

**THESE DE DOCTORAT DE L'UNIVERSITE DU LITTORAL CÔTE
D'OPALE**

Spécialité : Mathématiques Appliquées
Option : Statistiques Mathématiques

Présentée par

Mohammed Kadi ATTOUCH

Pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DU LITTORAL CÔTE D'OPALE

Sujet de la thèse :

**Estimation robuste de la fonction de régression
pour des variables fonctionnelles.**

Soutenue le 29 Janvier 2009 devant le jury composé de :

Benoit CADRE	ENS-Cachan, Bretagne-Sud	Examineur
Ali LAKSACI	Université Sidi Bel Abbès	Directeur de thèse
Djamel LOUANI	Université Reims-Champagne-Ardenne	Examineur
Wenceslao González MANTEIGA	Université Santiago de Compostelle	Rapporteur
André MAS	Université de Montpellier 2	Rapporteur
Elias OULD-SAID	Université Littoral Côte d'Opale	Directeur de thèse
Cristian PREDA	Ecole Polytechnique de Lille	Examineur
Dominique SCHNEIDER	Université Littoral Côte d'Opale	Examineur



2

A ma Femme

BIBLIOTHEQUE DU CERIST

A ma Famille

Cette thèse à été soutenue avec un avis favorable de la commission des thèses de la fédération de mathématiques de la région Nord Pas de Calais

Remerciements

Je tiens à exprimer toute ma gratitude à mes deux directeurs de thèse Monsieur Elias Ould-Saïd et Monsieur Ali Laksaci de m'avoir initié dans un domaine de la statistique que je ne connaissais pas, à savoir l'étude des variables fonctionnelles. Leurs conseils et leur expérience m'ont permis de mener à bien ce travail de thèse. Leurs qualités humaines et scientifiques m'ont permis d'aborder des voies peu explorées mais très motivantes.

Je tiens à exprimer ma sincère reconnaissance à Messieurs André Mas et à Wenceslao González Mantega qui ont accepté, malgré toutes leurs occupations, d'être les rapporteurs de ce manuscrit de thèse. Je leur suis gré de l'intérêt qu'ils ont porté à mon travail. Je souhaite les remercier pour leur lecture attentive de ce mémoire ainsi que pour leurs remarques et commentaires qui m'ont permis d'en améliorer le contenu.

Je suis très heureux que Monsieur Benoit Cadre ait accepté de faire parti de mon jury. Je tiens à le remercier pour l'intérêt qu'il a porté à mon travail.

Messieurs Djamel Louani, Cristian Preda et Dominique Schneider ont accepté de faire parti de mon jury de thèse, je tiens à les remercier vivement et je leur en suis très reconnaissant.

Mes remerciements s'adressent également aux membres des deux laboratoires de recherche qui m'ont permis de mener dans des conditions très favorables mes travaux de recherche : le laboratoire L.M.P.A. de l'université du littoral Côte d'Opale (France) et le laboratoire L.D.M. de l'université Djillali Liabès de Sidi Bel Abbès (Algérie).

En dernier lieu, je tiens à remercier vivement toute ma famille pour leur soutien constant. Mes premières pensées vont à ma femme qui a su m'apporter courage et confiance en moi, même lorsque l'avenir paraissait incertain. Je tiens à exprimer toute ma gratitude à mes parents, mes frères et sœurs pour leurs soutiens durant ces années. Je tiens enfin à remercier tous les membres de ma belle famille qui m'ont apporté soutien et réconfort le long de ces années de préparation de ma thèse. Je dédie à vous tous ce modeste travail.

0.1 Résumé

La régression robuste est une analyse de régression ayant la capacité d'être relativement insensible aux larges déviations dues à certaines observations aberrantes. Dans ce cadre, on se propose dans cette thèse d'étudier l'estimation robuste de la fonction de régression, dans le cas où les observations sont de nature fonctionnelle.

Dans un premier temps, on considère une suite d'observations indépendantes identiquement distribuées. Dans ce contexte, nous établissons la normalité asymptotique d'une famille d'estimateurs robustes basée sur la méthode du noyau. A titre illustratif, notre résultat est appliqué à la discrimination des courbes, aux problèmes de la prévision, et à la construction des intervalles de confiance.

Dans un second temps, nous supposons que les observations sont fortement mélangeantes, et nous établissons la vitesse de convergence presque complète ponctuelle et uniforme de cette famille d'estimateurs, ainsi que la normalité asymptotique.

Notons, que les axes structurels du sujet, à savoir la "dimensionnalité" et la corrélation des observations, la "dimensionnalité" et la robustesse du modèle, sont bien exploités dans cette étude. De plus, la propriété de la concentration de la mesure de probabilité de la variable fonctionnelle dans des petites boules est utilisée, cette mesure de concentration permet sous certaines hypothèses de proposer une solution originale au problème du fléau de la dimension et ainsi généraliser les résultats déjà obtenus dans le cadre multivarié.

Pour illustrer l'extension et l'apport de notre travail, nous explicitons dans des exemples comment nos résultats peuvent être appliqués aux problèmes non standard de la statistique non-paramétrique tel que la prévision de série temporelles fonctionnelles. Nos méthodes sont appliquées à des données réelles telles que l'économie et l'astronomie.

0.2 Abstract

The robust regression is an analysis of regression with capacity to be relatively insensitive to the large deviations due to some outliers observations. Within this framework, one proposes in this thesis studied the robust estimate of the function of regression, if the observations are at the same time independent, strongly mixing and the covariate is functional.

Firstly, one considers a simple of identically distributed independent observations. In this context, we establish the asymptotic normality of robust nonparametric estimators for regression function weighted on the kernel method. We apply this result to the discrimination of the curves, the prediction problems, and to the construction of a confidence interval.

Secondly, we suppose that the observations are strongly mixing, and we establish the almost surely convergence rate (pointwise and uniform) of the studied model. The asymptotic normality in this correlation case is also established.

Let us note, that the structural axes of the subject, namely "dimensionality" and the correlation of the observations, "dimensionality" and the robustness of the model, are exploited in this study. Moreover, as all asymptotic properties in nonparametric functional statistics, the concentration property of the probability measure of the functional variable is include in the dispersion part of all asymptotic results given in our robust study.

To illustrate the extension and the contribution of our work, we show by some real data applications how our results can be applied to the nonstandard problems of the nonparametric statistics such as the forecast of functional time series.

0.3 Liste des travaux et communications

1. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. (2007). Strong uniform convergence rate of robust estimator of the regression function for functional and dependent processes. *Preprint, LMPA No. 340, Juillet 2007. Univ. du Littoral Côte d'Opale*. Submitted.
2. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. (2008a). Asymptotic distribution of robust estimator for functional nonparametric models. *Comm. Statist. Theory and Methods*, In press.
3. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. (2008b). Asymptotic normality of a robust estimator of the regression function for functional time series. *Preprint, LMPA No. 378, Janvier 2008. Univ. du Littoral Côte d'Opale*. Submitted.

Communication international :

1. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. Sur l'approche non paramétrique robuste en statistique fonctionnelle : Convergence et normalité asymptotique. Journées de Statistique Théorique et Appliquée, Biskra, les 21 et 22 Avril 2007.
2. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. Asymptotic distribution of robust estimator for functional nonparametric models. 5èmes Journées de Statistique Fonctionnelle et Operationnelle (STAPH), Lille, les 21 et 22 Juin 2007.
3. M. ATTOUCH, A. LAKSACI, E. OULD-SAÏD. Asymptotic normality of robust nonparametric estimator for functional dependent data. First International Workshop on Functional and Operatorial Statistics IWFOS'2008, University Paul Sabatier, Toulouse, 19-21 June 2008.

Table des matières

0.1	Résumé	1
0.2	Abstract	2
0.3	Liste des travaux et communications	3
1	Introduction	9
1.1	Statistique non paramétrique fonctionnelle : Motivation et context bibliographique . . .	9
1.2	Analyse non paramétrique robuste	11
1.3	Art de la statistique non paramétrique robuste pour des données fonctionnelles	12
1.4	Quelques exemples de données fonctionnelles contenant des valeurs aberrantes	12
1.5	Description de la thèse	18
1.6	Brève présentation des résultats	18
1.6.1	Résultats : Cas i.i.d.	19
1.6.2	Résultats : Cas α mélangeant	20
2	Normalité asymptotique : Cas i.i.d.	29
2.1	Introduction	30
2.2	Hypotheses and results	32
2.3	Application	36
2.3.1	Application to prediction	36

2.3.2	Conditional Confidence Interval	37
2.3.3	Application to curves discrimination	37
2.4	Some special cases	38
2.4.1	Vectorial case	38
2.4.2	Fractal dimension process case	38
2.4.3	Continuous time process case	39
2.4.4	Hilbertian variable case	40
2.5	Numerical study	41
2.6	Appendix	44
3	Convergence presque sur : cas α-mélange	53
4	Normalité asymptotique : Cas α-mélange	83
4.1	Introduction	84
4.2	Hypotheses and results	86
4.3	Main result	88
4.4	Application	90
4.4.1	Application to functional times series prediction	90
4.4.2	Conditional Confidence curve	91
4.5	A real data application	91
4.6	Appendix	93
5	Annexe 1 : Conclusions et commentaires	103
5.1	Choix des paramètres	103
5.1.1	La robustification	103
5.1.2	Le paramètre de lissage	105
5.1.3	La semi-métrie	106

5.2	Généralisation des résultats au cas non-robustes	107
5.2.1	La normalité asymptotique : cas i.i.d.	107
5.2.2	La convergence presque complète : cas de mélange	108
5.2.3	La normalité asymptotique : cas de mélange	111
5.3	Perspectives	114
6	Annexe 2 : Quelques outils de probabilités	115
6.1	Notations et définitions	115
6.2	Sur les outils de probabilité utilisés	116
6.2.1	Inégalités type exponentiel	116
6.2.2	Théorème central limite	117
6.3	Sur la structure topologique considérée	119