur cette propriété asymptotique. À cette fin, nous générons des observations fonctionnelles en considérant le modèle fonctionnel non paramétrique suivant :

$$B_i = \Lambda(A_i) + \epsilon_i \text{ pour } i = 1, \dots, n,$$

où les  $\epsilon_i$  sont générés selon une distribution normale  $\mathcal{N}(0, 0.5)$ . Il est bien connu dans la littérature que le processus linéaire des variables quasi-associées satisfait la condition (H5). Ainsi, nous générons le régresseur fonctionnel quasi-associé comme suit

$$A_i(t) = \sum_{k=i+1}^{i+m} Z_k(t) \quad \text{où, } Z_k(t) = \sin(W_k * t) + V_k * t \quad t \in \left[0, \frac{\pi}{2}\right],$$

et  $(W_k)_k$  (respectivement  $(V_k)_k$ ) sont indépendants et identiquement distribués comme  $\mathcal{N}(1, 0.5)$  (respectivement  $\mathcal{N}(0, 1)$ ). Les courbes de  $A_i$  sont discrétisées dans une même grille composée de 100 points dans  $[0, \pi/2]$  et sont tracées sur la Figure 2.1 pour trois valeurs de  $m = 1$  (cas indépendant), 5 et 10.

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

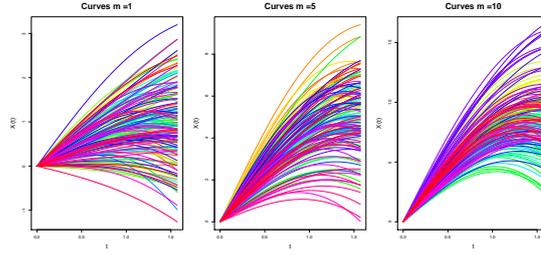


FIGURE 2.1 – Les courbes  $A_i(t), t \in [0, \pi/2]$  pour  $i = 1, \dots, 150$  et  $m = 1, 5, 10$

De plus, la variable scalaire  $B_i$  est calculée par l'opérateur de régression

$$\Lambda(a) = 5 \int_0^{\pi/2} \frac{1}{1 + |a(t)|} dt.$$

La Figure 2.1 affiche les courbes des tailles d'échantillons  $n = 150$ . Deuxièmement, nous devons sélectionner une semi-métrique appropriée  $d(\cdot, \cdot)$ , noyau  $L(\cdot)$ , paramètre de lissage  $k_{opt}$  pour l'estimateur  $k$ NN fonctionnel et  $h_{opt}$  pour le noyau CV estimateur. Pour cela, nous choisissons le noyau quadratique asymétrique défini comme  $L(u) = \frac{3}{4}(1 - u^2) 1_{[0,1]}(u)$ .

Du faite de la régularité des courbes  $A_i(t)$ , on considère la semi-métrique suivante basée sur la dérivée première

$$d^{deriv}(A_i, A_j) = \sqrt{\int_0^{\pi/2} (A'_i(t) - A'_j(t))^2 dt}, \quad \forall A_i, A_j \in \mathcal{G}.$$

Dans ce qui suit, nous avons divisé au hasard l'échantillon de 150 valeurs en deux parties : l'une est un échantillon d'apprentissage  $(A_i, B_i)_{i=1}^{100}$  qui est utilisé pour modéliser, et l'autre est un échantillon de test  $(A_j, B_j)_{j=101}^{150}$  qui est utilisé pour vérifier l'effet de prédiction.

D'une part, par l'échantillon d'apprentissage, nous pouvons sélectionner le paramètre optimal  $k_{opt}$  pour le noyau  $k$ NN et l'estimateur robuste, et le paramètre optimal  $h_{opt}$  pour le noyau classique CV et estimateur robuste par les procédures de validation croisée suivantes, respectivement.

Concrètement, on sélectionne  $k_{opt} = \arg \min_k CV_1(k)$ , où

$$CV_1(k) = \sum_{i=1}^n \left( B_i - \widehat{m}_{(-i)}^{kNN}(A) \right)^2$$

et

$$\hat{m}_{(-i)}^{kNN}(A) = \frac{\sum_{j=1, j \neq i}^n B_j L\left(\frac{d^{deriv}(A_j, A)}{E_k(A)}\right)}{\sum_{j=1, j \neq i}^n L\left(\frac{d^{deriv}(A_j, A)}{E_k(A)}\right)}.$$

Et le  $kNN$  robuste par  $k_{opt} = \arg \min_k CV_2(k)$ , où

$$CV_2(k) = \sum_{i=1}^n \left( B_i - \hat{\Lambda}_{(-i)}^{kNN}(X) \right)^2$$

et

$$\hat{\Lambda}_{(-i)}^{kNN}(A) = \arg \min_t \frac{\sum_{j=1, j \neq i}^n \rho(B_j, t) L\left(\frac{d^{deriv}(A_j, A)}{E_k(A)}\right)}{\sum_{j=1, j \neq i}^n L\left(\frac{d^{deriv}(A_j, A)}{E_k(A)}\right)}.$$

De même, nous choisissons  $h_{opt} = \arg \min_{h_L} CV(h_L)$  pour les méthodes CV classique et robuste, où

$$CV_3(h_L) = \sum_{i=1}^n \left( B_i - \hat{m}_{(-i)}^{CV}(A) \right)^2$$

et

$$CV_4(h_L) = \sum_{i=1}^n \left( B_i - \hat{\Lambda}_{(-i)}^{CV}(A) \right)^2,$$

où

$$\hat{m}_{(-i)}^{CV}(A) = \frac{\sum_{j=1, j \neq i}^n B_j L\left(\frac{d^{deriv}(A_j, A)}{h_L(A)}\right)}{\sum_{j=1, j \neq i}^n L\left(\frac{d^{deriv}(A_j, A)}{h_L(A)}\right)}$$

et

$$\hat{\Lambda}_{(-i)}^{CV}(A) = \arg \min_t \frac{\sum_{j=1, j \neq i}^n \rho(B_j, t) L\left(\frac{d^{deriv}(A_j, A)}{h_L(A)}\right)}{\sum_{j=1, j \neq i}^n L\left(\frac{d^{deriv}(A_j, A)}{h_L(A)}\right)}.$$

Un support théorique pour une telle procédure de validation croisée a été donné aussi bien pour les données dépendantes dans Hardle [104] que pour les données fonctionnelles dans Rachdi [162]. D'autre part, à parti de l'échantillon de test, nous pouvons calculer les valeurs de prédiction des variables de réponse désignées par  $(B_j)_{j=101}^{150}$ . Ainsi, les réponses prédites pour les quatre méthodes sont illustrées dans la Figure 2.2 où nous voyons que la

## CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

prévision de l'estimateur  $k$ NN est plus précise que celle du noyau CV sous l'échantillon dépendant fonctionnel quasi-associé.

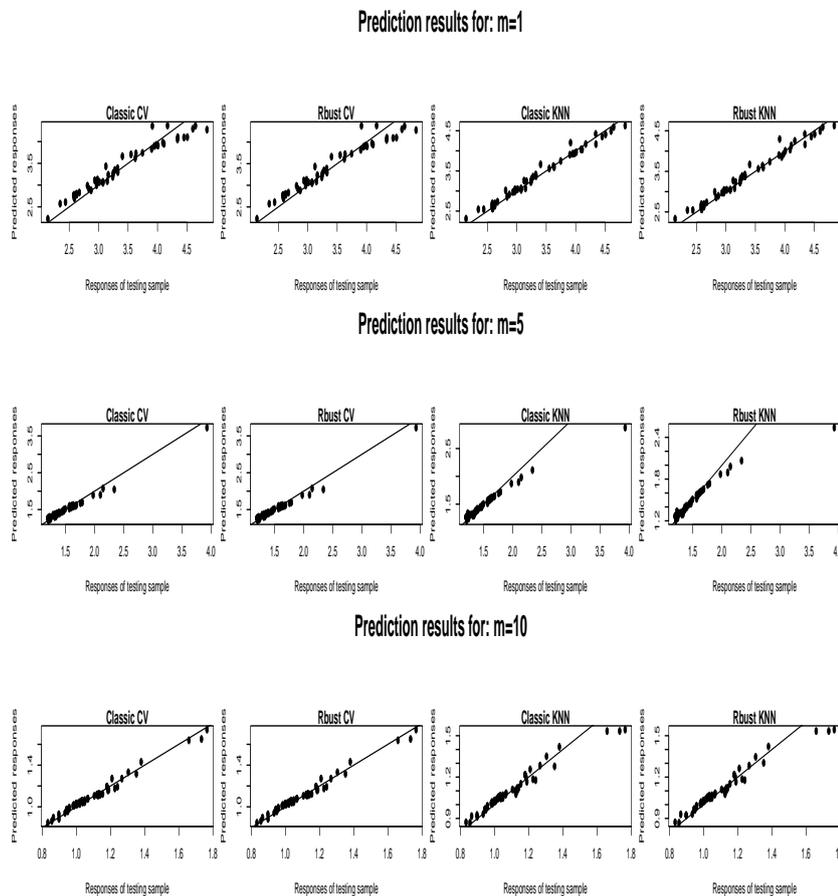


FIGURE 2.2 – Prédiction de l'estimateur classique et robuste avec les méthodes CV et  $k$ NN, respectivement.

Maintenant, nous réalisons 100 répliques indépendantes ce qui permet de calculer 100 valeurs pour  $MSE$  et d'afficher leur distribution au moyen d'un boxplot. La Figure 2.3 montre les boxplots du  $MSE$  des valeurs de prédiction.

Alors que, calculons la moyenne de  $MSE$  de l'estimateur  $k$ NN et de l'estimateur du noyau CV. Les résultats sont reportés dans le Tableau 2.1 suivant.

## CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

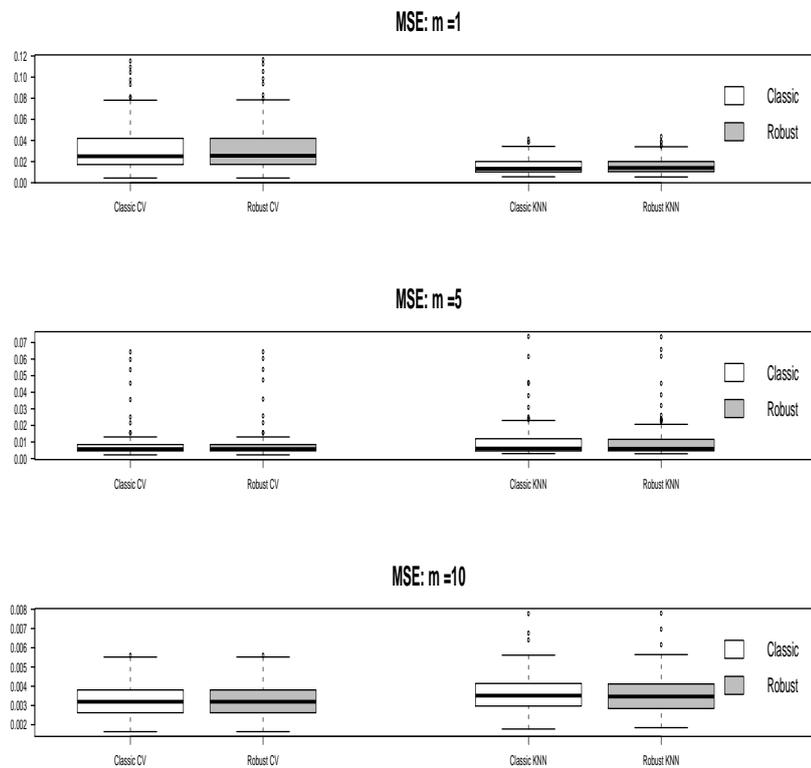


FIGURE 2.3 – Les boxplots des  $MSE$  des valeurs de prédiction par les quatre méthodes pour les différents degrés de dépendance sans valeurs aberrantes

On observe dans le Tableau 2.1 qu'en présence de valeurs aberrantes, la régression robuste  $k$ NN donne de meilleurs résultats que les autres modèles, en ce sens que, même si la valeur  $MSE$  de toutes les méthodes augmente sensiblement par rapport au nombre de les points perturbés, il restant très faible pour le  $k$ NN robuste.

CHAPITRE 2. ESTIMATION DE M-RÉGRESSION AVEC LA MÉTHODE DU K PLUS PROCHE VOISIN SOUS DES DONNÉES QUASI-ASSOCIÉES DANS LES STATISTIQUES FONCTIONNELLES

---

TABLE 2.1 – Comparaison entre les quatre méthodes en présence de valeurs aberrantes avec différents degrés de dépendance

m	Nombre de valeurs aberrantes	CV classique	CV robuste	$k$ NN classique	$k$ NN robuste
1	0	0.037514	0.037881	0.015731	0.016489
	6	502.040275	62.369308	558.648462	0.051654
	12	1772.638790	101.099854	1993.288220	0.161468
	18	3800.416704	180.703892	4157.818408	0.418262
5	0	0.008346	0.008406	0.010926	0.011399
	6	111.125504	11.549593	119.765797	0.035228
	12	400.821394	29.412965	420.740074	0.099072
	18	864.387731	61.053536	929.041219	0.261310
10	0	0.003400	0.003400	0.003780	0.003784
	6	43.848807	1.388947	46.273798	0.026864
	12	164.544865	10.976110	169.472108	0.080355
	18	360.790930	26.130611	376.862336	0.183755

## 2.6 Conclusion

L'approche de fiabilité  $k$ NN uniforme est une alternative de lissage qui permet le développement d'un estimateur adaptatif pour une variété de problèmes statistiques, y compris le choix de le paramètre de lissage.

Dans notre situation, en outre, la consistance uniforme n'est pas une simple extension de l'approche ponctuelle, car elle nécessite l'utilisation de méthodes et de techniques supplémentaires. L'hypothèse selon laquelle le paramètre de lissage dans la méthode  $k$ NN est une variable aléatoire ajoute à la complexité de ce problème. Dans le cas de résultats quasi-associés, l'innovation clé de cette approche est d'estimer la fonction de régression en mélangeant deux techniques statistiques essentielles : la méthode de M-Régression et les procédures  $k$ NN. Cette stratégie a permis le développement d'un nouvel estimateur qui combine les avantages des deux méthodes.

Pour résumer, le comportement de l'estimateur développé n'est pas affecté par le nombre de valeurs aberrantes dans la collecte de données. Par rapport à la méthode du noyau classique, le mélange de l'algorithme  $k$ NN et de la méthode robuste permet de réduire l'impact des valeurs aberrantes dans les résultats.

## CHAPITRE 3

LE TAUX DE CONVERGENCE DE LA FONCTION DE  
HASARD AVEC LA VARIABLE EXPLICATIVE  
FONCTIONNELLE : CAS DES DONNÉES SPATIALES  
AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

## **Le taux de convergence de la fonction de hasard avec la variable explicative fonctionnelle : cas des données spatiales avec la méthode des k plus proches voisins**

Dans ce chapitre, nous avons introduit un nouvel estimateur de hasard lorsque les co-variables sont de nature fonctionnelle. Cet estimateur est un mélange de la procédure  $k$  plus proches voisins ( $k$ NN) et de données fonctionnelles spatiales. Ensuite, les taux de convergence sont introduits lorsque l'échantillon considéré est collecté dans l'ordre spatial avec une structure de mélange. En théorie, il y a une estimation du point de risque puis une discussion sur les difficultés d'application, telles que le choix de le paramètre de lissage en fonction des données. De plus, une étude comparative basée sur des données simulées et réelles est également fournie pour illustrer les performances et l'utilité de l'approche  $k$ NN et pour prouver la grande sensibilité de l'approche  $k$ NN à la présence même d'une faible proportion de valeurs aberrantes dans les données.

### **3.1 Introduction**

L'opérateur de hasard est l'un des modèles les plus utilisés pour explorer la relation entre deux variables aléatoires. Dans cet article, nous traitons le problème de l'estimation non paramétrique de la fonction de hasard conditionnelle lorsque les observations sont spatiales avec l'approche  $k$ NN. L'estimation de la fonction de hasard joue un rôle très important dans les statistiques. En effet, il est utilisé en analyse de risque ou pour l'étude des phénomènes de survie dans de nombreux domaines tels que (médecine, géophysique, fiabilité, ...). La littérature sur l'estimation des fonctions de hasard est très abondante, lorsque les observations sont vectorielles. Prenons par exemple Watson et Leadbetter [190], Rousas [174], Lecoutre et Ould-Said [132], Estévez *et al.* [70], pour les références récentes. Dans tous ces travaux, les auteurs considèrent des observations indépendantes ou des données dépendantes de séries temporelles. Les premiers résultats sur l'estimation non paramétrique du hasard en statistiques fonctionnelles, ont été obtenus par Ferraty *et al.* [89]. Ils ont étudié la convergence presque complète d'un estimateur à noyau pour la fonction de hasard d'une variable aléatoire réelle conditionnelle à une variable explicative fonctionnelle. Quintela-del-Rio [159] a montré que l'estimateur à noyau présenté par Ferraty *et al.* [88] est fortement consistant et asymptotiquement normalement distribué. Une généralisation de ces résultats dans le cas des données spatiales a été obtenue par Laksaci et Mechab [126]. Plus précisément, ils ont étudié la convergence presque complète

d'une version adaptée de cet estimateur. Les mêmes auteurs ont traité le taux de convergence  $L^2$  en donnant l'expression exacte impliquée dans les termes principaux de l'erreur quadratique et la normalité asymptotique de l'estimateur construit (voir, Laksaci et Mechab [127]) , puis nous citons pour les avancées et références les plus récentes, Attouch *et al.* [7], Kara *et al.* [115], Tabti [183], Torkia *et al.* [184], Daoudi *et al.* [51], Guagui *et al.* [93].

L'importance de ce sujet de recherche est motivée par la croissance du nombre de problèmes concrets pour lesquels les données sont collectées dans un ordre spatial. De tels problèmes se rencontrent dans de nombreux domaines tels que l'épidémiologie, l'économétrie, les sciences de l'environnement et de la terre, l'agronomie, l'imagerie, etc. Les premiers résultats pris en compte pour la dépendance spatiale ont été obtenus par Tran [186]. Il a obtenu la normalité asymptotique pour l'estimateur du noyau de densité, alors que le problème de régression spatiale non paramétrique a été étudié par Lu et Chen [143] et Biau et Cadre [24], qui ont utilisé les poids de Nadaraya-Watson pour obtenir un estimateur de noyau, établissant la convergence faible et la distribution asymptotique. Le modèle d'auto-régression non paramétrique dans un contexte de prédiction sur des champs aléatoires a été étudié par Carbon *et al.* [35]. Nous renvoyons à Ferraty *et al.* [90] pour la convergence uniforme presque complète sur la composante fonctionnelle de ce modèle spatial non paramétrique.

Li et Tran [137] utilisent la technique  $k$ NN pour estimer la régression spatiale non paramétrique. Ils ont montré la normalité asymptotique de l'estimation du construit. Nous revenons à El Machkouri et Stoica [69], Robinson [170] et Dabo *et al.* [50] pour les avancées récentes et les références en analyse de données spatiales non paramétriques.

L'objectif principal de cette contribution est d'utiliser la procédure  $k$ NN. Il s'agit d'une approche alternative de lissage qui permet d'estimer l'opérateur de hasard avec le paramètre de lissage de bande variée fortement en fonction des données. Précisément, le paramètre de lissage est défini prioritairement en fonction de la distance entre la variable aléatoire fonctionnelle. L'algorithme  $k$ NN, une méthode interactive, permet d'explorer la composante topologique ainsi que la composante spectrale des données. Poussée par cette procédure sophistiquée sur la sélection de le paramètre de lissage, l'approche de lissage fonctionnel  $k$ NN a reçu beaucoup d'attention ces dernières années.

Le Livre de Gyorfi [100] est une analyse approfondie des estimateurs  $k$ NN dans le contexte de dimension finie. Les travaux dans ce domaine ont été lancés par Cover [45], et un grand nombre d'articles sont maintenant disponibles dans divers contextes d'estimation, notamment la régression, la discrimination, l'estimation de densité et de mode

### CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

et l'analyse de clustering. Puis les résultats importants dans ce sujet ont été établis par Burba *et al.* [32]. Ils ont donné le taux de convergence de la consistance presque complète de l'estimateur construit. En utilisant les mêmes techniques, Attouch et Bouabsa [10] ont établi la consistance presque complète de l'estimateur de mode conditionnel. On se réfère à Attouch et Belabed [7] pour la fonction de hasard conditionnelle, on fait référence à Collomb [42], Devroye et Wangner [60], Li [134], Moore *et al.* [153], Devroye et Wangner [59], Devroye *et al.* [58], Beirlant *et al.* [22], Laloë [129], Burba *et al.* [33], Tran [185], Lian [133], Attouch *et al.* [8], Kudraszow et Vieu [119] et nous citons pour les avancées et références les plus récentes Kara *et al.* [115], Almanjahie *et al.* [3], Bouabsa [30]. Cependant, la difficulté du lissage  $k$ NN réside dans le fait que le paramètre de lissage est une variable aléatoire, contrairement à la régression classique dans laquelle le paramètre de lissage est un scalaire déterministe. Ainsi, l'étude des propriétés asymptotiques de notre estimateur proposé est compliquée et nécessite des outils et des techniques supplémentaires.

Ainsi, le but de cet article nous montre que cela peut être utilisé pour approfondir l'estimation du hasard non paramétrique fonctionnel dans le cas d'ensembles de données spatiales. Cela est motivé par le fait que l'estimateur de hasard robuste présente plusieurs avantages par rapport à l'estimateur de régression par noyau classique. Le principal avantage de l'utilisation d'une régression robuste est qu'elle permet de réduire l'effet des données aberrantes. L'exploration et l'analyse de données en présence de valeurs aberrantes est un grand défi en statistique. En particulier, en statistiques spatiales plusieurs modèles robustes résistants à ces anomalies ont été étudiés. Concernant le cas de la modélisation non paramétrique, la première approche est donnée par Xu et Wang [191]. Ces derniers considèrent une estimation linéaire locale de la fonction de régression basée sur le plus petit écart absolu. Hallin *et al.* [102] établir la consistance et la normalité asymptotique d'une version spatiale de l'estimation linéaire locale des quantiles conditionnels. La version spatiale de la M-estimation de la fonction de régression a été introduite par Gheriballah *et al.* [95]. Ils ont obtenu la convergence presque complète et la normalité asymptotique de cette estimation. Dabo et Thiam [48] se sont attachés à étudier la régression quantile non paramétrique par la méthode  $L^1$ . Ils ont affirmé la faible consistance et la normalité asymptotique de l'estimateur construit.

Dans ce chapitre, sous certaines hypothèses générales, nous énonçons la convergence presque complète (avec taux). Notant que, ce travail est le lien entre l'étude de Laksaci et Mechab [126] avec Kara *et al.* [115] et Almanjahie *et al.* [3]. Par exemple, rappelons que la complexité de notre recherche vient du fait que le paramètre de lissage dans la

méthode  $k$ NN est une variable aléatoire. Précisément, le paramètre de lissage est défini prioritairement en fonction de la distance entre la variable aléatoire fonctionnelle. Une telle considération permet d'explorer la composante topologique ainsi que la composante spectrale des données.

Dans NFDA, le  $k$ NN hasard avec des données spatiales est nouveau. Nous présentons l'estimateur de notre modèle spatial avec la méthode  $k$ NN dans la deuxième Section. Dans la troisième Section, nous donnons les hypothèses, puis nous étudions la convergence presque complète de cet estimateur. Dans la quatrième Section, nous donnons tous les résultats et leurs preuves. Comme application nous traitons, dans la Sous-Section (2.5.1), l'estimation du risque maximum conditionnellement à une variable explicative fonctionnelle. Dans la Sous-Section (2.5.2), nous soulignons l'influence pratique directe des résultats de l'UIB sur le choix de le paramètre de lissage basée sur les données. Enfin, une étude de simulation est donnée dans la Sous-Section (2.5.4) prouvant les bonnes performances de notre estimateur puis la Sous-Section (2.5.5) est consacrée à une application de données réelles pour évaluer les performances de cette estimation. Nous terminons le chapitre par une conclusion.

## 3.2 Le modèle et son estimateur avec la méthode $k$ NN

Soit  $M$  un entier naturel dans  $\mathbb{N}^*$ . On considère le champ aléatoire  $W_{\mathbf{i}} = (A_{\mathbf{i}}, B_{\mathbf{i}})$ ,  $\mathbf{i} \in \mathbb{N}^M$  avec des valeurs dans  $\mathcal{G} \times \mathbb{R}$ , où  $(\mathcal{G}, d)$  est un espace semi-métrique de dimension éventuellement infinie. Dans ce contexte,  $(A_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^M}$  peut être une variable aléatoire fonctionnelle. Il est à noter que, depuis une bonne dizaine d'années, la communauté statistique s'est préoccupée de développer des modèles et des méthodes adaptés à ce contexte de données fonctionnelles. Alors que les premières études dans ce sens se sont principalement concentrées sur des modèles linéaires (voir Bosq [29], Ramsay et Silverman [167]), les développements récents (voir Ferraty et Vieu [79]) rapportent des modèles non paramétriques adaptés pour ce type de données.

Par la suite, on fixe un point  $a$  dans  $\mathcal{G}$  (respectivement, un compact  $\mathcal{Q} \in \mathbb{R}$ ), on suppose que les observations spatiales  $(A_{\mathbf{i}}, B_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^M}$  ont la même distribution que  $W = (A, B)$  et que la version régulière de la probabilité conditionnelle de  $B$  sachant  $A = a$  existe et admet une densité bornée par rapport à la mesure de Lebesgue sur  $\mathbb{R}$ , notée  $f_k^a$ . Avec la méthode  $k$ NN le paramètre fonctionnel étudié dans cet article, noté  $h_k^a$ , est défini, pour

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

tout  $b \in \mathbb{R}$  tel que  $F^a(b) < 1$ , par

$$h_k^a(b) = \frac{f_k^a(b)}{1 - F_k^a(b)},$$

où,  $F_k^a$  est la fonction de distribution conditionnelle de  $B$  sachant  $A = a$ , avec

$$F_k^a(\cdot) = \mathbb{P}(B \leq \cdot \mid A = a).$$

De plus, nous supposons que le champ aléatoire fonctionnel est observé sur l'ensemble  $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, \dots, i_M) \in \mathbb{N}^M, 1 \leq i_\ell \leq n_\ell, \ell = 1, \dots, M\}$ ,  $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M$  et nous estimons avec la méthode  $k$ NN la fonction de distribution conditionnelle par

$$\widehat{F}_k^a(b) = \frac{\sum_{\mathbf{i} \in I_{\mathbf{n}}} L(P_k^{-1}d(a, A_{\mathbf{i}})) J(h_J^{-1}(b - B_{\mathbf{i}}))}{\sum_{\mathbf{i} \in I_{\mathbf{n}}} L(P_k^{-1}d(a, A_{\mathbf{i}}))}, \quad \forall b \in \mathbb{R},$$

avec,  $L$  est un noyau et  $J$  est une fonction de distribution conditionnelle (cdf), définie par

$$J(\cdot) = \frac{\int_0^\cdot L(s)ds}{\int_0^{1/2} L(s)ds},$$

et  $P_k = P_{k, \mathbf{n}}$  (resp.  $h_J = h_{J, \mathbf{n}}$ ) est une suite de nombres réels positifs qui appartiennent à l'intervalle  $(x_{\mathbf{n}}, y_{\mathbf{n}})$  ( resp.  $(v_{\mathbf{n}}, s_{\mathbf{n}})$ ), avec  $\lim_{\mathbf{n} \rightarrow \infty} v_{\mathbf{n}} = \lim_{\mathbf{n} \rightarrow \infty} s_{\mathbf{n}} = 0$ .

Bien que, récemment, ce type d'estimation par noyau soit largement étudié dans la littérature, ces études concernent exclusivement des largeurs de bande fixes. Dans notre article, nous dérivons la consistance uniforme presque complète avec les vitesses de convergence correspondantes et mettons l'accent sur les applications pour la construction de paramètres de lissage automatiques basés sur les données dans la section.

De  $\widehat{F}_k^a$ , on déduit un estimateur de la densité conditionnelle, noté  $\widehat{f}_k^a$ , où

$$f_k^a(\cdot) = (F_k^a)'(\cdot).$$

Donc

$$\widehat{f}_k^a(b) = \frac{h_J^{-1} \sum_{\mathbf{i} \in I_{\mathbf{n}}} L(P_k^{-1}d(a, A_{\mathbf{i}})) J'(h_J^{-1}(b - B_{\mathbf{i}}))}{\sum_{\mathbf{i} \in I_{\mathbf{n}}} L(P_k^{-1}d(a, A_{\mathbf{i}}))}, \quad \forall b \in \mathbb{R},$$

où,  $J'$  est la dérivée de  $J$ . L'estimateur  $k$ NN de la fonction de hasard conditionnelle est

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

noté par

$$\widehat{h}_k^a(b) = \frac{\widehat{f}_k^a(b)}{1 - \widehat{F}_k^a(b)}, \quad \forall b \in \mathbb{R}.$$

Le but de ce travail est d'étudier la convergence presque complète de l'estimateur  $\widehat{h}_k^a$  vers  $h_k^a$ , lorsque le champ aléatoire fonctionnel  $(W_i)_{i \in \mathbb{N}^M}$  satisfait la condition de mélange suivante.

$$\left\{ \begin{array}{l} \text{Il y a une fonction } \Phi(s) \downarrow 0 \text{ quand } s \rightarrow \infty, \text{ tel que} \\ \forall Z, Z' \text{ sous-ensemble de } \mathbb{N}^M \text{ cardinal fini} \\ \alpha(\mathcal{B}(Z), \mathcal{B}(Z')) = \sup_{Y \in \mathcal{B}(Z), E \in \mathcal{B}(Z')} |\mathbb{P}(Y \cap E) - \mathbb{P}(Y)\mathbb{P}(E)| \\ \leq \Psi(\text{Card}(Z), \text{Card}(Z')) \Phi(\text{dist}(Z, Z')), \end{array} \right.$$

où,  $\mathcal{B}(Z)$  (resp.  $\mathcal{B}(Z')$ ) la tribu borélienne engendrée par  $(W_i, i \in Z)$  (resp.  $(W_i, i \in Z')$ ),  $\text{Card}(Z)$  (resp.  $\text{Card}(Z')$ ) est le cardinal de  $Z$  (resp.  $Z'$ ),  $\text{dist}(Z, Z')$  désigne la distance Euclidienne entre  $Z$  et  $Z'$  et  $\Psi$  une fonction symétrique :  $\mathbb{N}^2 \rightarrow \mathbb{R}^+$ , décroissante par rapport aux deux variables séparément et remplissant l'une des conditions suivantes :

$$\Psi(u, v) \leq C \min(u, v), \quad u, v \in \mathbb{N}, \quad (3.1)$$

ou

$$\Psi(u, v) \leq C(u + v + 1)^{\widetilde{\vartheta}}, \quad u, v \in \mathbb{N}, \quad (3.2)$$

pour certains  $\widetilde{\vartheta} \geq 1$  et  $C > 0$ , notez que ces conditions ont été utilisées par Tran [186] et qu'elles sont vérifiées par de nombreux modèles spatiaux (voir Guyon [98]).

Rappelons que lorsque l'équation (3.1) est vérifiée avec  $\Psi \equiv 1$  ou  $M = 1$ , le champ aléatoire  $W_i = (A_i, B_i)$  on dit que est fortement mélangeant (voir Doukhan [65] pour plus d'informations sur les qualités de mélange et des exemples).

De plus, nous supposons que le processus  $W$  vérifie la condition de mélange suivante :

$$\sum_{i=1}^{\infty} i^{\ell} \Phi(i) < \infty, \ell > 0. \quad (3.3)$$

Notons que les conditions (3.2) et (3.3) sont identiques aux conditions de mélange employées par Carbon *et al.* [36] et Tran [186].

### 3.3 Propriétés asymptotiques

Dans la suite, nous désignerons par  $C$  et/ou  $C'$  toute constante strictement positive et nous supposons que :  $\forall b \in \mathcal{S}, F_k^a(b) < 1$ . Rappelons que dans ce contexte spatial,  $\mathbf{n} \rightarrow \infty$  signifie que  $\min \{\mathbf{n}_\ell\} \rightarrow \infty$  et que pour chaque  $1 \leq j, \ell \leq M$  nous avons  $\infty > C > \left| \frac{\mathbf{n}_j}{\mathbf{n}_\ell} \right|$ .

Introduisons les hypothèses suivantes :

- (K1)
  - (K1-i)  $\mathbb{P}(A \in B(a, t)) = \varphi_a(t) > 0$  où  $B(a, t)$  est la boule fermée, centrée en  $a$  et de rayon  $t$ .
  - (K1-ii) Pour tous  $q \in (0, 1)$ ,  $\lim_{t \rightarrow 0} \frac{\varphi_a(qt)}{\varphi_a(t)} = \varsigma_a(q) < \infty$ .
- (K2) La fonction  $\Phi$  vérifie  $\sum_{i=1}^{\infty} i^\gamma \Phi(i) < \infty, \gamma > 5M$ .
- (K3)  $0 < \sup_{i \neq j} \mathbb{P}[(A_i, A_j) \in B(a, P_k) \times B(a, P_k)] \leq C (\varphi_a(P_k))^{(x+1)/x}$ , avec  $1 < x < \gamma M^{-1}$ .
- (K4) Rappelons que  $a$  est un élément fonctionnel fixe et  $N_a$  est un voisinage fixe de  $a$ . Le modèle non paramétrique pour la distribution conditionnelle et la densité conditionnelle est construit en supposant que  $C > 0$  tel que pour tout  $(b_1, b_2) \in \mathbb{R}^2$  et tout  $(a_1, a_2) \in N_a \times N_a$ , nous avons :
  - (K4-i)

$$|F_k^{a_1}(b_1) - F_k^{a_2}(b_2)| \leq C \left( d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4} \right), \beta_3 > 0, \beta_4 > 0.$$

- (K4-ii)

$$|f_k^{a_1}(b_1) - f_k^{a_2}(b_2)| \leq C \left( d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4} \right).$$

- (K5) Le noyau  $L$  a une première dérivée continue sur  $(0, 1/2)$  et est supporté à l'intérieur de  $(0, 1/2)$ . De plus, il existe des constantes  $0 < C < C' < \infty$ , telles que

$$0 < C \mathbb{I}_{(0,1/2)}(\cdot) \leq L(\cdot) \leq C' \mathbb{I}_{(0,1/2)}(\cdot) \quad \text{et} \quad L(1/2) - \int_0^{1/2} L'(s) \varsigma_a(s) ds > 0.$$

- (K6)  $J$  est une fonction de la classe  $C^2$  et un support compact.

— (K7) Il existe  $0 < \alpha < (\delta - 5M)/3M$  et  $\zeta_0 > 0$ , tels que

$$\lim_{\mathbf{n} \rightarrow \infty} \widehat{\mathbf{n}}^\alpha h_J = \infty \quad \text{et} \quad C \widehat{\mathbf{n}}^{\frac{(5+3\alpha)M - \delta}{\delta}} + \zeta_0 \leq h_J \varphi_a(P_k),$$

où,  $\widehat{\mathbf{n}} = n_1 \dots n_M$ .

— (K8)

— (K8-i)

$$\mathcal{L}_0 = \left\{ \cdot \mapsto \frac{1}{\int_0^1 L(s) ds} \int_0^{(\delta^{-1}(y \cdot))} L(s) ds, \text{ pour } \delta > 0 \right\} \text{ est une classe mesurable ponctuellement.}$$

— (K8-ii)

$$\mathcal{L}_0 \text{ est telle que } \sup_U \int_0^1 \sqrt{1 + \log \mathcal{N}(\varepsilon \|K_0\|_{U,2}, \mathcal{L}_0, d_U)} d\varepsilon < \infty,$$

où,  $K_0$  est la fonction enveloppe de l'ensemble  $\mathcal{L}_0$ .

— (K9)

— (K9-i)

$$\frac{\log \mathbf{n}}{\mathbf{n} \min(x_{\mathbf{n}}, \varphi_a(x_{\mathbf{n}}))} \rightarrow 0.$$

— (K9-ii)

$$\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \min(x_{\mathbf{n}}, \varphi_a(x_{\mathbf{n}}))} \rightarrow 0.$$

### Remarque sur les hypothèses

Notre recherche est le lien entre le travail de Kara *et al.* [115], et Laksaci et Mechab [126], donc plusieurs hypothèses sont les mêmes considérées dans toutes ces recherches.

## 3.4 Résultat et preuve

**Théorème 3.1.** *Sous les conditions (3.1) et (3.2) et les hypothèses (K1)-(K7), et (K9), on a que*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |\widehat{h}_k^a(b) - h_k^a(b)| = O(y_{\mathbf{n}}^{\beta_3}) + O(s_{\mathbf{n}}^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}} \right).$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

*Démonstration.* La démonstration de ce théorème repose sur le fait que

$$|\hat{h}_k^a(b) - h_k^a(b)| \leq \frac{1}{|1 - \hat{F}_k^a(b)|} \left[ |\hat{f}_k^a(b) - f^a(b)| + \frac{|f_k^a(b)|}{|1 - F_k^a(b)|} |\hat{F}_k^a(b) - F_k^a(b)| \right].$$

Le Théorème 3.1 est obtenu à partir des Théorèmes 3.2 - 3.3 et du Corollaire 3.1.1 ci-dessous.  $\square$

**Corollaire 3.1.1.** *Sous les conditions du Théorème 3.1, il existe un  $\delta > 0$  tel que :*

$$\sum_{\mathbf{n} \in \mathbb{N}^M} \mathbb{P} \left\{ \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |1 - \hat{F}_k^a(b)| < \delta \right\} < \infty.$$

*Démonstration.* Nous rappelons que pour tout  $P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})$ ,  $h_J \in (\nu_{\mathbf{n}}, s_{\mathbf{n}})$ , nous avons

$$|1 - \hat{F}_k^a(b)| \leq \frac{(1 - F_k^a(b))}{2} \text{ implique } |\hat{F}_k^a(b) - F_k^a(b)| \geq \frac{(1 - F_k^a(b))}{2}.$$

Par conséquent,

$$\begin{aligned} & \sum_{\mathbf{n} \in \mathbb{N}^M} \mathbb{P} \left( \inf_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}}), h_J \in (\nu_{\mathbf{n}}, s_{\mathbf{n}})} |1 - \hat{F}_k^a(b)| \leq \frac{1 - F_k^a(b)}{2} \right), \\ & \leq \sum_{\mathbf{n} \in \mathbb{N}^M} \mathbb{P} \left( \sup_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}}), h_J \in (\nu_{\mathbf{n}}, s_{\mathbf{n}})} |\hat{F}_k^a(b) - F_k^a(b)| \geq \frac{1 - F_k^a(b)}{2} \right) < \infty. \end{aligned}$$

L'application du Théorème 3.2 nous permet d'obtenir le résultat souhaité.  $\square$

**Théorème 3.2.** *Sous les hypothèses (K1), et (K7)-(K8-ii)-(K9-i), on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |\hat{F}_k^a(b) - F_k^a(b)| = O(y_{\mathbf{n}}^{\beta_3}) + O(s_{\mathbf{n}}^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \varphi_a(x_{\mathbf{n}})}} \right).$$

*Démonstration.* Les démonstrations sont basées, respectivement, sur les décompositions suivantes :

$$\begin{aligned} \hat{F}_k^a(b) - F_k^a(b) &= \frac{1}{\hat{F}_D^a} \left\{ (\hat{F}_N^a(b) - \mathbb{E} \hat{F}_N^a(b)) - (F_k^a(b) - \mathbb{E} \hat{F}_N^a(b)) \right\} + \\ &+ \frac{F_k^a(b)}{\hat{F}_D^a} (\mathbb{E} \hat{F}_D^a - \hat{F}_D^a). \end{aligned} \tag{3.4}$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

On définit alors,

$$\widehat{F}_D^a = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}[L(P_k^{-1}d(a, A_1))]} \sum_{i \in I_n} L(P_k^{-1}d(a, A_i)),$$

$$\widehat{F}_N^a(b) = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}[L(P_k^{-1}d(a, A_1))]} \sum_{i \in I_n} L(P_k^{-1}d(a, A_i)) J(h_j^{-1}(b - B_i)),$$

où, 1 est l'indice spatial des composantes fixes 1.

La preuve découle du Lemmes ci-dessous.  $\square$

**Lemme 3.1.** *Sous les hypothèses (3.1) et (3.2), ainsi que les hypothèses (K1), (K3), (K5) et (K7)-(K9-i), on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} |\widehat{F}_D^a - \mathbb{E}\widehat{F}_D^a| = O\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\varphi(x_n)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

*Démonstration.* Dans les preuves suivantes on notera pour tout  $i \in I_n$ ,

$$L_i(a) = L(P_k^{-1}d(a, A_i)), \quad J_i(b) = J(h_j^{-1}(b - B_i)),$$

et

$$J'_i(y) = J'(h_j^{-1}(b - B_i)).$$

Il faut montrer qu'il existe,  $\zeta_0 > 0$ , tel que

$$\sum_n \mathbb{P}\left\{ \sup_{x_n \leq P_k \leq y_0} \sqrt{\frac{n\varphi_a(x_n)}{\log n}} |\widehat{F}_D^a - \mathbb{E}[\widehat{F}_D^a]| \geq \zeta_0 \right\} < \infty, \text{ pour certains } y_0 > 0.$$

Suivant l'inégalité de Bernstein (voir, Dony et Einmahl [62], p. 321), la preuve est basée sur cette inégalité pour les processus empiriques, en définissant

$$h_{L,j} = 2^j x_n \text{ et } K(n) = \max\{j : h_{L,j} \leq 2y_0\},$$

ainsi

$$\begin{aligned} & \sup_{x_n \leq P_k \leq y_0} \sqrt{\frac{n\varphi_a(x_n)}{\log n}} |\widehat{F}_D^a - \mathbb{E}[\widehat{F}_D^a]| \\ & \leq \max_{1 \leq j \leq K(n)} \sup_{h_{L,j-1} \leq P_k \leq h_{L,j}} \sqrt{\frac{\mathbf{n}\varphi_a(P_k)}{\log n}} |\widehat{F}_D^a - \mathbb{E}[F_D^a]|. \end{aligned}$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

Pour écrire la différence, la démonstration s'appuie sur des concepts similaires à ceux utilisés par Carbon *et al.* [36]. Ainsi

$$\hat{F}_D^a(a) - \mathbb{E}[\hat{F}_D^a(a)] = \frac{1}{\hat{\mathbf{n}}\mathbb{E}[L_1(a)]} \sum_{i \in I_{\mathbf{n}}} \Theta_i(a),$$

où  $\Theta_i(a) = (1/\sqrt{\mathbf{n}}) \sum_{i=1}^{\mathbf{n}} (L_i - \mathbb{E}[L_i])$  correspond au processus empirique basé sur des variables  $A_1, A_2, \dots, A_{\mathbf{n}}$ . Ensuite, considérons la classe de fonctions suivante

$$\mathcal{L}_{L,j} = \{z \mapsto L(\gamma^{-1}d(a, z)) \text{ où } h_{L,j-1} \leq \gamma \leq P_{k,j}\}.$$

Donc,

$$\begin{aligned} \sup_{x_{\mathbf{n}} \leq P_k \leq y_0} \sqrt{\frac{n\varphi_a(x_{\mathbf{n}})}{\log n}} |\hat{F}_D^a - \mathbb{E}[\hat{F}_D^a]| &\leq \sum_{i=1}^{K(\mathbf{n})} \mathbb{P} \left\{ \frac{1}{\sqrt{n\varphi_a\left(\frac{P_{k,j}}{2}\right) \log n}} \|\sqrt{\mathbf{n}}\Theta_i(a)\|_{\mathcal{L}_{L,j}} \geq \zeta_0 \right\} \\ &\leq K(\mathbf{n}) \max_{j=1, \dots, k(\mathbf{n})} \mathbb{P} \left\{ \max_{1 \leq k \leq \mathbf{n}} \|\sqrt{k}\Theta_k(a)\|_{\mathcal{L}_{L,j}} \geq \zeta_0 \sqrt{n\varphi_a\left(\frac{P_{k,j}}{2}\right) \log n} \right\}. \end{aligned} \quad (3.5)$$

On considère la décomposition spatiale de Tran [186] sur les variables  $\Theta_i(a)$ , définies, pour l'entier fixe  $\varrho_{\mathbf{n}}$ , comme suit

$$\begin{aligned} V(1, \mathbf{n}, a, \mathbf{j}) &= \sum_{\substack{i_{\ell}=2j_{\ell}\varrho_{\mathbf{n}}+1 \\ \kappa=1, \dots, M}}^{2j_{\kappa}\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}} \Theta_i(a), \\ V(2, \mathbf{n}, a, \mathbf{j}) &= \sum_{\substack{i_{\kappa}=2j_{\kappa}\varrho_{\mathbf{n}}+1 \\ \kappa=1, \dots, M-1}}^{2j_{\kappa}\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}} \sum_{i_M=2j_M\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}+1}^{2(j_M+1)\varrho_{\mathbf{n}}} \Theta_i(a), \\ V(3, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k=2j_k\varrho_{\mathbf{n}}+1 \\ k=1, \dots, N-2}}^{2j_k\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}} \sum_{i_{N-1}=2j_{N-1}\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}+1}^{2(j_{N-1}+1)\varrho_{\mathbf{n}}} \sum_{i_N=2j_N\varrho_{\mathbf{n}}+1}^{2j_N\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}} \Theta_i(a), \\ V(4, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k=2j_k\varrho_{\mathbf{n}}+1 \\ k=1, \dots, N-2}}^{2(j_{N-1}+1)\varrho_{\mathbf{n}}} \sum_{i_{N-1}=2j_{N-1}\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}+1}^{2(j_N+1)\varrho_{\mathbf{n}}} \sum_{i_N=2j_N\varrho_{\mathbf{n}}+\varrho_{\mathbf{n}}+1} \Theta_i(a), \end{aligned}$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

et ainsi de suite. Les deux derniers termes sont les suivants

$$V(2^{M-1}, \mathbf{n}, a, \mathbf{j}) = \sum_{\substack{i_\kappa=2j_\kappa \varrho_{\mathbf{n}} + \varrho_{\mathbf{n}} + 1 \\ \kappa=1, \dots, M-1}}^{2(j_{\kappa+1})\varrho_{\mathbf{n}}} \sum_{i_M=2j_M \varrho_{\mathbf{n}} + 1}^{2j_M \varrho_{\mathbf{n}} + \varrho_{\mathbf{n}}} \Theta_{\mathbf{i}}(a),$$

$$V(2^M, \mathbf{n}, a, \mathbf{j}) = \sum_{\substack{i_\kappa=2(j_{\kappa+1})\varrho_{\mathbf{n}} \\ \kappa=1, \dots, M}}^{2(j_{\kappa+1})\varrho_{\mathbf{n}}} \Theta_{\mathbf{i}}(a).$$

Pour  $\tau_i = 2^{-1}n_i \varrho_{\mathbf{n}}^{-1}$ ,  $i = 1, \dots, M$  et  $\mathcal{H} = \{0, \dots, \tau_1 - 1\} \times \dots \times \{0, \dots, \tau_M - 1\}$ , nous posons

$$\Delta(\mathbf{n}, a, i) = \sum_{\mathbf{j} \in \mathcal{H}} V(i, \mathbf{n}, a, \mathbf{j}).$$

Sans perte de généralité, on peut écrire

$$|\widehat{F}_D^a(a) - \mathbb{E}[\widehat{F}_D^a(a)]| = \frac{1}{\widehat{\mathbf{n}} \mathbb{E}[L_1(a)]} \sum_{i=1}^{2^M} \Delta(\mathbf{n}, a, i). \quad (3.6)$$

Car même si  $n_i$  n'est pas exactement égal à  $2\tau_i \varrho_{\mathbf{n}}$ , on regroupe les variables restantes dans un bloc  $\Delta(\mathbf{n}, a, 2^M + 1)$  (cela ne changera pas la preuve voir Biau et Cadre [24]).

Maintenant, sous la dernière équation (3.6), pour tout  $\zeta > 0$ , on a

$$\mathbb{P}(|\widehat{F}_D^a(a) - \mathbb{E}[\widehat{F}_D^a(a)]| \geq \zeta) \leq 2^M \max_{i=1, \dots, 2^M} \mathbb{P}(\Delta(\mathbf{n}, a, i) \geq \zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]).$$

Il suffit donc de calculer

$$\mathbb{P}(\Delta(\mathbf{n}, a, i) \geq \zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]) \quad \text{pour tout } i = 1, \dots, 2^M.$$

Nous ne traitons que le cas  $i = 1$ . Pour cela, on numérote les variables ( $V(1, \mathbf{n}, a, \mathbf{j}); \mathbf{j} \in \mathcal{H}$ ) et on applique le Lemme 4.5 de Carbon *et al.* [36] sur les variables renumérotées.

Les variables avec la nouvelle numérotation seront notées  $D_1, \dots, D_N$ , où  $N = \prod_{\kappa=1}^M \tau_\kappa = 2^{-M} \widehat{\mathbf{n}} \varrho_{\mathbf{n}}^{-M} \leq \widehat{\mathbf{n}} \varrho_{\mathbf{n}}^{-M}$ . Remarquez que pour tout  $D_j$  il existe un certain  $\mathbf{j}$  dans  $\mathcal{H}$ , tel que

$$D_j = \sum_{\mathbf{i} \in I(1, \mathbf{n}, a, \mathbf{j})} \Theta_{\mathbf{i}}(a),$$

où,  $I(1, \mathbf{n}, a, \mathbf{j}) = \{\mathbf{i} : 2j_\kappa \varrho_{\mathbf{n}} + 1 \leq i_\kappa \leq 2j_\kappa \varrho_{\mathbf{n}} + \varrho_{\mathbf{n}}; \kappa = 1, \dots, M\}$ . La distance entre ces ensembles est supérieure à  $\varrho_{\mathbf{n}}^M$  et chaque ensemble contient des éléments  $\varrho_{\mathbf{n}}^M$ .

Le Lemme (4.5) de Carbon *et al.* [36], nous permet d'approximer  $D_1, D_2, \dots, D_N$  par des

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
 AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
 SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

variables aléatoires indépendantes  $D_1^*, \dots, D_N^*$  de même loi que  $D_{j=1, \dots, N}$  et telle que

$$\sum_{j=1}^N \mathbb{E} |D_j - D_j^*| \leq 2CN (\varrho_n^M \psi(N-1) \varrho_n^M, \varrho_n^M) \varphi(\varrho_n).$$

Alors, par les inégalités de Bernstein et de Markov, on a

$$\mathbb{P}(\Delta(\mathbf{n}, a, i) \geq \zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]) \leq \beta_1 + \beta_2,$$

où,

$$\begin{aligned} \beta_1 &= \mathbb{P}\left(\left|\sum_{j=1}^N D_j^*\right| \geq \frac{N\zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]}{2N}\right) \\ &\leq 2 \exp\left(-\frac{(\zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)])^2}{N \text{Var}[D_1^*] + C \varrho_n^M \zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]}\right), \\ \beta_2 &= \mathbb{P}\left(\sum_{j=1}^N |D_j - D_j^*| \geq \frac{\zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]}{2}\right) \leq \frac{C}{\zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)]} \sum_{j=1}^N \mathbb{E} |D_j - D_j^*|, \\ &\leq CN \varrho_n^M (\zeta \widehat{\mathbf{n}} \mathbb{E}[L_1(a)])^{-1} \psi((N-1) \varrho_n^M, \varrho_n^M) \varphi(\varrho_n). \end{aligned}$$

Comme  $\widehat{\mathbf{n}} = 2^M N \varrho_n^M$  et  $\psi((N-1) \varrho_n^M, \varrho_n^M) \leq \varrho_n^M$  donc pour  $\zeta = \zeta_0 \sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}} \varphi_a(P_k)}}$ ,

$$\beta_2 \leq \widehat{\mathbf{n}} \varrho_n^M (\log \widehat{\mathbf{n}})^{-1/2} (\widehat{\mathbf{n}} \varphi_a(P_k))^{-1/2} \varphi(\varrho_n).$$

Nous prenons  $\varrho_n C \left(\frac{\widehat{\mathbf{n}} \varphi_a(P_k)}{\log \widehat{\mathbf{n}}}\right)^{1/2M}$ , alors

$$\beta_2 \leq (\log \widehat{\mathbf{n}})^{-1} \widehat{\mathbf{n}} \varphi(\varrho_n). \quad (3.7)$$

D'après l'hypothèse (K7), il en résulte que

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \widehat{\mathbf{n}} \varphi(\varrho_n) < \infty.$$

Nous allons maintenant nous occuper de  $\beta_1$ . Pour ce faire, il suffit d'évaluer  $\text{Var}[D_1^*]$ . En effet,

$$\text{Var}[D_1^*] = \text{Var}\left[\sum_{\mathbf{i} \in I(1, \mathbf{n}, a, 1)} \Theta_{\mathbf{i}}(a)\right] = \sum_{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, a, 1)} |\text{Cov}(\Theta_{\mathbf{i}}(a), \Theta_{\mathbf{j}}(a))|.$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

Posons  $Z_n = \sum_{\mathbf{i} \in I(1, \mathbf{n}, a, \mathbf{1})} \text{Var} [\Theta_{\mathbf{i}}(a)]$  et  $T_n = \sum_{\mathbf{i} \neq \mathbf{j} \in I(1, \mathbf{n}, a, \mathbf{1})} |\text{Cov} (\Theta_{\mathbf{i}}(a), \Theta_{\mathbf{j}}(a))|$ . En vertu de (K1), on a  $\text{Var} [\Theta_{\mathbf{i}}(a)] \leq C (\varphi_a (P_k) + (\varphi_a (P_k))^2)$ , donc  $Z_n = O(\varrho_n^M \varphi_a (P_k))$ .

S'agissant de  $T_n$ , on utilise les techniques de Masry [148] et on considère les ensembles suivants

$$Q_1 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, a, \mathbf{1}) : 0 < \|\mathbf{i} - \mathbf{j}\| \leq \omega_n\},$$

$$Q_2 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, a, \mathbf{1}) : \|\mathbf{i} - \mathbf{j}\| > \omega_n\},$$

où,  $\omega_n$  est une suite réelle tendant vers  $+\infty$ . Ainsi,

$$T_n = \sum_{(\mathbf{i}, \mathbf{j}) \in Q_1} |\text{Cov} (\Theta_{\mathbf{i}}(a), \Theta_{\mathbf{j}}(a))| + \sum_{(\mathbf{i}, \mathbf{j}) \in Q_2} |\text{Cov} (\Theta_{\mathbf{i}}(a), \Theta_{\mathbf{j}}(a))| = T_n^1 + T_n^2.$$

D'une part, nous avons

$$\begin{aligned} T_n^1 &= \sum_{(\mathbf{i}, \mathbf{j}) \in Q_1} |\mathbb{E} [L_{\mathbf{i}}(a)L_{\mathbf{j}}(a)] - \mathbb{E} [L_{\mathbf{i}}(a)] \mathbb{E} [L_{\mathbf{j}}(a)]| \\ &\leq C \varrho_n^M \omega_n^M \varphi_a (P_k) \left( (\varphi_a (P_k))^{1/a} + \varphi_a (P_k) \right) \leq C \varrho_n^M \Omega_n^M \varphi_a (P_k)^{(a+1)/a}. \end{aligned}$$

D'autre part, nous avons  $T_n^2 = \sum_{(\mathbf{i}, \mathbf{j}) \in Q_2} |\text{Cov} (\Theta_{\mathbf{i}}(a), \Theta_{\mathbf{j}}(a))|$ .

Du faite que le noyau  $L$  est borné, à partir de Lemme 2.1-(ii) de Tran [186], on obtient

$$|\text{Cov} (\Delta_{\mathbf{i}}(a), \Delta_{\mathbf{j}}(a))| \leq C \varphi(\|\mathbf{i} - \mathbf{j}\|).$$

Ainsi

$$T_n^2 \leq C \sum_{(\mathbf{i}, \mathbf{j}) \in Q_2} \varphi(\|\mathbf{i} - \mathbf{j}\|) \leq C \varrho_n^M \sum_{\mathbf{i}: \|\mathbf{i}\| \geq \omega_n} \varphi(\|\mathbf{i}\|) \leq C \varrho_n^N \omega_n^{-Na} \sum_{\mathbf{i}: \|\mathbf{i}\| \geq \omega_n} \|\mathbf{i}\|^{Na} \varphi(\|\mathbf{i}\|).$$

Nous prenons,  $\omega_n = (\phi_x (h_K))^{-1/Na}$ , donc

$$T_n^2 \leq C \varrho_n^M \omega_n^{-Ma} \sum_{\mathbf{i}: \|\mathbf{i}\| \geq \omega_n} \|\mathbf{i}\|^{Ma} \varphi(\|\mathbf{i}\|) \leq C \varrho_n^N \phi_x (h_K) \sum_{\mathbf{i}: \|\mathbf{i}\| \geq \omega_n} \|\mathbf{i}\|^{Ma} \varphi(\|\mathbf{i}\|).$$

A partir de l'hypothèse (K2), on peut écrire  $T_n^2 \leq C \varrho_n^M \varphi_a (P_k)$ . De plus, avec ce choix de  $\omega_n$ , nous avons

$$T_n^1 \leq C \varrho_n^M \varphi_a (P_k),$$

où,

$$\text{Var} [D_1^*] = O(\varrho_n^M \varphi_a (P_k)).$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
 AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
 SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

En utilisant ce dernier résultat, avec les définitions de  $\varrho_n$ ,  $N$  et  $\zeta$ , on montre

$$\beta_1 \leq \exp(-C\zeta_0 \log \widehat{\mathbf{n}}).$$

L'hypothèse (K9-i) et le fait que  $L(n) \leq 2 \log n$ , d'après l'équation 3.5

$$\begin{aligned} K(n) \max_{j=1, \dots, K(n)} \mathbb{P} \left\{ \max_{1 \leq k \leq n} \|\sqrt{\mathbf{n}}\Theta_n(a)\|_{\mathcal{L}_{L,j}} > \zeta_0 \sqrt{\mathbf{n}\varphi_a\left(\frac{P_{k,j}}{2}\right) \log \mathbf{n}} \right\} \\ \leq C(\log \mathbf{n})\mathbf{n}^{-C'\zeta^2}. \end{aligned}$$

En choisissant maintenant  $\zeta_0$  tel que  $C'\zeta_0^2 > 1$ , on obtient :

$$\sup_{x_n \leq P_k \leq y_n} |\widehat{F}_D^a - \mathbb{E}[\widehat{F}_D^a]| = O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\varphi_a(x_n)}} \right).$$

□

**Lemme 3.2.** *Sous les hypothèses (K1), (K4-i), (K5) et (K7)*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h_j \leq s_n} |F_k^a(b) - \mathbb{E}\widehat{F}_N^a(b)| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}).$$

*Démonstration.* Étant donné que toutes les variables aléatoires sont distribuées de la même manière,

$$\forall b \in \mathcal{S}, \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h_j \leq s_n} |F_k^a(b) - \mathbb{E}\widehat{F}_N^a(b)| = \mathbb{E} \left[ L_1(a) \mathbb{I}_{B(a, P_k)}(A) [\mathbb{E}[J_1(b)/A] - F^a(b)] \right].$$

En intégrant par parties, on voit que

$$\mathbb{E}(J_1(b)/A) = \int_{\mathbb{R}} J(h_J^{-1}(b-z)) f_k^A(z) dz = h_J^{-1} \int_{\mathbb{R}} J'(h_J^{-1}(b-z)) F_k^A(z) dz.$$

Prenons en compte le changement de variable  $t = \frac{b-z}{h_J}$ , on obtient

$$\mathbb{E}(J_1(b)/A) = \int_{\mathbb{R}} J'(b) F_k^A(b - h_J t) dt,$$

donc

$$|\mathbb{E}(J_1(b)/A) - F_k^a(b)| \leq \int_{\mathbb{R}} J'(b) |F_k^A(b - h_J t) - F_k^a(b)| dt.$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

Sous l'hypothèse (K4), on a

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_K \leq y_n} \sup_{\nu_n \leq h, J \leq \nu_n} I_{B(a, P_k)}(A) |\mathbb{E}(J_1(b)/A) - F_k^a(b)| \leq \int_{\mathbb{R}} J'(b) (P_k^{\beta_1} + |t|^{\beta_2} h_J^{\beta_2}) dt.$$

Sachant que  $J'$  est une densité de probabilité, l'hypothèse (K6) complète donc la démonstration du Lemme 3.2, ce qui conduit finalement à

$$\forall b \in \mathcal{S}, \quad \sup_{x_n \leq P_K \leq y_n} \sup_{\nu_n \leq h, J \leq \nu_n} |F_k^a(b) - \mathbb{E}\widehat{F}_N^a(b)| \leq C (y_n^{\beta_3} + s_n^{\beta_4}).$$

□

**Lemme 3.3.** *Sous les mêmes conditions du Théorème 3.2, on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h, J \leq s_n} |\widehat{F}_N^a(b) - \mathbb{E}\widehat{F}_N^a(b)| = O\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}} \varphi_x(x_n)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

*Démonstration.* Par la compacité de  $S$ , on a

$$S \subset \bigcup_{j=1}^{\sigma_n} (\Gamma_j - \wp_n, \Gamma_j + \wp_n),$$

avec,  $\wp_n = \mathbf{n}^{-\alpha-1/2}$  et  $\sigma_n = O(\mathbf{n}^{\alpha+1/2})$ . Puis la monotonie de  $\mathbb{E}[\widehat{f}_N^a(b)]$  et  $\widehat{f}_N^a(b)$  donne, pour  $1 \leq j \leq \sigma_n$ , que

$$\begin{aligned} \mathbb{E}[\widehat{F}_N^a(\Gamma_j - \wp_n)] &\leq \mathbb{E}[\widehat{F}_N^a(\Gamma_j + \wp_n)], \\ \widehat{F}_N^a(\Gamma_j - \wp_n) &\leq \widehat{F}_N^a(\Gamma_j + \wp_n). \end{aligned}$$

A partir de la condition (K4-i), on obtient, pour tout  $b_1, b_2 \in S$ , que

$$|\mathbb{E}[\widehat{F}_N^a(b_1)] - \mathbb{E}[\widehat{F}_N^a(b_2)]| \leq C \left( d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4} \right).$$

Il s'ensuit que

$$\begin{aligned} &\sup_{t \in S} |\widehat{F}_N^a(b) - \mathbb{E}[\widehat{F}_N^a(b)]| \\ &\leq \max_{1 \leq j \leq \sigma_n, z \in \{\Gamma_j - \wp_n, \Gamma_j + \wp_n\}} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{F}_N^a(z)]| + 2C \frac{\wp_n}{h_J} \widehat{F}_D^a. \end{aligned}$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

Alors,

$$\wp_{\mathbf{n}} = o\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}}\right).$$

Donc, il ne reste plus qu'à prouver que

$$\begin{aligned} & \mathbb{P}\left(\sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq s_{\mathbf{n}}} \max_{1 \leq j \leq \sigma_{\mathbf{n}}, z \in \{\Gamma_j - \wp_{\mathbf{n}}, \Gamma_j + \wp_{\mathbf{n}}\}} |\widehat{F}_N^a(z) - \mathbb{E}[\widehat{F}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}}\right) \\ & \leq 2\sigma_{\mathbf{n}} \max_{1 \leq j \leq \sigma_{\mathbf{n}}, z \in \{\Gamma_j - \wp_{\mathbf{n}}, \Gamma_j + \wp_{\mathbf{n}}\}} \max_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \max_{\nu_{\mathbf{n}} \leq h, j \leq s_{\mathbf{n}}} |\widehat{F}_N^a(z) - \mathbb{E}[\widehat{F}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}}. \end{aligned}$$

Maintenant, regardons la quantité

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq s_{\mathbf{n}}} |\widehat{F}_N^a(z) - \mathbb{E}[\widehat{F}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\nu_{\mathbf{n}}\varphi_a(x_{\mathbf{n}})}}}\right),$$

pour tout  $z = \Gamma_j \mp \wp_{\mathbf{n}}, 1 \leq j \leq \sigma_{\mathbf{n}}$ . La preuve de l'inégalité ci-dessus est basée sur l'inégalité de Bernstein pour les processus empiriques, c'est-à-dire,  $\Theta_{i_1}(a) = \frac{1}{\sqrt{\mathbf{n}}} \sum_{i=1}^{\mathbf{n}} (L_i J_i - \mathbb{E}[L_i(a) J_i(a)])$ .

Alors, on obtient, pour tout  $z = \Gamma_j \mp \wp_{\mathbf{n}}, 1 \leq j \leq \sigma_{\mathbf{n}}$ , que

$$\mathbb{P}\left\{\sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq s_{\mathbf{n}}} |\widehat{F}_N^a(z) - \mathbb{E}[\widehat{F}_N^a(z)]| \geq \zeta'_0\right\} \leq \sigma_{\mathbf{n}}(\beta_1 + \beta_2).$$

La définition de  $\sigma_{\mathbf{n}}$ , l'équation 3.7 et l'hypothèse (K7) permettent d'écrire

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \widehat{\mathbf{n}}^{\alpha_+} \frac{1}{2} \beta_2 < \infty.$$

Alors qu'un choix approprié de  $\zeta_0$  donne

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \widehat{\mathbf{n}}^{\alpha_+} \frac{1}{2} \beta_1 < \infty.$$

Nous utilisons la même technique de démonstration que le Lemme 3.1, la seule différence est que nous utilisons  $\Theta_{i_1}(a)$  au lieu de  $\Theta_i(a)$ , cela conduit finalement à

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq s_{\mathbf{n}}} |F_N^a(b) - \mathbb{E}\widehat{F}_N^a(b)| = O_{p.co.}\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n}\varphi_a(x_{\mathbf{n}})}}}\right).$$

□

**Corollaire 3.2.1.** *Sous Les hypothèses du Lemme 3.1, nous avons que*

$$\sum_{\mathbf{n} \in \mathbb{N}^N} \mathbb{P} \left( \inf_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \widehat{F}_D^a < C \right) < \infty.$$

*Démonstration.* Par de simples arguments analytiques, on obtient à partir des conditions (K1) et (K5)-(K6)

$$\mathbb{E} \left[ \frac{1}{\mathbf{n}\varphi_a(P_k)} \sum_{i=1}^{\mathbf{n}} L(P_k^{-1}d(a, A_i)) \right] \rightarrow L(1/2) - \int_0^{1/2} L'(q)\varsigma_a(s)ds > 0.$$

Alors, pour  $\mathbf{n}$  assez grand il existe une constante  $C' > 0$ , telle que :

$$\mathbb{E} [\widehat{F}_D^a] \geq C' \quad \text{pour tout } P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}}).$$

En choisissant  $C = C'/2$ , nous obtenons :

$$\mathbb{P} \left( \inf_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})} \widehat{F}_D^a \leq C \right) \leq \mathbb{P} \left( \sup_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})} |\mathbb{E} [\widehat{F}_D^a] - \widehat{F}_D^a| \geq C \right)$$

et le Lemme 3.1 permet de conclure la démonstration du résultat recherché. □

**Théorème 3.3.** *Sous les hypothèses (K1),(K7) et,(K8-ii), (K9-ii) on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_j \leq s_{\mathbf{n}}} |\widehat{f}_k^a(b) - f_k^a(b)| = O(y_{\mathbf{n}}^{\beta_3}) + O(s_{\mathbf{n}}^{\beta_4}) + O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}} \right).$$

*Démonstration.* Les démonstrations sont basées, respectivement, sur les décompositions suivantes

$$\begin{aligned} \widehat{f}_k^a(b) - f_k^a(b) &= \frac{1}{\widehat{F}_D^a} \{ (\widehat{f}_N^a(b) - \mathbb{E}\widehat{f}_N^a(b)) - (f_k^a(b) - \mathbb{E}\widehat{f}_N^a(b)) \} + \\ &+ \frac{f_k^a(b)}{\widehat{F}_D^a} (\mathbb{E}\widehat{F}_D^a - \widehat{F}_D^a). \end{aligned} \tag{3.8}$$

On définit alors,

$$\widehat{f}_N^a(b) = \frac{1}{\widehat{\mathbf{n}}h_j \mathbb{E} [L(P_k^{-1}d(a, A_1))]} \sum_{i \in I_{\mathbf{n}}} L(P_k^{-1}d(a, A_i)) J'(h_j^{-1}(b - B_i)).$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

Ainsi, les deux Lemmes suivants sont nécessaires pour prouver le Théorème 3.3.  $\square$

**Lemme 3.4.** *Sous les hypothèses (K1), (K4-ii), (K5)-(K7) et (K9-ii) on a,*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |f_k^a(b) - \mathbb{E} \widehat{f_N^a}(b)| = O(y_{\mathbf{n}}^{\beta_3}) + O(s_{\mathbf{n}}^{\beta_4}).$$

*Démonstration.* En utilisant la stationnarité des données, la conditionnelle sur la variable explicative et le changement de la variable usuelle  $t = \frac{b-z}{h_J}$ , on obtient

$$h_J^{-1} \mathbb{E} (J'_1(b)/A) = h_J^{-1} \int_{\mathbb{R}} J'(b) f_k^A(b - h_J t) dt,$$

on en déduit que

$$|h_J^{-1} \mathbb{E} (J'_1(b)/A) - f_k^a(b)| \leq \int_{\mathbb{R}} J'(b) |f^A(b - h_J t) - f_k^a(b)| dt.$$

Sous la condition (K4-ii)

$$\forall b \in \mathcal{S}, \quad 1 = I = I_{B(a, P_k)}(A) |h_J^{-1} \mathbb{E} (J'_1(b)/A) - f_k^a(b)| \leq \int_{\mathbb{R}} J'(b) (h_k^{\beta_3} + |t|^{\beta_4} h_J^{\beta_4}) dt$$

Sachant que  $J'$  est une densité de probabilité, l'hypothèse (K6) complète donc la démonstration du Lemme 3.4, ce qui conduit finalement à

$$\forall b \in \mathcal{S}, \quad \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |f_k^a(b) - \mathbb{E} \widehat{f_N^a}(b)| \leq C (y_{\mathbf{n}}^{\beta_3} + s_{\mathbf{n}}^{\beta_4}).$$

$\square$

**Lemme 3.5.** *Sous les mêmes conditions du Théorème 3.3, on a*

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_J \leq s_{\mathbf{n}}} |\widehat{f_N^a}(b) - \mathbb{E} \widehat{f_N^a}(b)| = O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}} \right).$$

*Démonstration.* La démonstration est très similaire à celle du Lemme 3.3. En effet, nous considérons l'enveloppe de l'ensemble  $S$  suivante :

$$S \subset \bigcup_{j=1}^{\sigma_{\mathbf{n}}} (\Gamma_j - \wp_{\mathbf{n}}, \Gamma_j + \wp_{\mathbf{n}}),$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

avec  $\varrho_{\mathbf{n}} = \mathbf{n}^{-1/2+\frac{-3}{2}\alpha}$  et  $\sigma_{\mathbf{n}} = O\left(\mathbf{n}^{1/2+\frac{3}{2}\alpha}\right)$ . Puis la monotonie de  $\mathbb{E}[\widehat{f}_N^a(b)]$  et  $\widehat{f}_N^a(b)$

donne, pour  $1 \leq j \leq \sigma_{\mathbf{n}}$ , que

$$\begin{aligned}\mathbb{E}[\widehat{f}_N^a(\Gamma_j - \varrho_{\mathbf{n}})] &\leq \mathbb{E}[\widehat{f}_N^a(\Gamma_j + \varrho_{\mathbf{n}})] \\ \widehat{f}_N^a(\Gamma_j - \varrho_{\mathbf{n}}) &\leq \widehat{f}_N^a(\Gamma_j + \varrho_{\mathbf{n}}).\end{aligned}$$

Alors, à partir de la condition (K4-ii), on obtient, pour tout  $b_1, b_2 \in S$ , que

$$|\mathbb{E}[\widehat{f}_N^a(b_1)] - \mathbb{E}[\widehat{f}_N^a(b_2)]| \leq C \left( d(a_1, a_2)^{\beta_3} + |b_1 - b_2|^{\beta_4} \right).$$

Il s'ensuit que

$$\begin{aligned}&\sup_{z \in S} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{f}_N^a(z)]| \\ &\leq \max_{1 \leq j \leq \sigma_{\mathbf{n}}} \max_{z \in \{\Gamma_j - \varrho_{\mathbf{n}}, \Gamma_j + \varrho_{\mathbf{n}}\}} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{f}_N^a(z)]| + 2C \frac{\varrho_{\mathbf{n}}}{h_j^2} \widehat{F}_D^a.\end{aligned}$$

Donc,

$$\varrho_{\mathbf{n}} = o\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}}}\right).$$

Il ne reste plus qu'à prouver que

$$\begin{aligned}\mathbb{P}\left(\sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq \sigma_{\mathbf{n}}} \max_{1 \leq j \leq \sigma_{\mathbf{n}}} \max_{z \in \{\Gamma_j - \varrho_{\mathbf{n}}, \Gamma_j + \varrho_{\mathbf{n}}\}} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{f}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}}}\right) \\ \leq \left(2\sigma_{\mathbf{n}} \max_{1 \leq j \leq \sigma_{\mathbf{n}}} \max_{z \in \{\Gamma_j - \varrho_{\mathbf{n}}, \Gamma_j + \varrho_{\mathbf{n}}\}} \max_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \max_{\nu_{\mathbf{n}} \leq h, j \leq \sigma_{\mathbf{n}}} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{f}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}}}\right).\end{aligned}$$

Maintenant, regardons la quantité

$$\mathbb{P}\left(\sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h, j \leq \sigma_{\mathbf{n}}} |\widehat{f}_N^a(z) - \mathbb{E}[\widehat{f}_N^a(z)]| > \zeta \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}}}\right),$$

pour tout  $z = \Gamma_j \mp \varrho_{\mathbf{n}}$ ,  $1 \leq j \leq \sigma_{\mathbf{n}}$ .

La preuve de l'inégalité ci-dessus est basée sur l'inégalité de Bernstein pour les processus empiriques,

$$\Theta_{\mathbf{i}_2}(a) = \frac{1}{\sqrt{\mathbf{n}}} \sum_{i=1}^{\mathbf{n}} (L_i J'_i - \mathbb{E}[L_i J'_i]).$$

Alors, on obtient, pour tout  $z = \Gamma_j \mp \varphi_{\mathbf{n}}, 1 \leq j \leq \sigma_{\mathbf{n}}$ , que

$$\mathbb{P} \left\{ \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_{J \leq s_{\mathbf{n}}}} |\widehat{f}_{\mathbf{N}}^a(z) - \mathbb{E}[\widehat{f}_{\mathbf{N}}^a(z)]| \geq \zeta'_0 \right\} \leq \sigma_{\mathbf{n}}(\beta'_1 + \beta'_2),$$

où,  $\beta'_1 \leq \exp(-C(\zeta_0) \log \widehat{\mathbf{n}})$  et

$$B'_2 \leq \widehat{\mathbf{n}} \varrho_{\widehat{\mathbf{n}}}^M (\log \widehat{\mathbf{n}})^{-1/2} (\widehat{\mathbf{n}} h_{J \varphi_a}(P_k))^{-1/2} \varphi(\varrho_{\widehat{\mathbf{n}}}).$$

Nous prenons,  $\varrho_{\mathbf{n}} C \left( \frac{\widehat{\mathbf{n}} h_{J \varphi_a}(P_k)}{\log \widehat{\mathbf{n}}} \right)^{1/2M}$ , la définition de  $\sigma_{\mathbf{n}}$  sous l'hypothèse (K7), un bon choix de  $\zeta_0$  permet d'écrire

$$\sum_{\mathbf{n} \in I_{\mathbf{n}}} \widehat{\mathbf{n}}^{\frac{1}{2}} (\beta'_2 + \beta'_1) < \infty.$$

Nous utilisons la même technique de démonstration que le Lemme 3.1, la seule différence est que nous utilisons  $\Theta_{i_2}(a)$  au lieu de  $\Theta_i(a)$ , cela conduit finalement à

$$\sup_{b \in \mathcal{S}} \sup_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \sup_{\nu_{\mathbf{n}} \leq h_{J \leq s_{\mathbf{n}}}} |f_{\mathbf{N}}^a(b) - \mathbb{E} \widehat{f}_{\mathbf{N}}^a(b)| = O_{p.co.} \left( \sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_{\mathbf{n}} \varphi_a(x_{\mathbf{n}})}} \right).$$

□

## 3.5 Application

### 3.5.1 Estimation du point de risque

Dans cette partie, nous proposons d'estimer le point à haut-risque en  $\mathcal{S}$ , noté  $\eta(a)$ , défini comme

$$h_k^a(\eta(a)) = \max_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \max_{\nu_{\mathbf{n}} \leq h_{J \leq s_{\mathbf{n}}}} \max_{b \in \mathcal{S}} h_k^a(b). \quad (3.9)$$

Ce modèle a une solide expérience en statistique, en particulier en analyse des risques (voir Quintela-del-Rio [159]). Dans notre cadre fonctionnel, nous supposons qu'il existe un seul point  $\eta(a)$  dans  $\mathcal{S}$  avec  $\sup_{P_k \in (x_{\mathbf{n}}, y_{\mathbf{n}})}$ ,  $\sup_{h_{J \in (\nu_{\mathbf{n}}, s_{\mathbf{n}})}}$  et vérifié (3.9). L'estimateur naturel de  $\eta(a)$ , noté  $\hat{\eta}(a)$ , est le suivant :

$$\widehat{h}_k^a(\hat{\eta}(a)) = \max_{x_{\mathbf{n}} \leq P_k \leq y_{\mathbf{n}}} \max_{\nu_{\mathbf{n}} \leq h_{J \leq s_{\mathbf{n}}}} \max_{b \in \mathcal{S}} \widehat{h}_k^a(b). \quad (3.10)$$

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

En général, cette estimation n'est pas unique. Par conséquent, dans le reste de cet article,  $\hat{\eta}(a)$  sera utilisé pour désigner toute variable aléatoire vérifiant (3.10).

Pour analyser le taux de convergence presque complète de l'estimateur  $\hat{\eta}(a)$ , nous gardons les mêmes hypothèses de la section précédente et supposons que la fonction  $h_k^a$  est de classe  $C^2$  par rapport à  $b$ , comme indiqué ci-dessous.

$$h_k^{a'}(\eta(a)) = 0 \quad \text{et} \quad h_k^{a''}(\eta(a)) > 0. \quad (3.11)$$

Enfin, le Théorème 3.1 nous permet de déduire le Corollaire suivant

**Corollaire 3.3.1.** *Dans les conditions (3.1), (3.2), (K1)-(K7), (K4)-(K5) et (K9), et (3.11) nous avons*

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h_j \leq s_n} |\hat{\eta}(a) - \eta(a)| = O\left(y_n^{\beta_3/2}\right) + O\left(s_n^{\beta_4/2}\right) + O_{p.co.} \left( \left( \frac{\log n}{n \nu_n \varphi_a(x_n)} \right)^{1/4} \right).$$

*Démonstration.* Sous l'hypothèse (K5), le développement de Taylor de la fonction  $h_k^a$  au voisinage de  $\eta(a)$ , en particulier pour le point  $\hat{\eta}(a)$ , est

$$h_k^a(\hat{\eta}(a)) = h_k^a(\eta(a)) + (\hat{\eta}(a) - \eta(a))^2 \frac{h_k^{a''}(\eta^*)}{j!},$$

où,  $\eta^*$  est compris entre  $\hat{\eta}(a)$  et  $\eta(a)$ , il est donc nécessairement dans le compact  $\mathcal{S}$ . Ainsi

$$|\hat{\eta}(a) - \eta(a)|^2 \leq \frac{2!}{\min_{b \in \mathcal{S}, P_k \in (x_n, y_n), h_j \in (\nu_n, s_n)} h_k^{a''}(b)} |h_k^a(\hat{\eta}(a)) - h_k^a(\eta(a))|,$$

et avec l'utilisation d'arguments analytiques, il est démontré que

$$|h_k^a(\hat{\eta}(a)) - h_k^a(\eta(a))| \leq 2 \sup_{b \in \mathcal{S}, P_k \in (x_n, y_n), h_j \in (\nu_n, s_n)} |\hat{h}_k^a(b) - h_k^a(b)|.$$

En utilisant le résultat du Théorème 3.1, nous montrons que

$$\sup_{b \in \mathcal{S}} \sup_{x_n \leq P_k \leq y_n} \sup_{\nu_n \leq h_j \leq s_n} |\hat{\eta}(a) - \eta(a)| = O\left(y_n^{\beta_3}\right) + O\left(s_n^{\beta_4}\right) + O_{p.co.} \left( \sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}} \right).$$

□

### 3.5.2 Mise en place du paramètre de lissage uniforme pour le choix des résultats du paramètre de lissage

Lors de l'application d'un estimateur non paramétrique dans la pratique, on est confronté à la difficulté du choix de le paramètre de lissage. Une telle décision doit être basée sur des données, ce qui signifie que le paramètre de lissage seront utilisées dans la pratique sont des variables aléatoires qui varient en fonction de l'ensemble de l'échantillon statistique. Dans les modèles statistiques, la théorie asymptotique de le paramètre de lissage data-driven a toujours été liée à une sorte d'uniforme dans les études de paramètre de lissage (UIB). Notez, par exemple "le Lemme 1, p. 1473" de Härdle *et al.* [106], même si le mot UIB n'est apparu dans la littérature que bien plus tard.

Tout au long de cette section, nous verrons comment les résultats généraux de l'UIB rapportés précédemment dans ce travail peuvent être appliqués à un large éventail d'applications de le paramètre de lissage data-driven dans des données spatiales fonctionnelles avec des données de structure de mélange. Ceci sera décrit dans la Section 3.5.3 le problème d'estimation (cdf, fonctions de densité et de hasard).

### 3.5.3 Sélection du paramètre de lissage dans la fonction de distribution de hasard

Le même genre d'idées peut être développé pour traiter les paramètres de lissages aléatoires dans d'autres contextes que la régression. Pour éviter les répétitions fastidieuses, les situations étudiées dans la Section 3.2 seront présentées toutes ensemble. Soient  $\tilde{P}_k$  et  $\tilde{h}_J$  des variables aléatoires prenant respectivement des valeurs dans les intervalles  $(x_n, y_n)$  et  $(\nu_n, s_n)$ . On note  $\tilde{F}_k, \tilde{f}_k$  et  $\tilde{h}_k$  les estimateurs obtenus en incorporant les paramètres de lissages aléatoires  $\tilde{P}_k$  et  $\tilde{h}_J$  en  $\hat{F}_k, \hat{f}_k$  et  $\hat{h}_k$ . Alors, comme applications directes des Théorèmes 3.2, 3.3 et Théorème 3.1, on obtient les résultats suivants :

$$\tilde{F}_k^a(b) - F_k^a(b) = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \varphi_a(x_n)}}\right), \text{ p.co.} \quad (3.12)$$

$$\tilde{f}_k^a(b) - f^a(b) = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O\left(\sqrt{\frac{\log \mathbf{n}}{\mathbf{n} \nu_n \varphi_a(x_n)}}\right), \text{ p.co.} \quad (3.13)$$

et

$$\tilde{h}_k^a(b) - h_k^a(b) = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O\left(\sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}}\right), \text{ p.co.} \quad (3.14)$$

Comme le montrent les exemples ci-dessous, ces solutions génériques pour les paramètres de lissage aléatoires ont une applicabilité évidente dans les problèmes de sélection de le paramètre de lissage data-driven.

### Example1- Fonction de distribution conditionnelle

En se concentrant sur l'estimation c.d.f., une prédiction statistique peut être effectuée au moyen de la médiane conditionnelle, où  $\hat{B}_i^{(l)}$  est la prédiction leave-one-out de  $B_i$  est défini comme étant la solution de l'équation suivante :

$$\hat{F}_k^{A_i, -i}(b) = \frac{1}{2},$$

où,

$$\hat{F}_k^{a, -i} = \frac{\sum_{i \in I_n} \sum_{j \neq i} L(P_k^{-1}d(a, A_j)) J(h_J^{-1}(b - B_i))}{\sum_{i \in I_n} \sum_{j \neq i} L(P_k^{-1}d(a, A_j))}.$$

En pratique, sur la base de l'échantillon statistique, on utilise un paramètre de lissage data-driven, définie comme

$$\tilde{P}_k = \tilde{P}_k(A_1, B_1, \dots, A_M, B_M).$$

En conséquence, des résultats de cohérence peuvent être obtenus pour tout type de sélection automatique de paramètre de lissage data-driven en conséquence de (3.12). Une des méthodes les plus utilisée est la validation croisée, qui consiste à minimiser le critère de prédiction des moindres carrés suivant :

$$\left(\tilde{p}_{k,CV}^{(l)}, \tilde{h}_{j,CV}^{(l)}\right) = \arg \min_{P_k \in (x_n, y_n), h_J \in (\nu_n, s_n)} \sum_{i=1}^n \left(B_i - \hat{B}_i^{(l)}\right)^2.$$

### Example2- Fonction de densité

Le mode conditionnel peut être utilisé pour faire des prédictions à partir de l'estimateur de densité conditionnelle.  $B_i$  pourrait bien être prédit précisément comme la réponse au

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

problème de maximisation suivant :

$$\hat{B}_i^{(n)} = \arg \max_b \hat{f}_k^{A_i, -i}(b),$$

où l'approche de leave-one-out définit  $\hat{f}_k^{a, -i}$  est

$$\hat{f}_k^{a, -i} = \frac{h_J^{-1} \sum_{i \in I_n} \sum_{j \neq i} L(P_k^{-1} d(a, A_j)) J'(h_J^{-1}(b - B_i))}{\sum_{i \in I_n} \sum_{j \neq i} L(P_k^{-1} d(a, A_j))}.$$

Ensuite, il existe des choix plus évidents pour les paramètres de lissages qui est l'approche guidée (data-driven), qui consiste à

$$\left( \tilde{P}_{k,CV}^{(n)}, \tilde{h}_{J,CV}^{(n)} \right) = \arg \min_{P_k \in (\mathbf{x}_n, \mathbf{y}_n), h_J \in (\nu_n, s_n)} \sum_{i=1}^n \left( B_i - \hat{B}_i^{(n)} \right)^2.$$

Les résultats suivants sont des conséquences directes des Théorèmes 3.2, 3.3 et Théorème 3.1.

**Corollaire 3.3.2.** Soit  $j = 1$  ou  $j = 2$ , et notons  $\hat{F}_k^{(j),a}$ ,  $\hat{f}_k^{(j),a}$  et  $\hat{h}_k^{(j),a}$ , respectivement, les estimateurs  $\hat{F}_k$ ,  $\hat{f}_k$  et  $\hat{h}$  construit avec paramètres de lissages par validées croisées  $\tilde{P}_{k,CV}^{(j)}$  et  $\tilde{h}_{J,CV}^{(j)}$ . Alors, les trois résultats suivants tiennent :

- Sous les conditions du Théorème 3.2, on a que :

$$\left| \hat{F}_k^{(j),a}(b) - F_k^a(b) \right| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O\left( \sqrt{\frac{\log n}{n \varphi_a(x_n)}} \right), p.co.$$

- Sous les hypothèses du Théorème 3.3, on obtient que

$$\left| \hat{f}_k^{(j),a}(b) - f_k^a(b) \right| = O(y_n^{\beta_3}) + O(s_n^{\beta_4}) + O\left( \sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}} \right), p.co.$$

- Sous les conditions du Théorème 3.1, on a que

$$\left| \hat{h}_k^{(j),a}(b) - h_k^a(b) \right| = O(y_n^{\beta_3}) + O(s_n^{\beta_2}) + O\left( \sqrt{\frac{\log n}{n \nu_n \varphi_a(x_n)}} \right), p.co.$$

Ceux-ci sont particulièrement intéressants dans la pratique car ils automatisent entièrement les problèmes pratiques qui surviennent lors de l'utilisation des estimateurs  $\hat{F}_k$ ,  $\hat{f}_k$

et  $\hat{h}_k$ , notre travail semble être le premier dans ce domaine pour les paramètres de hasard, de distribution conditionnelle et de densité conditionnelle.

### 3.5.4 Application de données simulées

Dans cette section, nous comparons les performances en échantillon fini de la fonction de hasard  $k$ NN (K.H) dans le contexte de la prédiction fonctionnelle via la fonction de hasard classique (C.H) via une courte étude de Monte Carlo. Plus précisément, nous comparons l'efficacité à échantillon fini des deux fonctions de régression en tant qu'outils de prédiction spatiale. Afin de mettre en évidence la principale caractéristique de notre procédure, nous comparons leur sensibilité à la présence de valeurs aberrantes. Pour cela, nous considérons le modèle suivant

$$B_{\mathbf{i}} = 5 \log((4 - A_{\mathbf{i}})^2 + 2) + \epsilon_{\mathbf{i}}, \quad \mathbf{i} = (i_1, i_2),$$

où  $\epsilon_{(i_1, i_2)}$  suit une distribution normale de moyenne 0 et de variance 0,25. Par souci de simplicité, nous considérons le même processus spatial univarié  $A_1$  utilisé par Cressie [47] défini par

$$A_{\mathbf{i}} = \left( \sqrt{(2/m)} \right) \sum_{\varpi=1}^m (\cos(i_1 w_{\varpi,1} + i_2 w_{\varpi,2} + v_{\varpi})),$$

où  $w_{\varpi,s}, s = 1, 2 \quad \varpi = 1, 2 \dots 500$  sont indépendamment, identiquement distribués avec la distribution normale standard et sont indépendants de  $v_{\varpi}, \varpi = 1, \dots 500$  qui sont indépendamment et identiquement distribués avec la distribution uniforme sur  $[-\pi, \pi]$ . Rappelons que comme  $m \rightarrow \infty$ ,  $A_1$  est un processus spatial ergodique gaussien (voir, Cressie [47]) qui est un exemple de processus spatial  $\alpha$ -mélangeant.

Nous générons le champ aléatoire  $(A_1, B_{\mathbf{i}})$  sur les sites  $\hat{\mathbf{n}} = 30 \times 30$ . La corrélation spatiale est contrôlée dans notre analyse théorique par la condition de mélange fort, qui est définie comme (3.1)- (3.3). Ainsi, en pratique, la corrélation spatiale fonctionnelle est intégrée au calcul de l'estimateur en introduisant un ensemble de voisinage donné par les équations suivantes pour chaque site  $\ell$  :

$$U_{\ell} = \{\mathbf{i}, \text{ tel que } \varphi(\|\mathbf{i} - \ell\|) \leq \|\mathbf{i} - \ell\|^{-\ell}\}.$$

Puisque nous nous sommes limités au cas isotrope dans lequel la dépendance spatiale est simplement fonction de la distance entre les sites, nous pouvons procéder avec l'ensemble

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

de voisinage défini par :

$$U_\ell = \{\mathbf{i}, \|\mathbf{i} - \mathbf{k}\| \leq \xi_n\}$$

où  $\xi_n$  est une séquence appropriée de nombres réels positifs. La quantité  $\xi_n$  est sélectionnée de manière optimale sur les sites des voisins les plus proches par rapport à la norme euclidienne sur les coordonnées. Pour plus de détails sur le choix de  $U_\ell$ , voir Dabo et Yao [49]. Nous estimons, pour chaque site fixe  $\ell$ , la fonction de hasard conditionnelle  $k$ NN par :  $\forall b \in \mathbb{R}$ ,

$$\hat{h}_k^{A_\ell}(b) = \frac{h_J^{-1} \sum_{\mathbf{i}, \mathbf{j} \in I_n} L(P_k^{-1}d(a, A_\ell)) J' \left( \frac{b - B_{\mathbf{i}}}{h_J} \right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}{\sum_{\mathbf{i}, \mathbf{j} \in I_n} L(P_k^{-1}d(a, A_\ell)) J \left( \frac{b - B_{\mathbf{i}}}{h_J} \right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}, \quad (3.15)$$

où,  $\mathbb{I}_{U_\ell}$  est la fonction indicatrice de l'ensemble  $U_\ell$ .

Deuxièmement, nous devons sélectionner un semi-métrique approprié  $d(\cdot, \cdot)$ , noyau  $L(\cdot)$ . Ensuite, nous choisissons le noyau quadratique asymétrique défini comme  $L(u) = \frac{3}{4}(1 - u^2) 1_{[0,1]}(u)$ . Notons que du faite de la régularité des courbes  $A_i(t)$ , on utilise la semi-métrique suivante basée sur la dérivée première :

$$d^{deriv}(A_i, A_j) = \sqrt{\int_0^{\pi/2} (A'_i(t) - A'_j(t))^2 dt}, \quad \forall A_i, A_j \in \mathcal{G}.$$

Rappelons que l'estimateur classique du hasard conditionnel est défini comme suit :

$\forall b \in \mathbb{R}$ ,

$$\hat{h}_{\mathbf{n}}^{A_\ell}(b) = \frac{h_J^{-1} \sum_{\mathbf{i}, \mathbf{j} \in I_n} L(h_L^{-1}d(a, A_\ell)) J' \left( \frac{b - B_{\mathbf{i}}}{h_J} \right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}{\sum_{\mathbf{i}, \mathbf{j} \in I_n} L(h_L^{-1}d(a, A_\ell)) J \left( \frac{b - B_{\mathbf{i}}}{h_J} \right) \mathbb{I}_{U_\ell}(\mathbf{i}) \mathbb{I}_{U_\ell}(\mathbf{j})}. \quad (3.16)$$

L'efficacité des estimateurs est évaluée par l'erreur quadratique moyenne empirique (MSE), définit par

$$MSE(\hat{h}_k^{A_\ell}(b)) = \mathbf{n}^{-1} \sum_{i=1}^{\mathbf{n}} (h_k^{A_i}(B_i) - \hat{h}_k^{A_i}(B_i))^2, \quad MSE(\hat{h}_{\mathbf{n}}^{A_\ell}(b)) = \mathbf{n}^{-1} \sum_{i=1}^{\mathbf{n}} (h_{\mathbf{n}}^{A_i}(B_i) - \hat{h}_{\mathbf{n}}^{A_i}(B_i))^2.$$

Nous utilisons la procédure de validation croisée proposée par Xu [191] pour laquelle le

CHAPITRE 3. LE TAUX DE CONVERGENCE DE LA FONCTION DE HASARD  
 AVEC LA VARIABLE EXPLICATIVE FONCTIONNELLE : CAS DES DONNÉES  
 SPATIALES AVEC LA MÉTHODE DES K PLUS PROCHES VOISINS

---

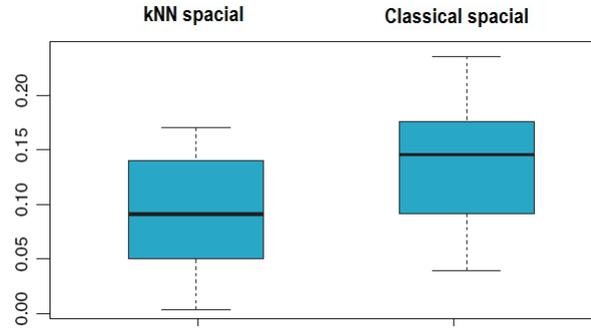


FIGURE 3.1 – Le MSE des deux modèles

paramètre de lissage  $h_L$  est choisie via la règle suivante :

$$h_{opt} = \arg \min_{h_L} CV_1(h_L) = \arg \min_{h_L} \sum_{i \in \mathcal{I}_n} | B_i - \hat{r}^{(-i)}(A_i) |,$$

où  $\hat{r}^{(-i)}$  est l'estimateur de leave-one-out de  $(\hat{h}_n^{A_\ell})$ .

Concrètement, pour le  $k$ NN robuste nous sélectionnons  $k_{opt} = \arg \min_k CV_2(k)$ , où

$$CV_2(k) = \arg \min_k \sum_{i \in \mathcal{I}_n} | B_i - \hat{q}^{(-i)}(A_i) |,$$

avec  $\hat{q}^{(-i)}$  est l'estimateur de leave-one-out de courbe de  $(\hat{h}_k^{A_\ell})$ .

Le boxplot de MSE des deux modèles est donnée à la Figure 3.1. Nous observons qu'il n'y a pas de différence significative entre ces prédicteurs spatiaux. Les deux prédicteurs sont fondamentalement équivalents et tous deux montrent le bon comportement. Maintenant, afin d'étudier les caractéristiques de notre approche, nous avons introduit des valeurs aberrantes artificielles en multipliant les valeurs de 15% de  $Y$  par 10. Nous donnons en graphique boxplot les erreurs (MSE) des deux modèles. De plus, nous voyons sur la Figure 3.2 que l'erreur de la fonction de hasard  $k$ NN est bien meilleure que l'erreur classique dans ce cas. De plus, en regardant les deux figures, il apparaît clairement que le MSE du modèle (C.H) a radicalement changé par rapport au cas (K.H). Cette affirmation confirme que la fonction de hasard  $k$ NN est plus robuste que la régression classique.

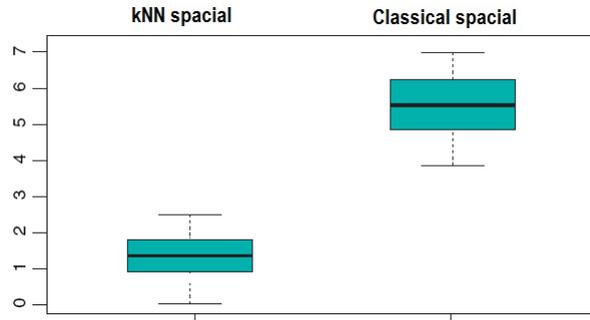


FIGURE 3.2 – MSE en présence de valeurs aberrantes

### 3.5.5 Application de données réelles

L'objectif principal de cette section est d'appliquer les résultats théoriques de la section précédente à des données réelles. En particulier, nous analysons l'efficacité de la fonction de hasard dans le contexte de la prédiction fonctionnelle spatiale  $k$ NN via la fonction de hasard classique.

Dans cet exemple de données réelles, nous recherchons un moyen d'estimer le logarithme des précipitations totales en fonction de la courbe de température maximale mensuelle. Le prédicteur fonctionnel  $A_i$  est la courbe des températures maximales mensuelles dans la  $i$ -ème station climatique (spécifiée par ses coordonnées géographiques) dans une période  $\Upsilon$ , et  $B_i$  est le logarithme du total précipitations dans la même station et la même période, selon les notations de la section précédente. Pour cela, nous avons utilisé les données mensuelles de température et de précipitations de 125 stations collectées en 2000 et 2010. Ces observations peuvent être trouvées à l'url suivante : <ftp://ftp.ncdc.noaa.gov/pub/data/ushcn/v2/monthly>. Dans la Figure 3.3, les covariables fonctionnelles sont données.

Afin d'évaluer l'efficacité de l'estimateur présenté, représenté par  $\hat{h}_n^{A_\ell}(b)$  dans l'équation (3.15) et de la comparer avec celle qui ne considère pas directement la distance entre les positions et est représentée par  $\hat{h}_n^{A_k}(b)$  dans l'équation (3.16), nous divisons nos données de manière aléatoire  $(A_i, B_i)_i$  en deux sous-ensembles différents :

- Échantillon d'apprentissage  $(A_i, B_i)_{i \in I}$  (114 stations),
- Échantillon de test  $(A_i, B_i)_{i \in I'}$  (11 stations).

Pour calculer l'estimateur pour les deux méthodes, nous prenons la même fonction de noyau quadratique et la semi-métrique  $d^{\text{deriv}}(\cdot, \cdot)$  dans l'étude de simulation, ainsi que les sélections des paramètres  $k_{opt}$  (de la méthode  $k$ NN) et  $h_{opt}$  (pour le noyau).

Comme mesure de précision, nous utilisons la procédure d'erreur quadratique moyenne

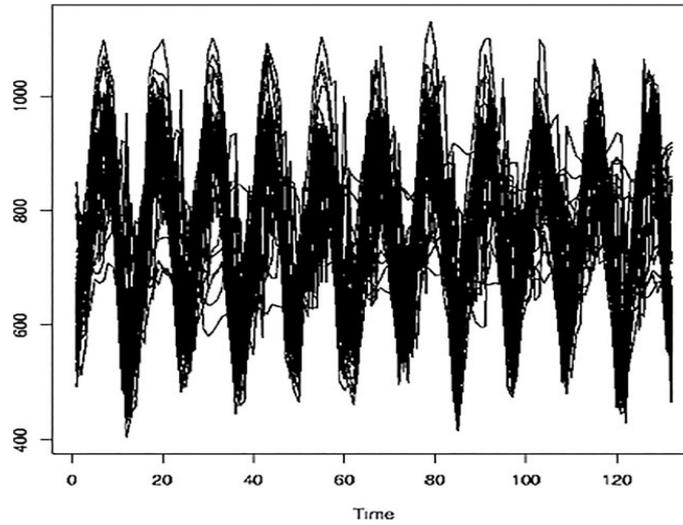


FIGURE 3.3 – Températures mensuelles les plus élevées sur 125 sites météorologiques aux États-Unis

(MSE), qui est décrite comme suit :

$$MSE(\hat{h}_k^{A_\ell}(b)) = \frac{1}{11} \sum_{i \in I'} (h_k^{A_i}(B_i) - \hat{h}_k^{A_i}(B_i))^2$$

et

$$MSE(\hat{h}_n^{A_\ell}(b)) = \frac{1}{11} \sum_{i \in I'} (h_n^{A_i}(B_i) - \hat{h}_n^{A_i}(B_i))^2.$$

Les résultats des deux méthodes sont tracés par rapport aux vraies valeurs dans la Figure 3.4, où les valeurs prédites sont tracées par rapport aux vraies valeurs.

La partie gauche de la Figure 3.4 illustre le cas dans lequel la corrélation spatiale  $k$ NN est utilisée, tandis que la moitié droite décrit le cas dans lequel la corrélation spatiale classique est utilisée. La ligne continue contrôle l'efficacité de la prédiction, en ce sens que l'efficacité de la méthode de prédiction est mesurée par la proximité des points tracés par rapport à la ligne continue. Il est évident que l'estimation utilisant la corrélation spatiale  $k$ NN est nettement meilleure et plus efficace que l'estimation traditionnelle. L'erreur quadratique moyenne obtenue respectivement dans les deux situations le prouve,  $MSE(\hat{h}_k^{A_k}(b)) = 0.21$  alors que,  $MSE(\hat{h}_n^{A_\ell}(b)) = 0.36$ .

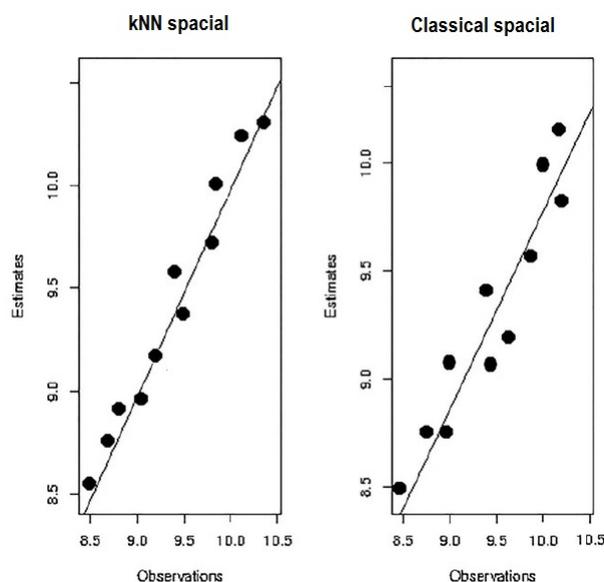


FIGURE 3.4 – Comparaison des résultats de prédiction entre les deux méthodes.

### 3.6 Conclusion

L'approche de fiabilité  $k$ NN uniforme est une alternative de lissage qui permet le développement d'un estimateur adaptatif pour une variété de problèmes statistiques, y compris le choix de le paramètre de lissage. Dans notre situation, de plus, la consistance uniforme n'est pas une extension directe de l'approche ponctuelle, car elle nécessite l'utilisation de méthodes et techniques supplémentaires. L'hypothèse selon laquelle le paramètre de lissage dans la méthode  $k$ NN est une variable aléatoire ajoute à la complexité de ce problème. Dans le cas des résultats spatiaux, l'innovation clé de cette approche est d'estimer la fonction de hasard en mélangeant deux techniques statistiques essentielles : l'approche spatiale et les procédures  $k$ NN. Cette stratégie a permis le développement d'un nouvel estimateur qui combine les avantages des deux méthodes. Pour résumer, Le comportement de l'estimateur développé et la comparaison des deux nuages de points indiquent que l'approche du  $k$ NN hazard avec des données spatiales donne des résultats légèrement meilleurs que l'approche classique du hazard avec des données spatiales. Cela est confirmé par le résultat de l'erreur quadratique moyenne (MSE).

## BIBLIOGRAPHIE

- [1] Abraham, C., Matzner-Løber, E. and Molinari, N., Cornillon, P. A. (2003) Scandinavian Journal of Statistics, vol. 30, issue 3, 581-595.
- [2] Almanjahie, I.M., Chikr, Z., Elmezouar, Laksaci, A., Rachdi, M. (2018). kNN local linear estimation of the conditional cumulative distribution function : Dependent functional data case. Comptes Rendus. Mathématique. Vol.10, issue 356, 1036-1039.
- [3] Almanjahie, I.M., Aissiri, K., Laksaci, A., Chiker, Z., el mezouar. (2020). The k nearest neighbors smoothing of the relative-error regression with functional regressor, Communications in Statistics - Theory and Methods. Vol.10, No.356, 1-14.
- [4] Altman, D. G. (1994). The scandal of poor medical research, British Medical Journal Publishing Group. Vol.308, 283-284.
- [5] Aneiros-Perez, G., Cardot, H., Estevez-Perez, G., Vieu, P. (2004). Maximum ozone concentration forecasting by functional non-parametric approaches. Environmetrics. Vol.15, 675–685.
- [6] Attaoui, S., Laksaci, A., Ould-Said, E. (2015). Asymptotic results for an M-estimator of the regression function for quasi-associated processes, Functional Statistics and Applications, Contributions to Statistics. Vol. 200, No. 67, 45-53.
- [7] Attouch, M.K, Belabed, F. (2014). The  $k$  nearest neighbors estimation of the conditional hazard function for functional data. REVSTAT. Vol.12, 273–97.
- [8] Attouch, M.K, Bouabsa, W., Chiker el mozoaur, Z. (2018). The k-nearest neighbors estimation of the conditional mode for functional data under dependency, International Journal of Statistics & Economics. Vol. 19, No. 1, 48–60.

## BIBLIOGRAPHIE

---

- [9] Attouch, M.K, Laksaci, A. and Ould Said E. (2009). Asymptotic distribution of robust estimator for functional nonparametric models. *Comm. Statist. Theory and Methods*. Vol.38, 1317-1335.
- [10] Attouch, M.K., Bouabsa, W. (2013). The  $k$ -nearest neighbors estimation of the conditional mode for functional data, *Rev. Roumaine Math. Pures Appl.* Vol. 58, No. 4, 393-415.
- [11] Attouch, M.K, Benchikh, T. (2012). Asymptotic distribution of robust  $k$ -nearest neighbour estimator for functional nonparametric models. *Mathematic Vesnic*. Vol.64, No. 4, 275-285.
- [12] Attouch, M.K., Laksaci, A. Ould Saïd, E. (2010). Asymptotic normality of a robust estimator of the regression function for functional time series data, *Journal of The Korean Statistical Society*. Vol. 39, No. 4, 489-500.
- [13] Attouch, M.K., Laksaci, A. Ould Saïd, E. (2012). Robust regression for functional time series data, *Journal of The Japan Statistical Society*. Vol. 42, No. 3, 125-143.
- [14] Attouch, M.K., Laksaci, A., Messabihi, N. (2015) Nonparametric relative error regression for spatial random variables, *Statistical papers*. Vol.58, No. 4, 987-1008.
- [15] Andrews, D. F. (1974). A robust method for multiple linear regression. *Thechnometrics*. Vol.16, 523-531.
- [16] Aurzada, F., Simon, T.(2007). Small ball probabilities for stable convolutions. *ESAIM Probab. Stat. (electronic)*. vol.11, 327-343 .
- [17] Azzedine, N., Laksaci, A., Ould Saïd, E. (2008). On the robust nonparametric regression estimation for functional regressor, *Statistics & Probability Letters*. Vol. 78, No. 18, 3216-3221.
- [18] Barlow, R. E. Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing, Probability Models*, Holt, Rinehart & Winston, New York.
- [19] Belatreche, N., Bouabsa, W., Attouch, M. K., Fetitah, O. (2022). M-Regression Estimation with the  $k$  Nearest Neighbors Smoothing under Quasi-associated Data in Functional Statistics. *Applications and Applied Mathematics : An International Journal (AAM)*. Vol. 17, No 2, 333 – 365.
- [20] Belatreche N., Bouabsa W., Attouch M. K. The converge rate of the hazard function with function explanatory variable : case of spacial data with the  $k$  Nearest Neighbor method. (2023). *Turkish Journal of Computer and Mathematics Education*. Vol. 14, No 2 , 180 – 202.

## BIBLIOGRAPHIE

---

- [21] Benko, M., Härdle, W., Kneip, A.(2005). Common Functional Component Modeling. CASE, Humboldt-Universität zu Berlin and Dep. of Statistics, Universität Bonn.
- [22] Berlinet, A., Biau, G., Rouvière, L. (2005). Optimal  $L_1$  bandwidth selection for variable kernel density estimates, *Statistics and Probability Letters*. Vol.74, 116-128.
- [23] Beirlant, J., Berlinet, A. and Biau, G. (2008). Higher order estimation at Lebesgue points, *Ann. Inst. Statist. Math.*, Vol. 90, No. 60, 651–677.
- [24] Biau, G. Cadre, B. (2004). Nonparametric spatial prediction, *Stat. Inference Stoch. Process.* Vol. 7 , 327-349.
- [25] Boente, G. and Fraiman, R. (1989). Robust nonparametric regression estimation, *J. Multivar. Anal.* Vol. 29, No. 2, 180-198.
- [26] Boente, G. and Fraiman, R. (1990). Asymptotic distribution of robust estimators for nonparametric models from mixing processes, *Ann. Stat.* Vol. 2, No. 18, 891-906.
- [27] Boente, G., Ganzalez-Manteiga, W. Pérez-Gonzalez, A. (2009). Robust nonparametric estimation with missing data, *J. Stat. Plann. Inference*. Vol. 139, No. 2, 571-592.
- [28] Bosq, D. (1991). Modelization, Nonparametric Estimation and Prediction for Continuous Time Processes. Part of the NATO ASI Series book series (ASIC). vol.335, 509-529.
- [29] Bosq, D. (2000). Linear Processes in Function Spaces : Theory and Applications, *Lecture Notes in Statistics* . Springer. 149.
- [30] Bouabssa, W. (2021). Nonparametric relative error estimation via functional regressor by the k nearest neighbors smoothing under truncation random, *Applications and Applied Mathematics : An International Journal (AAM)*. Vol. 16, No. 1, 97-116.
- [31] Bulinski, A. and Suquet, C. (2001). Normal approximation for quasi-associated random fields, *Statistics & Probability Letters*, Vol. 54, No. 2, 215-226.
- [32] Burba, F., Ferraty, F. and Vieu, P.(2008). Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique. *C. R. Math. Acad. Sci. Paris*. 346, no. 5-6, 339-342.
- [33] Burba, F., Ferraty, F. Vieu, P. (2009). k-nearest neighbour method in functional non-parametric regression, *J. Nonparametr. Stat.* Vol. 21, No. 4, 453-469.
- [34] Cai, Z.W. and Roussas, G.G. (1992). Uniform strong estimation under  $\alpha$ -mixing, with rates, *Statist. Probab. Lett.* Vol. 15, No. 6, 47-55.

## BIBLIOGRAPHIE

---

- [35] Carbon, M., Francq, C., Tran, L.T. (2007). Kernel regression estimation for random fields. *J. Statist. Plann. Inference*. Vol.137, 778-798.
- [36] Carbon, M., Tran, L.T., Wu, B.(1997). Kernel density estimation for random fields, *Statist. Probab. Lett.* Vol.36 , 115-125.
- [37] Cardot, H., Cénac, P., Monnez, J. (2012). A fast and recursive algorithm for clustering large datasets with k-medians. *Computational Statistics & Data Analysis*. Vol. 56, 1434-1449.
- [38] Cardot, H., Ferraty, F., Sarda, P. (1999). Functional linear model. *Statist.Probab. Lett.* Vol.45, 11-22.
- [39] Cérou, F., and Guyader, A. (2005). Nearest neighbor classification in infinite dimension. *ESAIM : P&S* 10, 340-255.
- [40] Chiou, J-M., Müller, H.-G., Wang, J-L., Carey, J.R. (2003). A functional multiplicative effects model for longitudinal data, with application to reproductive histories of female medies. *Statistica Sinica*. Vol.13, 1119- 1133.
- [41] Chiou, J.M., Müller, H.G. and Wang, J.L. (2003a). Functional quasi-likelihood regression models with smooth random effects. *J. Royal Statist. Soc.* Vol.65, 405-423.
- [42] Collomb, G. (1981). Estimation non paramétrique de la régression : Revue bibliographique, *Internat. Statist. Rev.* Vol. 18, No. 49, 75-93.
- [43] Collomb, G. Hardle, W. (1986). Strong uniform convergence rates in robust non-parametric time series analysis and prediction : Kernel regression estimation from dependent observations, *Stoch. Process. Appl.* Vol. 23, No. 1, 77-89.
- [44] Collomb, G. (1980). Estiamation de la regression par la methode des k plus proches avec noyau : quelques propriétés de convergence ponctuelle. *Statistique Non Parametrique Asymptotique*, J-P. Raoult, Ed. vol.821, 159–175.
- [45] Cover, T.M. (1968). Estimation by the nearest neighbor rule, *IEEE Trans. Inform. Theory*. Vol. 7, No. 14, 50-55.
- [46] Crambes, C., Delsol, L. Laksaci, A. (2008). Robust nonparametric estimation for functional data, *J. Nonparameter. Stat.*, Vol. 20, No. 7, 573-598.
- [47] Cressie, N.A. *Statistics for Spatial Data*. Wiley, New York, 1993.
- [48] Dabo-Niang, S., Thiam, B. (2010). Robust quantile estimation and prediction for spatial processes. *Stat Probab Lett.* Vol.80, 1447–1458

## BIBLIOGRAPHIE

---

- [49] Dabo-Niang, S., Yao, A.F. (2007) Kernel regression estimation for continuous spatial processes, *Math. Methods Statist.* Vol. 16, No. 15, 298-317.
- [50] Dabo-Niang, S., Ould-Abdi, S., Ould-Abdi, A., A. Diop. (2014). Consistency of a non-parametric conditional mode estimator for random fields, *Stat Methods Appl.* Vol. 16, No. 23, 1–39.
- [51] Daoudi, H., Mechab, B., Z. Chikr Elmezouar. (2018). Asymptotic normality of a conditional hazard function estimate in the single index for quasi-associated data. Vol. 4, No. 22, 513-530
- [52] Dauxois, J., Pousse, A., Romain, Y. (1982) . Asymptotic theory for the principal component analysis of a vector random function : some applications to statistical inference, *Journal of multivariate analysis*, Elsevier.
- [53] Dedecker, J., Doukhan, P., Lang, G., León, R.J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence : With Examples and Applications*, Lecture Notes in Statistics, Springer, New York.
- [54] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2013). Functional data : local linear estimation of the conditional density and its application. *Statistics*, 47(1), 26–44.
- [55] Dereich, S. (2003). High resolution coding of stochastic processes and small ball probabilities. PhD Thesis.
- [56] Derrar, S., Laksaci, A. and Ould Said, E. (2020). M-estimation of the regression function under random left truncation and functional time series model, *Statistical Papers*. Vol. 61, No. 3, 1181-1202.
- [57] Deville, J.C. (1974). Méthodes statistiques et numériques de l'analyse harmonique. *Ann. Insee.* 15, 3-101.
- [58] Devroye, L., Györfi, L., Krzyzak, A., Lugosi, G. (1994). On the strong universal consistency of nearest neighbor regression function estimates, *Ann. Statist.* Vol. 20, No. 22, 1371-1385.
- [59] Devroye, L., Wagner, T. (1977). The strong uniform consistency of nearest neighbor density, *Ann. Statist.* Vol. 10, No. 5, 536-540.
- [60] Devroye, L., Wagner, T. (1982). Nearest neighbor methods in discrimination. In *Classification, pattern recognition and reduction of dimensionality, Handbook of Statistics*. North-Holland, Amsterdam. Vol. 2.

## BIBLIOGRAPHIE

---

- [61] Devroye, L.P. (1978). The uniform convergence of nearest neighbour regression function estimators and their application in optimization. *IEEE Trans. Inform. Theory*. Vol.24, 142-151.
- [62] Dony, J., Einmahl, U. (2009). Uniform Bandwidth Consistency of Kernel Regression Estimators at a Fixed Point. *High Dimensional Probability. V : The Luminy*. Vol.53, No.5, 308-325.
- [63] Doob, J.L. (1953). *Stochastic Processes*, Wiley, New York.
- [64] Douge, L.(2010). Théorèmes limites pour des variables quasi-associées hilbertiennes, *Ann. Inst. Stat. Univ, Paris*. Vol.54, No.8, 51-60.
- [65] Doukhan, P., Lang, G., Surgailis, D., Tyssière, G. (2010). *Dependence in Probability and Statistics. Lecture Notes in Statistics*, Springer, Berlin.
- [66] Douge, L. (2010). Théorèmes limites pour des variables quasi-associées hilbertiennes. *Ann. I.S.U.P.*, Vol.54(1-2), 51-60.
- [67] Douge, L. (2018). Nonparametric regression estimation for quasi-associated hilbertian processes.
- [68] Einmahl,U. D.Mason, (2005). Uniform in bandwidth consistency of kernel-type function estimators, *The Annals of Statistics* , Vol. 33, No. 3, 1380-403.
- [69] El Machkouri, M. , Stoica, R. (2010). Asymptotic normality of kernel estimates in a regression model for random fields, *J Nonparametric Stat.* Vol. 22, No. 15, 955–971
- [70] Estévez-Pérez, G., Quintela-del-Ro, A. Vieu, P. (2002). Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples, *J. Statist. Plann. Inference*.Vol. 16, No. 16, 1-30.
- [71] Esary, J., Proschan, F. and Walkup, D. (1967). Association of random variables with applications, *Ann. Math. Stat.* Vol. 38, No. 5, 1466-1476.
- [72] Ezzahrioui, M., Ould-Saïd, E. (2008). Asymptotic normality of nonparametric estimator of the conditional mode for functional data. *Journal of Nonparametric Statistics*. 20, 3-18.
- [73] Fan, J. (1992). Design adaptive nonparametric regression. *J. of the American Statistician Association*, Vol.87, 998–1004.
- [74] Fan, J., Gijbels, I. (1996). *Local polynomial modelling and its applications.*, Chapman & Hall/CRC. Vol.66.

## BIBLIOGRAPHIE

---

- [75] Fan, T.C. Hu, Y.K. Truong. (1994). Robust nonparametric function estimation. *Scand. J. Statist.*, Vol.21 ,433-446.
- [76] Febrero, M., Galeano, P., González-Manteiga, W. (2007). A functional analysis of NOx levels : location and scale estimation and outlier detection. *Computational Statistics*. Vol.22, 411–427
- [77] Ferraty, F., Vieu, P. (2003). Curves discrimination : a nonparametric functional approach. *Computational Statistics & Data Analysis*. Vol.44, 161-173.
- [78] Ferraty, F., Vieu, P. (2004). Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination. *Nonparametric Statistics. The International Conference on Recent Trends and Directions in Nonparametric Statistics*. Vol.16, 111-125.
- [79] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis. Theory and practice*, New York : Springer-Verlag.
- [80] Ferraty, F. Vieu, P. (2006a). *Nonparametric modelling for functional data*. Springer-Verlag, New York.
- [81] Ferraty, F. et Vieu, P. (2006b). *Functional nonparametric statistics in action. The art of semiparametr.*
- [82] Ferraty, F., Mas, A. et Vieu, P. (2007). Nonparametric Regression on Functional Data : Inference and Practical Aspects. *Aust. NZ J. Stat.* Vol.49, 267-286.
- [83] Ferraty, F., Laksaci, A., Vieu, P. (2002). Conditional distribution in nonparametric functional models : theoretical and practical advances.
- [84] Ferraty, F., Laksaci, A., Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Stat. Inference Stoch. Process.* Vol. 9, No. 20, 47–76.
- [85] Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Sankhy a*, Vol.67(2), 378-398.
- [86] Ferraty, F., Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Acad. Sci.,Paris*. Vol.330, 139-142.
- [87] Ferraty, F., Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist. And Data Anal.* Vol.4, 545-564.
- [88] Ferraty, F., Peuch, A., Vieu, P. (2003). Modèle à indice fonctionnel simple. *C. R. A. S.Mathématiques Paris*. Vol.336(12), 1025-1028.

## BIBLIOGRAPHIE

---

- [89] Ferraty, F., Rabhi, A., P.Vieu. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle, *Rev. Roumaine Math. Pures Appl.* Vol. 53, No. 15, 1–18.
- [90] Ferraty, F. Tadj, A. Laksaci, A. and P.Vieu. (2010). Rate of uniform consistency for nonparametric estimates with functional variables, *J. Statist. Plann. Inference.* Vol. 140, No. 15, 335-352.
- [91] Florens, J. P., Larribeau, S., Mouchart, M. (1994). Bayesian Encompassing Tests of a Unit Root Hypothesis. *Econometric Theory.* Vol.10 , 747-763.
- [92] Frank, I.E., Friedman, J.H. (1993). A statistical view of some chemometrics regression tools. *Technometrics.* Vol.35, 109-135.
- [93] Gagui, A., Chouaf, A. On the nonparametric estimation of the conditional hazard estimator in a single functional index, *STATISTICS IN TRANSITION new series.* Vol.23, No.2, 89–105.
- [94] Gao, F., Li, W.V., (2007). Small ball probabilities for the Slepian Gaussian fields. *Trans. Amer. Math. Soc.* Vol.359 (3), 1339-1350 (electronic).
- [95] Gheriballah A, Laksaci A, Rouane, R. (2010). Robust nonparametric estimation for spatial regression. *J Stat Plan Inference.* Vol.140, 1656–1670.
- [96] Gheriballah, A., Laksaci, A., Sekkal, S. (2013). Nonparametric M-regression for functional ergodic data, *Statistics & Probability Letters.* Vol. 83, No. 6, 902-908.
- [97] Grenander, U. (1981). *Abstract Inference*, Wiley, New York.
- [98] Guyon, X. (1987). Estimation d'un champ par pseudo-vraisemblance conditionnelle : Etude asymptotique et application au cas Markovien, *Proceedings of the Sixth Franco-Belgian Meeting of Statisticians.*
- [99] Györfi, L., Härdle, W., Sarda, P. and Vieu, P. (1989). *Nonparametric Curve Estimation from Time Series*, Lecture Notes in Statistics, Springer, New York. Vol. 60.
- [100] Györfi, L., Kohler, M., Krzyzak, A., Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*, Springer, New York.
- [101] Hall, P., Vial, C. (2006) Assessing the finite dimensionality of functional data. *Journal of the Royal Statistical Society : Series B (Statistical Methodology).* vol.68, 689-705.
- [102] Hallin, M., Lu, Z., Yu, K. (2009). Local linear spatial quantile regression. *Bernoulli.* Vol.15, 659–686.

## BIBLIOGRAPHIE

---

- [103] Hampel, F. R. (1971). A general qualitative definition of robustness. *Ann. Math. Stat.* Vol.42, 1887-1896.
- [104] Härdle, W. Vieu, P. (1992). Kernel regression smoothing of time series, *Journal of Time Series analysis*, Vol. 13, No.3, 209-232.
- [105] Härdle, W. Tsybakov, A. (1988). Robust nonparametric regression with simultaneous scale curve estimation, *Ann. Stat.* Vol. 16, No. 1, 120-135.
- [106] Härdle, W., Marron, J.S. (1985). Optimal Bandwidth Selection in Nonparametric Regression Function Estimation, *Annals of Statistics*. Vol. 13, No. 15, 1465-1481.
- [107] Hastie, T., Mallows, C. (1993). Discussion of "A statistical view of some chemometrics regression tools." by Frank, I.E. and Friedman, J.H. *Technometrics*. Vol.35, 140-143.
- [108] Hastie, T., Buja, A. Tibshirani, R. (1995). Penalized discriminant analysis. *Annals of Statistics*. Vol.23, 73-102.
- [109] Holmstrom, I. (1963). On a method for parametric representation of the state of the atmosphere. *Tellus*. Vol.15, 127-149.
- [110] Huber, P.J. (1973). Robust regression. Asymptotics, conjonctures and monte Carlo. *Ann. Stat.* Vol.43, 799-821.
- [111] Huber, P.J. (1981). *Robust statistics*. John Wiley.
- [112] Huber, P. (1964). Robust estimation of a location parameter, *Ann. Math. Stat.* Vol. 35, No. 1, 73-101.
- [113] Jong-Dev, K., Proschan, F. (1983). Negative association of random variables, with applications, *Ann. Stat.* Vol. 11, No. 1, 286-295.
- [114] Kallabis, R.S., Neumann, M.H. (2006). An exponential inequality under weak dependence, *Bernoulli*. Vol. 12, No. 2, 333-350.
- [115] Kara, Z., Laksaci, A., Vieu, P. (2017). Data-driven kNN estimation in nonparametric functional data analysis, *Journal of Multivariate Analysis*, Vol. 153, No. 85, 176-188.
- [116] Kirkpatrick, M. Heckman, N. (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* Vol.27, No.4, 429-450.
- [117] Krasker, W. S., Welsch, R. E. (1982). Efficient bounded influence regression estimation. *J. A. S. A.* Vol.77, 595-604.

## BIBLIOGRAPHIE

---

- [118] Kneip, A., Utikal, K. (2001). Inference for density families using functional principal component analysis. *Journal of the American Statistical Association*, Taylor & Francis. 519-542.
- [119] Kudraszow, N.L., Vieu, P. (2013). Uniform consistency of kNN regressors for functional variables, *Statistics & Probability Letters*. Vol. 83, No. 153, 1863-1870.
- [120] Laïb, N. and Louani, D. (2010), Nonparametric kernel regression estimation for functional stationary ergodic data : asymptotic properties. *Journal of Multivariate Analysis*. Vol.101, 2266- 2281.
- [121] Laïb, N. and Louani, D. (2011), Rates of strong consistencies of the regression function estimator for functional stationary ergodic data, *Journal of Statistical Planning and Inference*. Vol.141, 359-372.
- [122] Laïb, N., Ould-Saïd, E. (2000). A robust nonparametric estimation of the autoregression function under ergodic hypothesis. *Canad. J. Statist.* Vol.28, 817-828.
- [123] Laksaci, A. (2005). Contribution aux modèles non paramétriques conditionnels pour variables explicatives fonctionnels. Thèse de doctorat, université de Toulouse.
- [124] Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Annales de l'ISUP*. Vol.51, 69-80
- [125] Laksaci, A. Meref, F. (2009). Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes C. R. Acad. Sci. Paris, Ser. I. Vol.347, 1075-1080 .
- [126] Laksaci, A. Mechab, B. (2010). Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales, *REV. ROUMAINE MATH. PURES APPL.* Vol. 55, No. 1, 35–51.
- [127] Laksaci, A., Mechab, B. (2014). Conditional hazard estimate for functional random fields. *Journal of Statistical Theory and Practice*. Vol.8, 192-200.
- [128] Laloë, T. (2008). A k-nearest neighbor approach for functional regression. *Statistics & Probability Letters*. Vol.78, 10, 1189-1193.
- [129] Laloë, T. (2008). A k-nearest neighbor approach for functional regression, *Statistics & Probability Letters* . Vol. 78, No. 10, 1189-1193.
- [130] Lancaster, T. (1990). *The econometric analysis of transition data*. Cambridge university press. Issue.17.

## BIBLIOGRAPHIE

---

- [131] Laurini, M. P. (2014). Dynamic functional data analysis with non-parametric state space models. *Applications Statistics*. Vol.41, 142-163.
- [132] Lecoutre, J.P. Ould-Said, E. (1992). Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure, *C.R. Math. Acad. Sci. Paris*. Vol. 314, No. 15, 295-300.
- [133] Lian, H. (2011). Convergence of Functional k-Nearest Neighbor Regression Estimate with Functional Responses, *Electronic Journal of Statistics*. Vol. 5, No. 133, 31-40.
- [134] Li, J.P. (1985). Strong convergence rates of error probability estimation in the nearest neighbor discrimination rule, *J. Math. (Wuhan)*. Vol. 15, No. 5, 113-118.
- [135] Li, J., Tran, L.T. (2007). Hazard rate estimation on random fields, *J. Multivariate Anal.* Vol. 98, 1337-1355.
- [136] Li, J., Tran, L.T. (2009). Nonparametric estimation of conditional expectation, *J. Statist. Plann. Inference*. Vol. 139, No. 36, 164-175.
- [137] Li, J. Tran, L.T. (2009). Nonparametric estimation of conditional expectation, *J Stat Plan Inference*. Vol. 139, No. 35, 164–175.
- [138] Li, W.V., Shao, Q.M. (2001). Gaussian processes : inequalities, small ball probabilities and applications. In *Stochastic processes : Theory and Methods*, Ed. C.R. Rao and D. Shanbhag. *Hanbook of Statistics*, 19, North-Holland, Amsterdam.
- [139] Lian, H. (2011). Convergence of functional k-nearest neighbor regression estimate with functional responses. *Electronic Journal of Statistics*. Vol.5, 31-40.
- [140] Lifshits, M.A., et Simon, T. (2005). Small deviations for fractional stable processes. *Ann. Inst. H. Poincaré Probab. Statist.* Vol.41, 725-752.
- [141] Lifshits, M.A., Linde, W., Shi, Z. (2006). Small deviations of Riemann- Liouville processes in  $L_q$ -spaces with respect to fractal measures. *Proc. London Math. Soc.* (3) . Vol.92 (1), 224-250.
- [142] Lu, Z., and Cheng, P. (1998). Strong consistency of nearest neighbor kernel regression estimation for stationary dependent samples. *Science in China Series A : Mathematics* 41. Vol.977, 918-926.
- [143] Lu, Z. Chen, X. (2004). Spatial kernel regression : weak consistency, *Stat Probab Lett.* Vol. 68, No. 30, 125–136.
- [144] Mack, Y. P. (1981). Local properties of k-NN regression estimates. *SIAM J. Algebr Discrete Methods*. Vol.2 , 311-323.

## BIBLIOGRAPHIE

---

- [145] Massim, I. and Mechab, B. (2016). Local linear estimation of the conditional hazard function. *International Journal of Statistics and Economics*. Vol.17, 1-11.
- [146] Masry, E. (2002). Multivariate probability density estimation for associated processes : Strong consistency and rates, *Stat. Probab. Lett*, Vol. 58, No. 2, 205-219.
- [147] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : asymptotic normality. *Stochastic Process. Appl.* Vol.115, No.1, 155-177.
- [148] Masry, E. (1986). Recursive probability density estimation for weakly dependent stationary processes, *IEEE Trans. Inform. Theory*. Vol. 32 , No. 16, 254-267.
- [149] Matula, P. (1992). A note on the almost sure convergence of sums of negatively dependent random variables, *Statistics & Probability Letters*. Vol. 15, No. 3, 209-213.
- [150] Mebsout, M., Attouch, M.K., Fetitah, O. (2020). Nonparametric M-regression with scale parameter for functional dependent data, *Applications and Applied Mathematics : An International Journal (AAM)*. Vol. 15, No. 2, 846-874.
- [151] Mechab, W., Laksaci, A. (2016). Nonparametric relative regression for associated random variables. *Metron*. Vol.74(1), 75-97.
- [152] Molennar, P., Boomsma, D. (1987). The genetic analysis of repeated measures ii. The KarhunenLoeve expansion. *Behavior Genetic*. Vol.17, 229-242.
- [153] Moore, D., Yackel, J. (1977). Consistency properties of nearest neighbor density function estimators, *Ann. Statist.* Vol. 20, No. 5, 143-154.
- [154] Neiros-Perez, G., Cardot, H., Estevez-Perez, G., Vieu, P. (2004). Maximum ozone concentration forecasting by functional nonparametric approaches. *Environmetrics*. Vol.15, 675-685.
- [155] Obukhov. A.(1960). The statistically orthogonal expansion of empirical functions. *American Geophysical Union*, 288-291.
- [156] Parzen, E. (1962). On Estimation of a Probability Density Function and Mode, *Institute of Mathematical Statistics*. Vol.33, No.3 , 1065-1076
- [157] Poskitt, DS., Hall, P., Presnell, B. 2001. Technometrics, Specification of echelon-form VARMA models. H Lütkepohl, DS Poskitt. Vol.43 (1), 1-9.
- [158] Presnell, Brett, and Hall, Peter. (1998). Applications of intentionally biased bootstrap methods. *Documenta Mathematica* : 257-266.
- [159] Quintela-del-Ro, A. (2006). Nonparametric estimation of the maximum hazard under dependence conditions, *Statist. Probab. Lett*. Vol. 76, No. 15, 1117-1124.

## BIBLIOGRAPHIE

---

- [160] Quintela-del-Ro, A. (2008). Hazard function given a functional variable : Non-parametric estimation under strong mixing conditions, *J. Nonparametr. Stat.* Vol. 20, No. 15, 413-430.
- [161] Rabhi, A., Benaissa, S., Hamel, E. and Mechab, B. (2013). Mean square error of the estimator of the conditional hazard function. *Appl. Math. (Warsaw)*. Vol.4, 405-420.
- [162] Rachdi, M., Vieu, P. (2007). Nonparametric regression for functional data : automatic smoothing parameter selection. Vol. 13, No. 137, 2784-2801.
- [163] Ramsay, J. (1982). When the data are functions, *Psychometrika*. Vol.47, 379–396.
- [164] Ramsay, J., Silverman, B. (1997). *Functional Data Analysis*, Springer- Verlag, New York.
- [165] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis (Second Edition)*. SpringerVerlag, New York.
- [166] Ramsay, J. (2000). Functional components of variation in handwriting. *Journal of the American Statistical Association*. Vol.95, 9-15.
- [167] Ramsay, J.O., Silverman, B.W. (2002). *Applied Functional Data Analysis ; Methods and Case Studies*, Springer-Verlag. New York.
- [168] Ramsay, J.O., Silverman, B.W. (1997). *Functional data analysis 1997* New York, Springer.
- [169] Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrika*. Vol.51, 83-90.
- [170] Robinson, P.M. (2011). Asymptotic theory for nonparametric regression with spatial data, *J Econom.* Vol. 165, No. 15, 5–19.
- [171] Robinson, P.M. (1983). *Robust Nonparametric Autoregression*, In : *Robust and Nonlinear Time Series Analysis*, Lectures Note in Statistics, Springer, New York.
- [172] Robinson, P.M. (1984). Robust nonparametric regression. In : Franke J., Hardle W. and Martin D. (Eds), *Robust and Nonlinear Time Series Analysis*, Lecture Notes in Statistics. Springer, Berlin. Vol.26, 247-255.
- [173] Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. *Institute of Mathematical Statistics*. Vol. 27, No. 3, 832-837
- [174] Roussas, G. (1989). Hazard rate estimation under dependence conditions, *J. Statist. Plann. Inference*. Vol. 22, No. 15, 81-93.

## BIBLIOGRAPHIE

---

- [175] Roussas, G. (1989). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann.* Vol.18, 135–149.
- [176] Roussas, G. (1991). Kernel estimates under association : strong uniform consistency, *Statistics & Probability Letters.* Vol. 12, 215-224.
- [177] Roussas, G. (2000). *Prediction Under Association.* In : *Limnios, N., Nikulin, M.S. (eds.) Recent Advances in Reliability Theory : Methodology, Practice and Inference.* Birkhäuser, Boston.
- [178] Royall. (1966), R.M. A class of nonparametric estimates of a smooth regression function, Ph.D. Diss., Stanford University.
- [179] Royall, R.M. A class of nonparametric estimates of a smooth regression function, Ph.D. Diss., Stanford University.
- [180] Sauder, C., Cardot, H., Disenhaus, C. et 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. *Agr. Sci.* Vol.151, 578-589.
- [181] Stone, C. J. (1977). Consistent nonparametric regression. *Ann. Statist.* Vol.5, 595-645.
- [182] Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* Vol.10.
- [183] Tabti, H., Ait Saïdi, A. (2018). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure. *Commu. Stat. Theory and Methods.* Vol.47, 816-838.
- [184] Torkia, M., Mechab, B., Massim, I. (2018). Quadratic error of the conditional hazard function in the local linear estimation for functional data. *Afr. Stat.* Vol. 3, No. 13, 1759-1777.
- [185] Tran, L. T., Wehrens, R., Buydens, L. (2006). kNN-kernel density-based clustering for high-dimensional multivariate data, *Comput. Statist. Data Anal.* Vol. 51, No. 81, 513-525.
- [186] Tran, L. T. (1990). Kernel density estimation on random fields, *J. Multivariate Anal.* Vol. 34, No. 25, 37-53.
- [187] Tran, L. T. (1993). Nonparametric Estimation for Time Series by Local Average Estimators. *The Annals of Statistics.* Vol.42, No. 2, 1040-1057.

- [188] Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*. Vol.23, 19-23.
- [189] Tukey, J.W. (1961). Curves as parameters, and touch estimation, Proc. 4th Berkeley Sympos. Math. Statist. and Prob.
- [190] Watson, G.S., Leadbetter, M.R. (1964). Hazard analysis, *Sankhya*. Vol. 26 , No. 25, 101-116.
- [191] Xu, R., Wang, J. (2008).  $L^1$ - estimation for spatial nonparametric regression. *J Nonparametric Stat.* Vol.20, 523–537.

## Abstract

Robust estimation presents an alternative method to classical regression methods, which is particularly relevant when observations are affected by the presence of outliers. We will focus more particularly in this thesis on the estimation of some robust functions in the case where the observations are of possibly infinite dimension using the  $k$  nearest neighbors method. In which the response variable is real while the explanatory variable is functional. The non-parametric estimation by the method of  $k$  nearest neighbors, offers the particularity of being relatively insensitive to large deviations due to a few observations of atypical values. We use for this a selection criterion for the choice of the optimal parameters in order to improve the quality of the forecast, by adding the scale parameter which will be unknown and by proposing the optimality by the method of  $k$  nearest neighbors, these two parameters together make it possible to control the bias of the estimator and thus improve the quality of the estimate.

First, we consider a sequence of observations of identically distributed strictly stationary dependent random variables. In this context, we construct an estimator of the regression function by relating the M-estimating approach with the quasi-association setting using the kNN method, and we study the asymptotic properties of this estimator.

Secondly, we developed a new hazard estimator in cases where the co-variables are functional in nature. This estimator combines spatial functional data with the  $k$  nearest neighbors method. under some general assumptions, we state almost complete convergence with the rate of convergence. By way of illustration, our result is applied to the discrimination of curves, to forecasting problems, and to the construction of confidence intervals. Our research focuses on the phenomenon of concentration of the probability measure of the functional variable on small balls. This concentration measure propose, under certain hypotheses, a solution to the problem of the curse of dimension as well as the inexistence of a probability density .

We can consider that the study presented in this thesis is a contribution to the extension of the work in progress in an unlimited dimension both in theoretical and practical aspects.

## Résumé

L'estimation robuste présente une méthode alternative aux méthodes de régression classiques, qui est particulièrement pertinente quand observations sont affectées par la présence de données aberrantes. Nous nous intéresserons plus particulièrement dans cette thèse à l'estimation de quelques fonctions robustes dans le cas où les observations sont de dimension éventuellement infinie en utilisant la méthode des  $k$  plus proches voisins. Dans lesquels la variable réponse est réelle tandis que la variable explicative est fonctionnelle. L'estimation non-paramétrique par la méthode des  $k$  plus proches voisins, offre la particularité d'être relativement insensible aux large déviation dues à quelques observations de valeurs atypiques. Nous utilisons pour cela un critère de sélection pour le choix des paramètres optimales afin d'améliorer la qualité de la prévision, en ajoutant le paramètre d'échelle qui sera inconnu et en proposant l'optimalité par la méthode des  $k$  plus proches voisins.

Dans un premier temps, nous considérons une suite d'observations de variables aléatoires dépendantes strictement stationnaires distribuées de manière identique. Dans ce contexte, nous construisons un estimateur de la fonction de régression en reliant l'approche de M-estimation avec le cadre de quasi-association à l'aide de la méthode kNN, et nous étudions les propriétés asymptotiques de cet estimateur. Dans un second temps, nous avons développé un nouvel estimateur de hasard dans les cas où les co-variables sont de nature fonctionnelle, cet estimateur combine des données fonctionnelles spatiales avec la méthode des  $k$  plus proches voisins. Sous certaines hypothèses générales, nous énonçons la convergence presque complète avec le taux de convergence. A titre d'illustration, notre résultat est appliqué à la discrimination des courbes, aux problèmes de la prévision. Nos recherches portent sur le phénomène de concentration de la mesure de probabilité de la variable fonctionnelle sur des petites boules. Cette mesure de concentration propose, sous certaines hypothèses, une solution au problème de la malédiction de dimension ainsi que de l'inexistence d'une densité de probabilité.

Nous pouvons considérer que l'étude présentée dans cette thèse est une contribution à l'extension des travaux en cours dans une dimension illimitée tant dans les aspects théoriques que pratiques.

## ملخص

يقدم التقدير القوي طريقة بديلة لطرق الانحدار الكلاسيكية، والتي تكون ذات صلة بشكل خاص عندما تتأثر الملاحظات بوجود القيم المتطرفة. سنركز بشكل خاص في هذه الأطروحة على تقدير بعض الوظائف القوية في الحالة التي تكون فيها الملاحظات ذات أبعاد لا نهائية باستخدام طريقة  $k$  الأقرب للجيران. حيث يكون متغير الاستجابة حقيقياً بينما يكون المتغير التوضيحي وظيفياً. التقدير غير البارامترى بطريقة  $k$  الأقرب للجيران، يقدم خصوصية كونه غير حساس نسبياً للانحرافات الكبيرة بسبب ملاحظات قليلة للقيم غير النمطية. نستخدم لهذا معيار اختيار لاختيار العلامات المثلى من أجل تحسين جودة التنبؤ، عن طريق إضافة معلمة المقياس التي ستكون غير معروفة وباقتراح الأمثل بطريقة  $k$  أقرب جيران، هذين الاثني معا تجعل من الممكن التحكم في تحيز المقدّر وبالتالي تحسين جودة التقدير. أولاً، نأخذ في الاعتبار سلسلة من الملاحظات لمتغيرات عشوائية تابعة ثابتة بشكل متتابع وموزعة بشكل متماثل. في هذا السياق، نقوم ببناء مقدر لوظيفة الانحدار من خلال ربط نهج  $M$ -تقدير مع اعدادات شبه الارتباط باستخدام طريقة  $kNN$ ، وندرس الخصائص المقاربة لهذا المقدّر. ثانياً، قمنا بتطوير مُقدّر جديد للمخاطر في الحالات التي تكون فيها المتغيرات المشتركة وظيفية بطبيعتها. يجمع هذا المقدّر بين البيانات الوظيفية المكانية وطريقة  $k$  الجيران الأقرب. في ظل بعض الافتراضات العامة، نعلن التقارب الكامل تقريباً مع معدل التقارب. على سبيل التوضيح، يتم تطبيق نتيجتنا على تمييز المنحنيات، والتنبؤ بالمشكلات، وبناء فترات الثقة. يركز بحثنا على ظاهرة تركيز المقياس الاحتمالي للمتغير الوظيفي على الكرات الصغيرة. يقترح مقياس التركيز هذا، في ظل فرضيات معينة، حلاً لمشكلة لعنة البعد بالإضافة إلى عدم وجود كثافة احتمالية. يمكننا أن نعتبر أن الدراسة المقدمة في هذه الأطروحة هي مساهمة في توسيع العمل الحالي في البعد غير المحدود من الناحيتين النظرية والعملية.