N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBÈS

THESE DE DOCTORAT DE 3ème CYCLE

Présentée par :

ALTENDJI Belkais

Domaine : Mathématiques Informatique

Filière : Mathématiques

Intitulé de la formation :Statistique, Mathématiques apliquées à l'économie et à la finance

Intitulée

Analyse non paramétrique spatio-fonctionnelle sur des données censurées

Soutenue le : …05/03/2019...

Devant le jury composé de :

Président :

Mr MECHAB Boubaker Maitre de Conférences A à L'Université S.B.A.

Examinateurs :

Mr BENAISSA Samir Professeur à L'Université S.B.A

Mr GUENDOUZI Toufik Professeur à L'Université de Saida

Mr BENCHIKH Tawfik Professeur à L'Université S.B.A

Directeur de thèse :

Mr LAKSACI Ali Professeur à L'Université S.B.A

Co-Directeur de thèse :

Mr RACHDI Mustapha Professeur à L'Université de Grenoble

Année universitaire : 2018/2019

Qu'est-ce qu'une idée ?
c'est une image qui se peint dans mon cerveau

Voltaire

REMERCIEMENTS

Soyons reconnaissants aux personnes qui nous donnent
du bonheur ; elles sont les charmants jardiniers
par qui nos âmes sont fleuries.
Marcel Proust
Le temps met tout
en lumière.
Thalès

Le remerciement est une affaire lucrative, vertueuse et rentable et le seul moyen de se délivrer d'une tentation, c'est d'y céder paraît-il ! Alors j'y cède en disant en grand Merci aux personnes qui ont cru en moi et qui m'ont permis d'arriver au bout de cette thèse.

En préambule, je tiens à remercier Dieu le tout puissant et miséricordieux, qui m'a donné la force et la patience d'accomplir ce modeste travail.

Je tiens à dire ma reconnaissance envers Monsieur le Professeur Ali LAKSACI Le directeur de notre laboratoire qui, malgré les prérogatives qui sont siennes, a accepté sans réserve, de diriger cette thèse, il s'y est grandement impliqué par ses directives, ses remarques et suggestions, mais aussi par ses encouragements dans les moments clés de son élaboration. Je tiens à le remercier aussi pour cette liberté qu'il a permise, sans laquelle le chercheur ne saurait affirmer sa manière de penser et de procéder, sa manière d'être, bref toute sa personnalité. sa rigueur scientifique et sa clairvoyance m'ont beaucoup appris. Ils ont été et resteront des moteurs de mon travail de chercheur.

Je ne manquerais pas non plus de dire un grand merci aux membres du jury qui ont accepté, sans réserve aucune, d'évaluer cette thèse à sa juste valeur, et de

me faire part de leur remarques sûrement pertinentes qui, avec un peu de recul, contribueront, sans nul doute, au perfectionnement du présent travail.

J'adresse mes sincères remerciements à tous les professeurs, à M.Samir TALHA et à M.Boubakar MECHAB intervenants et qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réfléxions et ont accepté à me rencontrer et répondre à mes questions durant mes recherches.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma maman, celle qui est toujours présente et continue de l'être pour faire mon bonheur. Merci pour t'être sacrifiée pour que tes enfants grandissent et prospèrent. Merci de trimer sans relâche, malgré les péripéties de la vie, au bien-être de tes enfants, elle est ma maman qui m'a fait comprendre que la vie n'est pas faite que de problèmes qu'on pourrait résoudre grâce à des formules mathématiques et des algorithmes, qui m'a donné un magnifique modèle de labeur et de pérsévérance. Je suis redevable d'une éducation dont je suis fière. Merci pour avoir fait de moi ce que je suis aujourd'hui. En ce jour mémorable, pour moi ainsi pour toi, reçoit ce travail en signe de ma vie reconnaissance et mon profond estime. Puisse le tout puissant te donner la santé, bonheur et longue vie afin que je puisse te combler à mon tour. je t'aime.

Autant de phrases et d'espressions aussi éloquentes soit-elle ne sauraient exprimer ma gratitude et ma reconnaissance à mon unique sœur Khaoula.
En souvenir d'une enfance dont nous avons partagé les meilleurs et les plus agréables moments, pour toute la complicité et l'entente qui nous unissent, ce travail est un témoignage de mon attachement et de mon amour. Merci pour ton soutien, ton encouragement et tes conseils durant toute ma vie, je t'aime ma confidente.

Pour toute l'ambiance dont tu m'as entouré, pour toute la spontanéité et ton talon chaleureux, je te dédie mon cher frère Mouaouia ce travail.
Merci pour avoir été l'homme de ma vie, merci pour ta tendresse et ton amour. Puisse Dieu le tout puissant exhausser tous tes vœux.

Est-ce un bon endroit pour dire ce genre de choses ? Je n'en connais en tous cas pas de mauvais. Je vous aime. votre présence, par votre amour, pour donner du goût et du sens à notre vie de famille, ô ! de si petite famille.

Mes remerciements tendent également à ma grand-mère maternelle pour l'expression des vœux qu'elle n'a céssé de formuler dans ses prières. Que Dieu vous préserve la santé et la longue vie.

Mes tantes, mes ancles mes cousins, mes cousines en Syrie (Alep) étaient mon sourir lors de ma nostalgie du pays, merci pour vos encouragements à supporter l'éloignement du pays et de la famille, merci pour vous trouver où j'ai terriblement besoin d'un petit mot, d'un petit geste, aussi humble soit-il, de soutien moral, merci énormement ! Que Dieu vous préserve la santé, le bonheur, vous protège et vous procure toutes et tous la longue vie.

Enfin, j'espère du fond du cœur que tout ce petit monde, mon monde à moi, trouve ici un mot de reconnaissance, et que chacun se reconnaisse en ce qui le concerne. J'espère aussi que l'effort déployé dans le présent travail réponde aux attentes des uns et des autres.

Je suis néanmoins seule et unique responsable des oublis, des lacunes et des faiblesses que puisse contenir la présente étude. Ceci étant, les propos contenus dans cette thèse n'engagent que ma propre responsabilité.

Une pensée pour terminer ces remerciements pour toi qui n'a pas vu l'aboutissement de mon travail mais je sais que tu en aurais été très fier de ta fille !!!

À la mémoire de mon père
Ahmed Farouk
Il a été toujours dans mon esprit et dans mon coeur
je te dédie aujourd'hui ma réussite.
Que Dieu le miséricordieux, vous accueille dans son
éternel paradis.

Table des matières

CHAPITRE 1

Introduction générale

On ne connaît pas complètement une science
tant qu'on n'en sait pas l'histoire.
Auguste Comte

Sommaire

1.1 Résumé

Dans cette thèse on s'intérésse à l'estimation non paramétrique de la fonction régression en présence des données incomplètes (données tronquées à gauche) pour une variable expliquative fonctionnelle et variable réponse réelle.
Dans un premier temps, on étudie la convergence presque complète de l'estimateur de la fonction de régression relative construit par la méthode de calcul de l'erreur quadratique relative dans le cas d'un modèle aléatoire de données tronquées à gauche à variable explicative fonctionnelle. On établit aussi la normalité asymptotique de son estimateur construit.
Dans un second temps, une généralisation sera étudié au cas des données α-mélangeantes, de même on construit son estimateur de la régression relative et on établit sa convergence presque complète.
Au final, une étude sur des données numériques sera données pour validé la performance de son estimateur par rapport à l'estimateur de la régression classique.

1.2 abstract

In this thesis we are interested in the non parametric estimation of the regression function in the presence of incomplete data (truncated data on the left) for a functional explanatory variable and a real response variable.
At first, we investigate the relationship between a functional random covariable and a scalar response which is subject to left-truncation by another random variable. Precisely, we use the mean squared relative error as a loss function to construct an non parametric estimator of the regression operator of this functional truncated data. Under some standard assumptions in Functional Data Analysis (FDA), we establish the almost sure consistency with rate of the constructed estimator as well as its asymptotic normality. In a second step, a generalization will be studied

in the case of strong mixing data , we construct our relative regression estimator and we establish its almost sure consistency with rate and its asymptotic normality. Then a simulation study, on finite-sized samples, was carried out in order to show the efficiency of our estimation procedure and to highlight its superiority over the classical kernel estimation, for different levels of simulated truncated data.

1.3 Étude bibliographique

La précision de l'estimateur pour les méthodes non paramétrique devient une problématique de recherche dont les statiticiens espèrent apporter de nouveau pour ce domaine. Camphell et al. (1989) qui ont pu élablir les premiers résulats dont ils ont utilisé l'outil de classification pour la regression relative. En 1999, Ruiz Valasco a montré l'effficacité asymptotique de la regression relative logistique dans un contexte paramétrique. Park et al. (1998) ont utilisé des techniques de régression non paramétrique pour avoir un estimateur de l'espérance conditionelle d'une variable réponse inversible. Jones et al. (2008) ont pu construire à l'aide de la méthode du noyau un estimateur consistant, ils ont pu avoir des propriétés asymptotiques dans le cas des variables aléatoires i.i.d. L'approche sur l'estimation de la survie relative des patients atteint d'un cancer colorectal a été étudié par Giorgi et al. (2008).

L'aspect non paramétrique, qui ne fait aucune hypothèse sur la loi, ni sur les paramètres. Ses connaissances sur le modèle ne sont pas exactes, ce qui est souvent le cas dans la pratique. Dans cette situation, il est naturel de vouloir estimer, soit généralement la fonction de répartition ou la densité (pour le cas continu) : c'est l'objectif de l'estimation fonctionnelle. Si la fonction de répartition empirique F_n résout le problème statistique fondamental de la distribution de probabilité (associée à un échantillon $X_1, X_2, \ldots, X_n$ de variables aléatoires réelles indépendantes et de même loi) en fonction des valeurs numériques observées, elle est par contre limitée pour décrire visuellement les caractères de l'échantillon. Pour Deheuvels (1980), c'est l'estimation la plus naturelle dans le cas où on ne fait, extérieurement à l'échantillon, aucune hypothèse restreignant le choix de cette distribution à une famille particulière. Cependant, elle n'en est pas moins limitée pour décrire les caractères de l'échantillon (ce qui occasionne une perte d'information sur la loi recherchée) Le lecteur intéressé est renvoyé de Bosq et Lecoutre (1987) ou Praksa Rao (1983).

Dans le cadre des données fonctionnelles, on pourrait citer à cet égard le papier J.Deville (1974) qui fût un des premiers (le premier peut-être) à envisager des variables statistiques continues/fonctionnelles. C'est une des raisons pour lesquelles un nouveau champ de la statistique dédié à l'étude de données fonctionnelles, a soulevé un grand défi au début des années quatre-vingt. En fait, et ce sont Ramsay et Silverman (1997), puis par les différents ouvrages de Bosq (2000), Ramsay et Silverman (2002, 2005) et Ferraty et Vieu (2006) a été popularisé ce domaine. On dit que c'est un des domaines de la statistique qui est en plein essor comme en témoignent les travaux publiés et/ou cités dans des revues de premiers rangs,..., etc. De plus, le statisticien a pu étudier des variables fonctionnelles à partir de son échantillon initial construites même si les données dont il dispose ne sont pas de nature fonctionnelle. Un exemple classique est celui où l'on observe plusieurs échantillons de données réelles indépendantes et où l'on est ensuite amenés à comparer les densités de ces différents échantillons ou bien à considérer des modèles où elles interviennent (Ramsay et Silverman, 2002). Dans le contexte particulier de l'étude des séries temporelles, l'approche introduite par Bosq (1991) fait apparaître une suite de données fonctionnelles dépendantes qui modélisent la série chronologique observée. Cette approche consiste tout d'abord à considérer le processus non pas à travers sa forme discrétisée mais comme étant un processus à temps continu puis à le découper en un échantillon de courbes successives.
Les tous premiers travaux dans lesquels on retrouve l'idée de considérer les données fonctionnelles sont relativement anciens. Rao (1958) et Tucker (1958) ont envisagé l'analyse en composantes principales et l'analyse factorielle pour des données fonctionnelles, en considérant explicitement les données fonctionnelles comme un type particulier de données. Par la suite, Ramsay (1982) a dégagé la notion de données fonctionnelles et a soulevé la question de l'adaptation des méthodes utilisées en

analyse statistique de données multivariées (en dimension finie) au cadre fonctionnel.

A partir de là, les travaux portant sur la statistique des données fonctionnelles ont commencé à se multiplier pour finalement aboutir, aujourd'hui, à des ouvrages devenus des références en la matière. Par exemple, les monographies de Ramsay et Silverman (2002) et 2005), Ferraty et Vieu (2006) présentent une collection importante de méthodes statistiques spécifiques aux variables fonctionnelles dans les cadres linéaire et non linéaire. De même, Bosq (1991) a contribué au développement de méthodes statistiques permettant l'analyse de variables aléatoires fonctionnelles dépendantes (processus autorégressifs hilbertiens). Citer plutôt, les travaux de Cuevas et al. (2002) qui se sont intéressés au problème de la régression linéaire d'une variable fonctionnelle sur un ensemble de données fonctionnelles déterministes "fixed functional design". D'autre part, Benhenni et al. (2010) ont considéré le problème d'estimation de l'opérateur de régression quand les données fonctionnelles sont déterministes et les erreurs sont corrélées. Cardot et al. (2005) quant à eux, ils ont proposé un estimateur non paramétrique de l'opérateur de régression quand le facteur prédictif est réel et la variable réponse est une courbe. Par ailleurs, l'étude du modèle de régression non linéaire est beaucoup plus récente que celle du cas linéaire. Ferraty et Vieu (2000) ont établi les premiers résultats sur l'estimation non paramétrique de l'opérateur de régression non linéaire. Ces résultats ont ensuite été prolongés par Ferraty et al. (2002) en traitant le cas de données dépendantes et en établissant des convergences fortes de l'estimateur à noyau de la régression. À leur tour, Niang et Rhomari (2003) ont étudié la convergence en norme $\mathcal{L}^p$ de l'estimateur de l'opérateur de régression et ont expérimonté leur résultats à la discrimination et à la classification de courbes. Rachdi et al. (2008) ont traité le problème d'estimation non paramétrique de l'opérateur de régression quand les erreurs vérifient des propriétés de longue mémoire. Ils ont établi aussi la convergence en probablilité ponctuelle puis uniforme de l'estimateur à noyau opératoriel. Une autre contribution basée sur la construction d'un critère de choix automatique et optimal du paramètre de lissage pour l'estimateur de la régression quand le régresseur est de type fonctionnel a été menée par Rachdi et Vieu (2005, 2007). Tandis qu'El Methni et Rachdi (2011) ont établi l'estimation locale d'une moyenne pondérées de l'opérateur de régression pour des données fonctionnelles déterministes. Ouassou et Rachdi (2010) ont amélioré ensuite cette estimation par l'estimateur de Stein.

La littérature sur les autres caractéristiques conditionnelles est aussi très importante. Citons, par exemple, Collomb et al. (1987) ont montré la convergence uniforme sur tout compact de $\mathbb{R}^p$ pour un processus ϕ-mélangeant. Quintela et Vieu (1997) ont pu estimer le mode conditionnel comme étant le point annulant la dérivée d'ordre un de l'estimateur de la densité conditionnelle et ont établi la convergence presque complète de cette estimateur pour le processus α-mélangeant.
Parmi tous ces types de dépendances, l'α-mélange est la plus faible et la moins restrictive de toutes ces formes de mélanges, autrement dit, toute suite de variable aléatoire sera alors forcément α-mélangeante. Les processus mélangeants sont très mises en valeur dans la pratique statistique. Dans la littérature, on trouve un nombre conséquent de travaux consacrés à l'étude des variables mélangeantes réelles dont les monographies de Doukhan (1994) ; Bosq (1996 ;1998) ; Rio (2000) et Yoshira (2004) donnent un point de vue global.

1.4 Quelques définitions préliminaires

1.4.1 Estimation par la méthode du noyau

Soit $(X_1, X_2, \ldots, X_n)$; un échantillon de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) de densité inconnue f. On définit l'estimateur à noyau de f par :

$$f_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

où $h := h(n)$ est une suite de nombres réels positifs (dépendant de n), appelés fenêtres ou largeurs de fenêtre, qui contrôlent le lissage de la courbe estimée, et K est une fonction bornée, intégrable, d'intégrale égale à 1, appelée noyau. Si de plus $\lim_{|u|\to\infty} uK(u) = 0$, K est appelé noyau de Parzen-Rosenblatt (et f_n l'estimateur à noyau de Parzen-Rosenblatt).

Définition 1.4.1. Un noyau $K(x)$ est dit d'ordre r, $(r \geq 1)$ si les fonctions $x \to x^j K(x), j = \overline{1,r}$ sont intégrables et vérifient $\int K(x)\mathrm{d}x$, et $\int x^j K(x)\mathrm{d}x = 0, j = \overline{1,r}$.

1.4.2 Régression par la méthode du noyau

Par analogie, elle consiste à remplacer la densité de probabilité par un noyau quelconque de type décrit, la régression par la méthode de noyau s'écrit sous la forme

$$\hat{r}_n(t) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)}.$$

La vitesse de convergence, au sens de l'erreur quadratique moyenne, est également à l'ordre de $n^{-5/4}$.

1.4.3 Donnée fonctionelle

Définition 1.4.2 (Ferraty et Vieu(2006))**.** : 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.

$\mathbf{X} = \{X_t, t \in \mathcal{T}\}$, où $\mathcal{T}$ est un ensemble continu pouvant représenter par exemple un intervalle de temps ou un spectre de longueurs d'ondes. On parle alors de l'analyse statistique de courbes ou de l'analyse de données fonctionnelle.

$$X_t : \mathcal{P} \to \mathbb{R}$$

– $\mathcal{T} \subseteq \mathbb{R}$: courbe

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

Définition 1.4.3. Un modèle fonctionnel est dit paramétrique si C est indexable par un nombre fini de paramètres appartenant à $\mathcal{F}$ où C n'est qu'un sous-ensemble de $\mathbb{F}_{\mathcal{F}}^{\mathcal{F}'}$ où $\mathbb{F}_{\mathcal{F}}^{\mathcal{F}'}$ l'ensemble des fonctions définies sur l'espace fonctionnel $\mathcal{F}$ et à valeurs dans l'espace $\mathcal{F}'$.
Un modèle fonctionnel est dit non-paramétrique dans le cas contraire.

1.4.4 Données incomplètes

Par ailleur, le problème des données manquantes, incomplètes ou erronées est très vaste et a suscité beaucoup d'intérêt parmi les statisticiens ces dernières années. L'attitude vis-à-vis de ce type de données a longtemps été soit de les éliminer, soit de minimiser le mauvais impact qu'elles pourraient avoir sur des procédures statistiques adaptées à des données complètes. Dans le domaine des durées de survie, les données sont souvent incomplètes à cause de deux phénomènes distincts : la censure et la troncature.

Dans leur acception la plus générale, ces deux notions ont la signification suivante. Premièrement, et c'est la troncature, on n'observe x que si elle appartient à un sous-ensemble B de ses valeurs possibles. On dit que x est tronquée par B.

Deuxièmement, et il s'agit là de la censure, même dans le cas où x appartient à B, on n'observe pas x complètement ; on sait seulement de cette variable qu'elle appartient à un sous-ensemble A de B. On dit qu'elle est censurée par A. Dans le cas où un n-échantillon de X soit $x_1, x_2, \ldots, x_n$ est concerné, à chacun des x_i sont associés deux ensembles B_i et A_i le premier qui tronque x_i et le second qui le censure. La plupart du temps, x_i est une variable réelle positive, et A_i et B_i sont des demi-droites du type $(-\infty, c_i)$ ou $(c_i, +\infty)$, ce qui correspond à des censures ou troncatures dites gauches ou droites, et c'est donc ce cas qui sera illustré par des exemples au paragraphe suivant. Pour des censures et troncatures par intervalles, on pourra voir Turnbull (1976), Groeneboom (1992), Frydman (1994) et Alioum et Commenges (1994).

Données censurées

1. Censure de type I (fixée)

   Nous sommes dans le cas décrit précédemment où C est une valeur constante fixée.

2. Censure de type II (attente)

   Elle est présente quand on décide d'observer les durées de survie des n patients jusqu'à ce que k d'entre eux soient décédés et d'arrêter l'étude à ce moment là. Soient $Y_{(i)}$ et $T_{(i)}$ les statistiques d'ordre des variables Y_i et T_i. La date de censure est donc $Y_{(k)}$ et on observe les variables suivantes :

$$T_{(1)} = Y_{(1)}$$

$$\vdots$$

$$T_{(k)} = Y_{(k)}$$

$$T_{(k+1)} = Y_{(k)}$$

$$\vdots$$

$$T_{(n)} = Y_{(k)}.$$

3. Censure de type III (ou censure aléatoire de type I)

   Soit $C_1, \cdots, C_n$ des v.a i.i.d. On observe les variables $Y_i = T_i \wedge C_i = \min(T_i, C_i)$. L'information disponible peut être résumée par :

   - la variable Y_i réellement observée,
   - l'indicatrice $\delta_i = \mathbb{I}_{\{T_i \leq C_i\}}$

     où

$$\delta_i = \begin{cases} 1 & \text{si on observe les "vraies" durées (l'évènement est observé ie; } Y_i = T_i, \\ 0 & \text{on observe les durées incomplètes (l'individu est censuré) ie; } Y_i = C_i. \end{cases}$$

   La censure aléatoire est la plus courante. Par exemple, lors d'un essai thérapeutique, elle peut être engendrée par :

   a) la perte de vue : le patient quitte l'étude en cours et on ne le revoit plus (à cause d'un déménagement, le patient décide de se faire soigner ailleurs). Ce sont des patients "perdus de vue".

   b) l'arrêt ou le changement du traitement : les effets secondaires ou l'inefficacité du traitement peuvent entraîner un changement ou un arrêt du traitement. Ces patients sont exclus de l'étude.

   c) la fin de l'étude : l'étude se termine alors que certains patients sont toujours vivants (ils n'ont pas subi l'événement). Ce sont des patients "exclus-vivants". Les "perdus de vue" (et les exclusions) et les "exclus-vivants" corresp ondent à des observations censurées mais les deux mécanismes sont de nature différente (la censure peut être informative chez les "perdus de vue").

Exemple 1.4.1 (KOZIOL ET GREEN (1976))**.** un essai clinique est réalisé sur 211 individus atteints du cancer de la prostate (phase 4) traités par oestrogène (hormone). A la fin de l'étude, 90 meurent du cancer de la prostate, 105 meurent d'autres causes et 16 sont encore vivants. Les censurés à droite sont les 105 + 16 (= 121) individus qui ne sont pas morts du cancer de la prostate (objet de l'étude).

Données tronquées Ce mécanisme empêche l'observation de la variable T entièrement (en général les valeurs extrêmes), et engendre une perte d'information (on n'étudie qu'un sous-échantillon). Il y a troncature si l'observation de la variable d'intérêt T n'a lieu que conditionnellement à un événement A. On dit qu'il y a une troncature

- à droite lorsque T n'est observable que si elle est inférieure à un seuil C positif fixé ou aléatoire,
- à gauche lorsque T n'est observable que si elle est supérieure à C. En général, on observe le couple (T, C), avec $T \geq C$ ou inversement, selon le cas. La troncature gauche apparait par exemple, dans la détection de réserves pétrolières (voir Woodroofe (1985)) : On ne connaît pas le nombre total de gisements, mais on peut observer les n gisements suffisamment grands (supérieurs à C) pour être exploités,
- par intervalle quand une durée est tronquée à droite et à gauche, on dit qu'elle est tronquée par intervalle. Par exemple, on rencontre ce type de troncature lors de l'étude des patients d'un registre : les patients diagnostiqués avant la mise en place du registre ou répertoriés après la consultation du registre ne seront pas inclus dans l'étude.

1.5 Estimation de la fonction de survie

Si l'on ne peut pas supposer a priori que la loi de la durée de survie obéit à un modèle paramétrique, on peut estimer la fonction de survie S grâce à plusieurs méthodes non-paramétriques dont la plus intéressante est celle de Kaplan-Meier. Cet estimateur est aussi appelé PL (Produit-Limite) car il s'obtient comme la limite d'un produit. Il est fondé sur la remarque suivante : si $t' < t$, la probabilité de survivre au-delà de l'instant t est égale au produit suivant :

$$S(t^+) = \mathbb{P}(X > t | X > t').S(t')$$

Si l'on renouvelle l'opération en choisissant une date t'' antérieure à t', on aura de même $S(t') = \mathbb{P}(X > t' | X > t'').S(t'')$, et ainsi de suite. Si l'on choisit pour les dates où l'on conditionne celle où s'est produit un événement, qu'il s'agisse d'une mort, d'une censure ou d'une troncature, on aura seulement à estimer des quantités de la forme :

$$\mathbb{P}(X > T_{(i)} | X > T_{(i-1)}).S(t') = p_i$$

or p_i est la probabilité de survivre pendant l'intervalle de temps $I_i =]T_{(i-1)}, T_{(i)}]$ quand on était vivant au début de cet intervalle.
Notant, comme précédemment, R_i le nombre des sujets qui sont vivants (donc "à risque" de mourir) juste avant l'instant $T_{(i)}$, ce qui peut aussi s'écrire :

- vivant à l'instant $T_{(i)}$ ou
- sujets de $R(T_{(i)})$ en désignant par $R(t)$ l'ensemble des sujets à risque à l'instant t^- et M_i le nombre de mort à l'instant $T_{(i)}$, $q_i = 1 - p_i$ est la probabilité de mourir pendant l'intervalle I_i sachant que l'on était vivant au début de cet

intervalle. Alors l'estimateur naturel de q_i est

$$\hat{q}_i = \frac{M_i}{R_i}.$$

1.6 Mesure de dépendance

Dans le cadre de l'estimation non paramétrique, il convient de modéliser la dépendance entre les variables aléatoires. Le type de "mélange" de dépendance est ainsi largement utilisé dans la littérature.
Les observations dépendantes sont plus ajustées à la réalité. Il existe, alors, de nombreuses notions de dépendance. On s'intéresse ici à celles qui s'expriment en termes de coefficients de mélange entre les tribus engendrées par le passé et le futur de la suite de variables alétoires $(X_n)_{n\leq 1}$. Rosen-blatt (1956) a définit les coefficients de α-mélange α_n, mesurent l'écart entre les indicatrices des événements qui appartenant respectivement à la tribu engendrée par le futur des variables après l'instant n et celle engendrée par le passé des variables avant l'instant zéro.

Les suites α-mélangeantes ont beaucoup d'intétrêt ; les processus linéaires sont sous certaines conditions α-mélangeants et leurs coefficients de mélange ont un ordre de grandeur explicite, il en est de même pour les chaînes de Markov. Cependant, il est techniquement pas aussi facile d'évaluer ces coefficients α_n des suites mélangeantes (voir Doukhan pour une étude complète de ce sujet).

1.6.1 Notion de mélange

Les différentes notions de mélange sont reliées à des mesures sous-jacentes de la dépendance entre les σ-algèbres. Plus présisément, soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, prenant $\mathcal{L}$ et $\mathcal{L}'$ deux sous-algèbres de $\mathcal{F}$ Plusieurs mesures de dépendance entre $\mathcal{L}$ et $\mathcal{L}'$ sont définies comme suit :

Définition 1.6.1. Le cœfficient de mélange fort ou d'α-mélange introduit par Rosenblatt (1956)

$$\alpha(\mathcal{L}, \mathcal{L}') := \sup \Big\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|; A \in \mathcal{L}, B \in \mathcal{L}' \Big\}.$$

La suite est dite alpha-mélangeante ou fortement mélangeante si le coefficient d'alpha-mélange vérifie $\lim\limits_{n\to\infty} \alpha(n) = 0$.

On peut définir deux types de mélanges forts :

- à décroissance géométrique, s'il existe $c > 0$ et $\rho \in]0, 1[$ tels que $\alpha(n) \leq c\rho^n$. Les processus AR et ARMA sont des exemples de processus géométriquement fortement mélangeants.
- à décroissance arithmétique d'ordre $s > 0$, s'il exite un nombres réel c strictement positif tels que $\alpha(n) \leq \dfrac{c}{n^s}$.

Propriété 1.6.1. L'échelle de ce cœfficient est définie par l'inégalité suivante :

$$0 \leq \alpha(\mathcal{L}, \mathcal{L}') \leq \frac{1}{4}.$$

Les résultats obtenus dans le cas de l'α-mélange vont donc concerner une classe plus large de processus.

Propriété 1.6.2. $\alpha(\mathcal{A}, \mathcal{B}) = 0$, ceci équivalent à l'indépendance des tribus $\mathcal{A}$ et $\mathcal{B}$.

1.6.2 Les conditions du mélange fort

Soit $(X_1, X_2, \ldots, X_k), k \in \mathbb{Z}$, une suite de variables aléatoires pas nécessairement stationnaire.

Pour $-\infty \leq j \leq l + \infty$, on définit la $\sigma-$algèbre $\mathcal{F}_j^l = \sigma(X_k, j \leq k \leq l \quad (k \in \mathbb{Z}))$.
Pour tout $n \geq 1$, définissons le coefficients de dépendance suivants :

$$\alpha(n) = \sup_{k \in \mathbb{Z}} \quad \alpha\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{+\infty}\right).$$

Remarque 1.6.1. Dans le cas de la stricte stationnarité de X ; on a simplement

$$\alpha(n) = \alpha\left(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_{n}^{+\infty}\right)$$

1.6.3 Inégalités de type exponentiel

L'outil qu'on va utiliser de manière déterminante dans les problèmes de convergence presque complète est l'inégalité exponentielle ci-dessous que l'on peut rencontrer dans Rio (2000) sous l'appellation inégalité de Fuk-Nagaev. Cette inégalité est en fait une extension au cadre de variables fortement mélangeantes de l'inégalité de Bernstein que l'on peut trouver sous diverses formes dans H$\ddot{o}$effding (1963. En fait, le développement des résultats en estimation non paramétrique de régression sous dépendance s'est fait en parallèle avec celui de telles inégalités. Depuis, la puissance de telles inégalités a été sérieusement améliorée (c.f. Bosq,1993), jusqu'à l'inégalité de Rio (2000). Le lecteur trouvera dans Rio (2000) mais aussi dans Bosq (1996) plusieurs versions de cette inégalité, dont certaines sous des énoncés plus généraux que ce qui va suivre. Cependant nous en tiendrons ici a cette version simplifiée qui est largement suffisante dans notre contexte, et que l'on trouve sous cette forme dans Rio (2000, p :87,formule (6 :19b)).

Inégalité exponentielle pour le cas i.i.d. Soit $U_1, U_2, \cdots, U_n$ une suite de variables aléatoires indépendantes identiquement distribuées, et $\sigma_n^2 = \mathbb{E}(U_j^2)$. s'il existe

$M = M_n < \infty$, tel que $U_1 \leq M$. on a :

$$\forall \epsilon > 0, \quad \mathbb{P}\left(\frac{1}{n}\Big|\sum_{j=1}^{n} U_j\Big| > \epsilon\right) \leq 2\exp\left(-\frac{\epsilon^2 n/2}{\sigma^2 + \epsilon M}\right)$$

si $u_n = \dfrac{\sigma_n^2 \log n}{n}$, tel que $\lim\limits_{n\to\infty} u_n = 0$, avec $M/\sigma^2 < \infty$, on a donc :

$$\frac{1}{n}\sum_{j=1}^{n} U_j = O(\sqrt{u_n}).$$

Inégalité exponentielle pour le cas dépendant.

Lemme 1.6.1. *(Kallabis and Neumann (2006)) Soit $U_1, U_2, \cdots, U_n$ une suite de variables aléatoires $\mathbb{E}(U_j) = 0$ et $\mathbb{P}(|U_j| \leq M) = 1$, pour tout $j = 1, \cdots, n$ et certaines $M < \infty$. Soit $\sigma_n^2 = Var\left(\sum_{j=1}^{n} U_j\right)$ supposer, en outre, qu'il existe $K < \infty$ et $\beta > 0$ tel que, pour tous les u-triplets $(s_1, \cdots, s_u)$ et tous les v-triplets $(t_1, \cdots, t_v)$ avec $1 \leq s_1 \leq \cdots \leq s_u \leq t_1 \leq \cdots \leq t_v \leq n$, l'inégalité suivante est satisfaite :*

$$\Bigg|Cov(U_{s_1}\cdots U_{s_u}, U_{t_1}\cdots U_{t_v})\Bigg| \leq K^2 M^{u+v-2} v e^{-\beta(t_1 - s_u)}.$$

Donc,

$$\mathbb{P}\left(\Big|\sum_{j=1}^{n} U_j\Big| > t\right) \leq \exp\left\{-\frac{t^2/2}{A_n + B_n^{1/3} t^{5/3}}\right\}$$

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

Lemme 1.6.2. *Inégalité de type Fuk-Nagaev :*

soit $(U_i, i \geq 1)$ *une suite de v.a.r centrées et de cœfficient de mélange fort* $\alpha(n) = \mathcal{O}(n^{-a})$, $a > 1$, *et* $\forall n \in \mathbb{N}$, $1 \leq i \leq n$, $|U_i| < \infty$. *Alors pour tout* $r > 1$ *et* $\epsilon > 0$

$$\mathbb{P}\left[\Big|\sum_{i=1}^{n} U_i\Big| > \epsilon\right] \leq c\left[\left(1 + \frac{\epsilon^2}{rS_n}\right)^{-r/2} + nr^{-1}\left(\frac{2r}{\epsilon}\right)^{a+1}\right]$$

avec

$$S_n = \sum_{1\leq i\leq n}\sum_{1\leq j\leq n} |Cov(U_i, U_j)|.$$

Théorème 1.6.1. *Soit* $X_1, \ldots, X_n$ *des variables aléatoires indépendantes centrées et réduites et soit* $S_n = \sum_{i=1}^{n} X_i$. *Les inégalités suivante tiennent :*

i) (Inégalité de Hoeffding)

Si $a_i \leq X_i \leq b_i$; $i = 1, \ldots, n$ *où* $a_1, b_1, \ldots, a_n, b_n$ *sont des constantes ensuite,*

$$\mathbb{P}(|S_n| \geq t) \leq 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right), \qquad t > 0.$$

ii) (Inégalité de Cramer)

S'il exite $c > 0$ *tel que ;*

$$\mathbb{E}(|X_i|^p) \leq c^{p-2} p! \mathbb{E}(X_i^2) < +\infty$$

$i = 1, \ldots, n;\ \ p = 3, 4, \ldots.$

iii) (Inégalité de Bernstein)

$$\mathbb{P}(|S_n| \geq t) \leq 2\exp\left(-\frac{t^2}{4\sum_{i=1}^{n}\mathbb{E}(X_i^2) + 2ct}\right), \qquad t > 0.$$

Théorème 1.6.2. *Inégalité de Rio :*

Soit $(X_n, n \in \mathbb{N})$ une famille de variables aléatoire réelles indépendantes telle que $\mathbb{E}(X_i) = 0$. Supposons qu'il existe des constantes positives A et B telles que, pour tout t positif,

$$\log \mathbb{E}\left[\exp(tS_n)\right] \leq \frac{At^2}{2(1 - Bt)}$$

avec $(S_n)_n$ une martingale adaptée à la filtration $(\mathcal{F}_n)_n$ telle quue $S_0 = 0$, alors :

$$\mathbb{P}(S_n^* \geq t) \leq \exp\left(-\frac{t^2}{2(A + Bt)}\right)$$

où : $S_n^* = \max(0, S_1, \ldots, S_n)$.

Théorème 1.6.3. *Inégalité de Rouassas et Ioannids :*

Soit s_i, t_i et ξ_i, $i = i, \ldots, n$ dans

$$\begin{cases} 1 = s_1 < t_1 < s_2 < t_2 < \cdots < s_n < t_n \\ \textit{avec} \quad s_{i+1} - t_i \geq \pi, \qquad i = 1, \ldots, n-1, \end{cases}$$

et ξ_i est une variable aléatoire $\mathcal{F}_{s_i}^{t_i} = \sigma(X_j, j = s_i, \ldots, t_i)-$mesurable, $i = 1, \ldots, n$ respectivement et on suppose que

$$|\xi_i| \leq M_i \quad p.s \quad i = 1, \ldots, n.$$

Then, sous ϕ_4-mélange,

(i)

$$|\mathbb{E}(\xi_1,\ldots,\xi_n)| - \mathbb{E}(\xi_1)\ldots\mathbb{E}(\xi_n) < 4(n-1)\phi_4(\pi)\prod_{i=1}^{n} M_i$$

pour des v.a réelles, et

(ii)

$$|\mathbb{E}(\xi_1,\ldots,\xi_n)| - \mathbb{E}(\xi_1)\ldots\mathbb{E}(\xi_n) < 16(n-1)\phi_4(\pi)\prod_{i=1}^{n} M_i$$

pour des v.a complexes.

Inégalité de Markov. Soit $g : \mathbb{R} \to \mathbb{R}$ est une fonction tel que $g(X)$ est est une vraie variable aléatoire, $\mathbb{E}(g(X)) < \infty$, et

$$\mathbb{P}[0 \leq g(X) \leq M] = 1$$

où $M \in [0, \infty]$. Si $g(x)$ est une fonction non décroissante sur $\mathbb{R}$, donc, pour tout $a \in \mathbb{R}$,

$$\frac{\mathbb{E}(g(X)) - g(a)}{M} \leq \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(g(X))}{g(a)}$$

si $g(x)$ est une fonction non décroissante sur $[0, \infty)$, et $g(x) = g(-x)$ pour tout $a \geq 0$ donc,

$$\frac{\mathbb{E}(g(X)) - g(a)}{M} \leq \mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(g(X))}{g(a)}$$

où $0/0 = 1$.

Lemme de Borel-Cantelli.

Lemme 1.6.3. *Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'événements,*

i) Si $\sum_n \mathbb{P}(A_n) < \infty$ alors $\mathbb{P}(\limsup A_n) = 0$.

ii) Si les A_n sont indépendents et si $\sum_n \mathbb{P}(A_n) = \infty$ *alors* $\mathbb{P}(\limsup A_n) = 1$, *c'est à dire que* $\mathbb{P}$ *p.s.* $\omega \in (A_n)$ *infiniment souvent.*

1.6.4 Inégalités de moment

Inégalité de Hölder. Soit $r > 1$, tels que $\frac{1}{r} + \frac{1}{s} = 1$ et $X \in \mathcal{L}^p, Y \in \mathcal{L}^p$, alors :

$$\mathbb{E}(|XY|) \leq \mathbb{E}(|X|^r)^{1/r}\mathbb{E}(|Y|^s)^{1/s}.$$

c$_r$-inégality(Loève 1963).

$$\mathbb{E}(|X+Y|^r) \leq c_r \left[\mathbb{E}(|X|^r) + \mathbb{E}(|Y|^r)\right]$$

où :

$$c_r = \begin{cases} 1 & \text{si} \quad 0 < r \leq 1, \\ 2^{r-1}, & \text{si} \quad r > 1. \end{cases}$$

La moyenne de l'inégalité c$_r$.

$$\mathbb{E}(\frac{1}{2}|X+Y|^r) \leq (\frac{1}{2})^r \left[\mathbb{E}(|X|^r) + \mathbb{E}(|Y|^r)\right] \qquad \text{si} \qquad 0 < r \leq 1$$

$$\leq \frac{1}{2}\left[\mathbb{E}(|X|^r) + \mathbb{E}(|Y|^r)\right], \qquad \text{si} \qquad r > 1.$$

Théorème central limite de Lypunov.

Théorème 1.6.4. $\log\left[\mathbb{E}(|X|)^r\right]$ *est une fonction convexe de r, i.e pour tout* $\lambda \in [0,1]$,

$$\log\left[|X|^{(\lambda r+(1+\lambda)r)s}\right] \leq \lambda \log\left[|X|^r\right] + (1-\lambda)\log\left[|X|^s\right].$$

Limites inférieures sur les moments de la somme Si

$$\mathbb{E}\left[|X|^r\right] < \infty, \mathbb{E}\left[|X|^s\right] < \infty$$

et $\mathbb{E}\left[|X \backslash Y|\right] = 0$, donc,

$$\mathbb{E}\left[|X+Y|^r\right] \geq \mathbb{E}\left[|X|^r\right], \qquad \text{pour} \qquad r \geq 1.$$

Remarque.— Souvent, A_n représente une propriété dépendant de n. On veut savoir si elle est vériée une infinité de fois.

Démonstration. **i)** La suite $\cup_{m \geq n} A_m$ est une suite décroissante d'ensembles, donc

$$\mathbb{P}(\limsup_n A_n) = \lim_n \mathbb{P}(\cup_{m \geq n} A_m).$$

Mais, $\mathbb{P}(\cup_{m \geq n} A_m) \leq \sum_{m \geq n} \mathbb{P}(A_m)$. Comme c'est une série convergente, la série donc tend vers 0.

ii) On regarde,

$$\mathbb{P}((\cup_{m \geq n} A_m)^c) = \mathbb{P}(\cap_{m \geq n} A_m^c),$$

mais grâce à l'indépendance des (A_n),

$$\mathbb{P}(\cap_{m \geq n}^k A_m^c) = \prod_{m=n}^k (1 - \mathbb{P}(A_m)) \leq \exp(-\sum_n \mathbb{P}(A_m)), \qquad \forall k \geq n.$$

Ce dernier terme tend vers 0 quand k tend vers l'infini. Donc pour tout n, on a

$$\mathbb{P}(\cup_{m \geq n} A_m) = 1$$

pour tout n, ce qui conclut la preuve.

□

1.7 Présentation du cadre de travail et des résultats

1.7.1 Problématique

Le travail présenté dans ce mémoire a pour objectifs d'établir les propriétés asymptotiques de l'estimateur à noyau de la régression fonctionnelle de l'erreur relative pour les données tronquées. Plus précisément, on évoque la convergence presque complète avec les taux, ainsi que la normalité asymptotique de cet estimateur sachant que la caractéristique principale de son approche est de développer un modèle de prévision alternative à la régression classique qui n'est pas sensible à la présence des valeurs aberrantes.
L'approche adoptée dans ce travail pour l'estimation de l'erreur relative de régression fonctionnelle dans un modèle aléatoire de troncature gauche est une synthèse d'une série de peu de recherches (Demongeot et al. (2016), Horrig et Ould-Said (2011-2014), Hellal et Ould-Said (2016)).

1.7.2 Contribution de la thèse

Le présent travail traite de l'analyse statistique non paramtrique des données fonctionelle. Cette thèse est organisée en 4 chapitres, suivis d'une brève conclusion et décrits succintement comme suit :

- **Chapitre 1.** On y présente une introduction générale dans laquelle certaines notions intervenant dans la suite, et nécessaires pour une meilleure compréhension de la thèse, sont rappelées ou développées.
- **Chapitre 2.** On s'intéresse à la relation entre une variable explicative aléatoire fonctionnelle et une réponse scalaire qui est soumise à une troncature à gauche par une autre variable aléatoire dont on a établi la convergence presque complète avec le taux de l'estimateur construit ainsi que sa normalité asymptotique.

- **Chapitre 3.** Est consacré à l'extension des résultats de chapitre 2 au cas des données fortement mélangeantes. On donne des résultats sur la convergence forte et la normalité asymptotique d'un estimateur à noyau dans le cas de $\alpha-$mélange.
- **Chapitre 4.** Qui est une application numérique, dont le bute est faire la comparaison entre l'estimateur de la regression relative $\tilde{r}(x)$ et l'estimateur de la regression classique $r(\hat{x})$, ensuite on fait le teste de la normalité asymptotique des deux estimteur par la méthode de simulation.

Liste des travaux

Publications

1. Functional data analysis : estimation of the relative error in functional regression under random left-truncation model.

CHAPITRE 2

Analyse des données fonctionnelles : Estimation de l'erreur relative de régression fonctionnelle dans un modèle aléatoire de troncature gauche

La science consiste à passer d'un étonnement à un autre.
Aristote

Sommaire

Ce chapitre fait l'objet d'une publication acceptée par " JOURNAL OF NONPARAMETRIC STATISTICS", dont on va exposer la partie théorique.

abstract

In this paper, we investigate the relationship between a functional random covariable and a scalar response which is subject to left-truncation by another random variable. Precisely, we use the mean squared relative error as a loss function to construct an non-parametric estimator of the regression operator of this functional truncated data. Under some standard assumptions in Functional Data Analysis (FDA), we estiablish the almost sure consistency with rate of the constructed estimator as well as its asymptotic normality.
keywords Functional data analysis ; Censored data ; Truncated data ; Small ball probability ; Functional regression ; Relative error ;

2.1 Introduction

The general framework of this paper is the analysis of the effect of a functional random covariate X on a scalar random response Y which is subject to left-truncation by another random variable T. Recall that the analysis of incomplete functional data (censored and/or truncated data) has a great importance in practice. In particular, such kind of data occur in many fields of applied statistics, for instance, in medicine, biometry, astronomy, economics, epidemiology, etc.

Notice that, there are several ways for explaining the link between two variables. Often, the nonparametric analysis of the relationship between X and Y is modeled by the following consideration :

$$Y = r(X) + \epsilon, \tag{2.1}$$

where r is an operator which is defined from a semi-metric space $(\mathcal{F}, d)$, equipped with a semi-metric d, to $\mathbb{R}$ and ϵ is a random error variable. Usually, the operator r is estimated by minimizing the following loss function :

$$\mathbb{E}\left[(Y - r(X))^2|X\right].$$

However, this kind of loss function is very sensitive to outliers. Indeed, this loss function treats all variables as having an equal weight. Thus, the presence of large outliers can lead to irrelevant results. For this reason, in this contribution, we overcome this drawback by using an alternative loss function based on the squared relative error which is defined, for $Y > 0$, by :

$$\mathbb{E}\left[\left(\frac{Y - r(X)}{Y}\right)^2|X\right]. \tag{2.2}$$

Clearly this criterion is more interesting than the least squared error, when the range of predicted values is large. In particular, this kind of obtained regression is less sensitive to the presence of outliers (see Demongeot et al., 2016). Moreover, the solution of the minimization of (2.2) can be explicitly expressed by :

$$r(x) = \frac{\mathbb{E}[Y^{-1}|X = x]}{\mathbb{E}[Y^{-2}|X = x]}.$$

Recall that the robustification of the nonparametric functional regression is an important subject in Nonparametric Functional Data Analysis (NFDA). Most developed methods are based on a M-estimation approach. One can see, for instance, Azzedine et al. (2008) and Attouch et al. (2009) for previous results and, Gheriballah et al. (2013) and Derrar et al. (2016) for recent advances and references.

The literature on the relative error regression in NFDA is still limited. At the best of our knowledge, only the paper by Demongeot et al. (2016) have paid attention to investigate this problem when the regressors are of functional kind. They proved the almost complete consistency and the asymptotic normality of a kernel estimator of r. On the other hand, the literature on the nonparametric analysis of incomplete functional data is still quite restricted. There are very few results on this topic (see, for instance, Horrigue and Ould-Said (2011 and 2014) for the conditional quantiles estimation in censorship model and functional regressors, and Helal and Ould-Said (2014) for the same model with truncated data). We return to Ling et al (2015, 2016) for some results on the classical regression or modal regression for functional ergodic data when missing responses are present. Among the lot of papers on functional statistics or nonparametric analysis of incomplete data we only refer to Ferraty and Vieu (2006), Zhang (2013), Cuevas (2014),Hsing and Eubank (2015), Bongiorno et al (2014), Aneiros et al (2017), Ould-Said and Lemdani (2006), Wang et al. (2012a).

The main aim of this paper is to establish asymptotic properties of the kernel estimator of the functional relative error regression for truncated data. More precisely, we establish the almost sure convergence with rates, as well as the asymptotic normality of this estimator. Notice that the main feature of our approach is to develop a prediction model alternative to the classical regression which is not sensitive to the presence of the outliers.
The rest of the paper is organized as follows. Section 2 is devoted to the presentation of the model and to the formulation of the estimator of r. The needed assumptions and main results are given in Section 3. Finally, the proofs of main results are relegated to the last section.

2.2 The functional truncated model and its estimator

Let (X_i, Y_i) for $i = 1, \ldots, N$ be N independent and identically distributed (i.i.d.) copies of (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$ where the sample size N is deterministic, but unknown. This sample is not completely observed, we observe only n variables (among N). Specifically, we suppose that the response observations $(Y_i)_{i=1,\ldots,N}$ are left-truncated by $(T_i)_{i=1,\ldots,N}$, in sense that the random variable (r.v.) of interest Y is interfered by the r.v.

T, such that both quantities Y and T are observable only if $Y \geq T$. Furthermore, we suppose that the variable of interest Y and the truncating variable T are independent and have, respectively, unknown distribution functions F and G. Obviously, in this context of truncated data, our estimator is constructed by the observed variables $(X_i, Y_i, T_i)_{i=1,\ldots,n}$, among the N variables. The size of the actually observed sample, n is a $Bin(N, \tau)$ random variable, with $\tau := \mathbb{P}(Y \geq T)$. Of course if $\tau = 0$, no data can be observed and therefore, we suppose that $\tau > 0$.

We point out that if the original data (Y_i, T_i) for $i = 1, \ldots, N$ are independent and identically distributed (i.i.d.), the observed data (Y_i, T_i) for $i = 1, \ldots, n$ are still i.i.d. (see Ould-Said and Lemdani 2006).

Following the same reasoning as Lemdani and Ould-Said (2007), the kernel estimator of the truncated relative error regression of $r(x)$ is based on the Lynden-Bell estimator, G_n, of the cumulative distribution function G of the r.v. T. To express this estimate, we point out that, according to Stute (1993) the distribution functions of Y and T, under the left-truncation condition, are expressed, respectively, as follows :

$$F^*(y) := \tau^{-1} \int_{-\infty}^{y} G(v)dF(v) \quad \text{and} \quad G^*(t) := \tau^{-1} \int_{-\infty}^{\infty} G(t \wedge v)dF(v)$$

where $t \wedge v = \min(t, v)$, and are estimated by their empirical analogues :

$$F_n^*(y) = n^{-1} \sum_{i=1}^{n} 1\!\!1_{\{Y_i \leq y\}} \quad \text{and} \quad G_n^*(y) = n^{-1} \sum_{i=1}^{n} 1\!\!1_{\{T_i \leq y\}}.$$

Then, we define

$$R(y) := \tau^{-1} G(y)[1 - F(y)]$$

which is estimated empirically by

$$R_n(y) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{\{T_i \le y \le Y_i\}}.$$

Thus, the distribution functions F and G are estimated by the following respective Lynden-Bell (1971) estimators :

$$F_n(y) = 1 - \prod_{s \le y} \left[1 - \frac{F_n^*(s)}{R_n(s)} \right] \quad \text{and} \quad G_n(y) = 1 - \prod_{s > y} \left[1 - \frac{G_n^*(s)}{R_n(s)} \right].$$

Then, He and Wang (1998) showed that τ can be estimated by

$$\tau_n := \frac{G_n(y)(1 - F_n(y))}{R_n(y)}$$

which is independent of y.

Now, the kernel estimator of the truncated relative error regression of $r(x)$ is defined by :

$$\widetilde{r}(x) := \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K(h^{-1} d(x, X_i)) Y_i^{-1}}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K(h^{-1} d(x, X_i)) Y_i^{-2}}.$$

where K is a kernel which satisfied some conditions, $h = h_{K,n}$ is a sequence of positive numbers

The asymptotic properties of F_n and G_n were studied by Woodroofe (1985), while the almost sure consistency of τ_n was obtained by He and Yang (1998). Noting that as N is unknown and n is known (although random), the stochastic consistency in truncation model would be stated with respect to the conditional probability $\mathrm{I\!P}(\cdot)$ (related to the n-sample) instead of the probability measure $\mathbb{P}(\cdot)$ (related to the N-sample). We denote by $\mathrm{I\!E}(\cdot)$ and $\mathbb{E}(\cdot)$ the respective expectation operators of $\mathrm{I\!P}(\cdot)$ and $\mathbb{P}(\cdot)$. Finally, let us point out that the asymptotic properties of F_n, G_n and τ_n are obtained under identifiability conditions $a_G \le a_F$ and $b_G \le b_F$, where a_W and b_W are the endpoints of the support of the distribution function W which are defined by :

$$a_W = \inf\{x : W(x) > 0\} \quad \text{and} \quad b_W = \sup\{x; W(x) < 1\}.$$

Let $a < b$ be two real numbers such that $[a, b] \subset [a_F, b_F]$. We point out that we need this strict inclusion because we need the uniform consistency of the distribution $G(\cdot)$ of the truncated r.v. T, which is stated over a compact set (see Remark 6 in Woodroofe, 1985).

2.3 Main results

All along the paper, when no confusion is possible, we denote by C and C' any generic positive constant and by a superscript $(^*)$ any characteristic of the observed data. For a fixed x in $\mathcal{F}$, and $r > 0$, we put $B(x, r) := \{x' \in \mathcal{F} \ : \ d(x', x) < r\}$ and we denote by $K_1 = K\left(d(x, X_1)/h\right)$ and $r_\gamma(u) = \mathbb{E}\left[Y^{-\gamma}|X = u\right]$ for $\gamma = 1, 2$.

2.3.1 Asymptotic properties

Our main first result is the pointwise almost sure convergence. In order to state these latter, we will need some assumptions which are gathered together for easy reference.

(H1) $\mathbb{P}(X \in B(x, s)) =: \phi_x(s) > 0$ for all $s > 0$ and $\lim_{s \to 0} \phi_x(s) = 0$.

(H2) For all $(x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, we have :

$$|r_\gamma(x_1) - r_\gamma(x_2)| \le C\ d^{k_\gamma}(x_1, x_2) \ \text{ for } \ k_\gamma > 0,$$

where $\mathcal{N}_x$ denotes a neighborhood of the fixed $x \in \mathcal{F}$.

(H3) The kernel K is a measurable function which is supported within $(0, 1)$ and such that :

$$0 < C \le K(\cdot) \le C' < \infty.$$

(H4) The smoothing parameter $h := h_n \to 0$ when $n \to +\infty$ and verifies :

$$\frac{n\phi_x(h)}{\log n} \to \infty \text{ as } n \to +\infty.$$

(H5) The inverse conditional moments of the response variable verify :

$$\text{for all } m \geq 2,\ \mathbb{E}[Y^{-m}|X = x] < C < \infty.$$

Remarks on the assumptions.

As in all asymptotic studies in NFDA, our assumptions cover the three fundamental aspects of our study such as the functional data kind, the functional space and the parameters involved in the kernel method (the kernel function and the smoothing parameter). Specifically, Assumption (H1) is the same as that used by Ferraty and Vieu (2006) which is linked to the functional structure of the functional covariate. Then, assumptions (H2) and (H5) characterize the functional space of the nonparametric model. Assumptions (H3) and (H4) concern the kernel $K(\cdot)$ and the smoothing parameter h and are technical conditions. All the considered conditions are usually assumed in this context of nonparametric regression analysis in FDA.

Théorème 2.3.1. *Under assumptions (H1)-(H5), we have :*

$$|\widetilde{r}(x) - r(x)| = O(h^{k_1}) + O(h^{k_2}) + O_{a.s.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right). \tag{2.3}$$

Proof Theorem 2.3.1. The proof is based on the following decomposition

$$\begin{aligned}
\widetilde{r}(x) - r(x) \quad = \quad & \frac{1}{\widehat{\Psi}_D(x)}\left(\widehat{\Psi}_N(x) - \widetilde{\Psi}_N(x)\right) + \frac{1}{\widehat{\Psi}_D(x)}\left(\widetilde{\Psi}_N(x) - \mathbb{E}\left[\widetilde{\Psi}_N(x)\right]\right) \\
& + \frac{1}{\widehat{\Psi}_D(x)}\left(\mathbb{E}\left[\widetilde{\Psi}_N(x)\right] - r_1(x)\right) + \frac{r(x)}{\widehat{\Psi}_D(x)}\left\{\left(\widetilde{\Psi}_D(x) - \widehat{\Psi}_D(x)\right)\right. \\
& \left. + \left(\mathbb{E}\left[\widetilde{\Psi}_D(x)\right] - \widetilde{\Psi}_D(x)\right) + \left(-\mathbb{E}\left[\widetilde{\Psi}_D(x)\right] + r_2(x)\right)\right\}
\end{aligned}$$

where

$$\widehat{\Psi}_N(x) = \frac{\tau_n}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{d(x, X_i)}{h}\right) Y_i^{-1}$$

$$\widehat{\Psi}_D(x) = \frac{\tau_n}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{d(x, X_i)}{h}\right) Y_i^{-2},$$

$$\widetilde{\Psi}_D(x) = \frac{\tau}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} \frac{1}{G(Y_i)} K\left(\frac{d(x, X_i)}{h}\right) Y_i^{-2}$$

and

$$\widetilde{\Psi}_N(x) = \frac{\tau}{n\mathbb{E}[K_1]} \sum_{i=1}^{n} \frac{1}{G(Y_i)} K\left(\frac{d(x, X_i)}{h}\right) Y_i^{-1}.$$

So, Theorem 2.3.1's result is a consequence of the following intermediate results, for which the proofs are given in the Appendix.

Lemme 2.3.1. *Under assumptions (H1) and (H3)-(H5), we have :*

$$\left|\widetilde{\Psi}_D(x) - \mathbb{E}\left[\widetilde{\Psi}_D(x)\right]\right| = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.s.$$

and

$$\left|\widetilde{\Psi}_N(x) - \mathbb{E}\left[\widetilde{\Psi}_N(x)\right]\right| = O\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right), \quad a.s.$$

Lemme 2.3.2. *Under assumptions (H1)-(H4), we have :*

$$\left|\mathbb{E}\left[\widetilde{\Psi}_N(x)\right] - r_1(x))\right| = O\left(h^{k_1}\right),$$

and

$$\left|\mathbb{E}\left[\widetilde{\Psi}_D(x)\right] - r_2(x))\right| = O\left(h^{k_2}\right).$$

Lemme 2.3.3. *Under assumptions (H1) and (H3)-(H5), we have :*

$$\left|\widehat{\Psi}_D(x) - \widetilde{\Psi}_D(x)\right| = O\left(\sqrt{\frac{1}{n}}\right), \quad a.s.$$

and

$$\left|\widehat{\Psi}_N(x) - \widetilde{\Psi}_N(x)\right| = O\left(\sqrt{\frac{1}{n}}\right), \quad a.s.$$

2.3.2 Asymptotic normality

This section is devoted to the study of the asymptotic normality of $\widetilde{r}(x)$. This last property is obtained under the following assumption.

(A1) The concentration property (H1) holds. Moreover, there exists a function $\chi_x(\cdot)$ such that :

$$\text{for all } s \in [0,1], \quad \lim_{r \to 0} \frac{\phi_x(sr)}{\phi_x(r)} = \chi_x(s).$$

(A2) For $\gamma \in \{1,2\}$, the functions $\Psi_\gamma(\cdot) = \mathbb{E}\left[r_\gamma(X) - r_\gamma(x)\middle| d(x,X) = \cdot\right]$ are derivable at 0.

(A3) The kernel function K satisfies (H3) and is a differentiable function on $]0,1[$ where its first derivative function K' is such that :

$$-\infty < C < K'(\cdot) < C' < 0.$$

(A4) The small ball probability satisfies : $n\phi_x(h) \to \infty$, when $n \to +\infty$.

(A5) For $m \in \{1,2,3,4\}$, the functions $g_m(\cdot) = \mathbb{E}[G^{-1}(Y)Y^{-m}|X = \cdot]$ are continuous in a neighborhood of x.

Théorème 2.3.2. *Assume that assumptions (A1)-(A5) hold, then, for any $x \in \mathcal{A}$, we have :*

$$\left(\frac{n\phi_x(h)}{\sigma^2(x)}\right)^{1/2}(\widetilde{r}(x) - r(x) - hB_n(x) - o(h)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty.$$

where $\xrightarrow{\mathcal{D}}$ means the convergence in distribution,

$$B_n(x) = \frac{(\Psi_1'(0) - r(x)\Psi_2'(0))\,\beta_0}{\beta_1 r_2(x)} \tag{2.4}$$

$$\sigma^2(x) = \tau \frac{(g_2(x) - 2r(x)g_3(x) + r^2(x)g_4(x))\,\beta_2}{\beta_1^2}$$

and

$$\mathcal{A} = \{x \in \mathcal{F},\ g_2(x)g_3(x)g_4(x)r_2(x)r(x) \neq 0\}$$

with

$$\beta_0 = K(1) - \int_0^1 (sK(s))'\chi_x(s)ds \text{ and } \beta_j = K^j(1) - \int_0^1 (K^j)'(s)\chi_x(s)ds$$

for $j = 1, 2$.

Proof of Theorem 2.3.2. We write :

$$\widetilde{r}(x) - r(x) = \frac{1}{\widehat{\Psi}_D(x)}\left[D_n - A_n\left(\widehat{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]\right)\right] + A_n$$

where

$$A_n = \frac{1}{\mathbb{E}[\widetilde{\Psi}_D(x)]r_2(x)}[\mathbb{E}[\widetilde{\Psi}_N(x)]r_2(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]r_1(x)]$$

and

$$D_n = \frac{1}{r_2(x)}\left[\left[\widehat{\Psi}_N(x) - \mathbb{E}[\widetilde{\Psi}_N(x)]\right]r_2(x) + \left[\mathbb{E}[\widetilde{\Psi}_D(x)] - \widehat{\Psi}_D(x)\right]r_1(x)\right].$$

Then

$$\widetilde{r}(x) - r(x) - A_n = \frac{1}{\widehat{\Psi}_D(x)} \left[D_n + A_n [\widehat{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]] \right]. \tag{2.5}$$

Therefore, Theorem 2.3.2 is a consequence of the convergence rate of Lemma 2.3.3 and the following intermediate results, for which the proofs are also postponed to the appendix.

Lemme 2.3.4. *Under the hypotheses of Theorem 2.3.2, we obtain :*

$$\left(\frac{n\phi_x(h)}{r_2^2(x)\sigma^2(x)} \right)^{1/2} \left(\left[\widetilde{\Psi}_N(x) - \mathbb{E}[\widetilde{\Psi}_N(x)] \right] r_2(x) + \left[\mathbb{E}[\widetilde{\Psi}_D(x)] - \widetilde{\Psi}_D(x) \right] r_1(x) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Lemme 2.3.5. *Under the hypotheses of Theorem 2.3.2, we have,*

$$A_n = hB_n(x) + o(h).$$

Lemme 2.3.6. *Under the hypotheses of Theorem 2.3.2, we obtain :*

$$\widetilde{\Psi}_D(x) \to r_2(x), \text{ in probability,}$$

and

$$\left(\frac{n\phi_x(h)}{r_2(x)^2\sigma^2(x)} \right)^{1/2} A_n \left(\widetilde{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)] \right) \to 0, \text{ in probability.}$$

2.4 Discussions and conclusions

2.4.1 Some concluding remarks

– *On the incomplete functional data analysis*

It is well known that the biggest motivation of the functional data analysis is the recent technological development on the measuring instruments allowing to collect a data over a thinner discretization grid. However, despite this technological advance the functional data are not totally observed, they observed only over limited grids.

Thus, the incomplete functional data (missing data, error in variables, censored data, surrogate data, truncated data, ...) is more realistic situations for functional data analysis. Therefore, the development of some statistical methods/models adapted to the incomplete functional data is really necessary for a practical purpose. In this paper we have developed a new model over a specific situation of incomplete functional data that is the functional truncated data. The application area of this latter is very large. In fact there are two preliminaries steps to reconstruct a functional variable (recording and interpolation). So, in addition to the truncation causes in the recording step, the basis functions of the interpolation step, usually used in truncated way, can introduce also some truncated functional observation. Thus, in conclusion we can say that, the practical interest of the present contribution stems from the nice features of the studied model (robustness, and scale-independency) and the diversity of the application scope of the functional truncated data.

On the other hand, from theoretical of point of view, the principal gain is the fact that we have established in this new complicated situation (new model and incomplete data case) the same convergence rate of the classical regression in complete data case by using a standard assumptions and a simple arguments.

– *On the practical implementation of the estimator*

As all smoothing approach, the choice of the bandwidth parameter is a fundamental issue for the practical implementation of this method. The asymptotic bias and variance, given in the asymptotic normality result, are very useful in this context. Indeed, it suffices to take the smoothing parameter that is minimize the the leading term of the mean quadratic error expressed by

$$B_n^2(x)h^2 + \frac{\sigma^2(x)}{n\phi_x(h)}$$

Of course, the practical utilization of this method requires the estimation of the unknown quantities $B_n(x)$ and $\sigma^2(x)$. A pilot estimator can be obtained by using the bootstrap procedure. Alternatively, we can avoid this additional calculation by adapting the cross-validation procedure proposed by Hardle,and Marron (1995) and used for the truncated data by Moreira et al. (2016). Specifically, we propose to take a smoothing parameter solution of the following minimization problem

$$\min_h CV(h) \qquad \text{with} \quad CV(h) = \sum_{i=1}^{n} (Y_i - \widetilde{r}_i(X_i))^2$$

where $\widetilde{r}_i$ is the leave-one-out version of $\widetilde{r}$ computed by removing the i-th datum from the initial sample.

As empirical estimate of the quadratic error, this cross-validation rule has shown its performance for various functional kernel estimators (see, for instance Rachdi and Vieu (2007) for the classical regression or Baillo nd Grané (2009) for the local linear model). However, the asymptotic optimality of this approach for the relative error regression is still open question even in the nonfunctional case. This issue is an important prospect of the present work.

2.4.2 Some prospects

- **The kNN method** The KNN method is an alternative smoothing approach which offers an adaptive estimator. In sense that the smoothing parameter is defined from the data. This feature is very intersting in practice. So it would be very important in the next future to study the asymptotic properties of the kNN estimate of the relative error regression. Nothing that the difficulty of this issue resides on the fact that the bandwidth parameter on the kNN method is a random variable.
- **The uniform consistency** This asymptotic property is an indispensable preliminary step in many statistical problems such as the bandwidth selection, the statistical test, additive modeling or multi-step estimation. However, unlike to the multivariate case, the uniform consistency is not a simple extension of the pointwise one, but it requires some additional tools and techniques (see, Ferraty et al.(2010)

for a deeper discussion on this topic). So it would be very interesting to adapt the ideas of this last work for the truncated error relative regression in the next future.

- **The UIB consistency** The uniform in bandwidth (UIB) consistency is very useful asymptotic result it allows to derive the asymptotic properties of the estimate even if the bandwidth parameter is not fixed. This kind of consistency has been studied for various functional nonparametric model by Using a similar ideas to prove the UIB consistency is an important question in the future.
- **The bootstrap procedure** In statistical analysis, the bootstrap procedure is usually used to provide a computational estimates. In the literature of functional statistics, there exist several extensions of this approach (see, for instance for the parametric case, and or of the nonparametric situation). In the two last works, the authors give some functional version of the wild bootstrap procedure. Providing an adaptive version to the relative error regression of this algorithm is an interesting perspective of the present work.

2.5 Appendix

First of all, we state the following Lemma which can be found in the monograph by Ferraty and Vieu (2006).

Lemme 2.5.1. *Let* (Z_i) *be a sequence of i.i.d. centred random variables such that, for all* $m \geq 2$, *there exists* $C_m > 0$, $\mathbb{E}(\mid Z_1^m \mid) < C_m a^{2(m-1)}$, *then*

$$\textit{for all } \varepsilon > 0, \mathbb{P}\left(\left|\sum_{i=1}^{m} Z_i\right| > \varepsilon n\right) \leq 2\exp\left(-\frac{\varepsilon^2 n}{2a^2(1+\varepsilon)}\right)$$

Proof of Lemma 2.3.1. The proofs of both cases are very similar. Thus, we give only the proof of the first one. It suffices to apply the exponential inequality, given in Lemma 2.5.1, on the following sequence (Z_i) :

$$Z_i = \frac{1}{\mathbb{E}[K_1]}\frac{\tau}{G(Y_i)}K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2} - \mathbb{E}\left[\frac{\tau}{G(Y_i)}\frac{1}{\mathbb{E}[K_1]}K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}\right].$$

The main key of the inequality in Lemma 2.5.1 is based on the evaluation of the expectation $\mathbb{E}(\mid Z_1^m \mid)$ for all $m \geq 2$. Indeed, by using the same arguments as in Lemma 6.3 in Ferraty and Vieu (2006, p.65) we get

$$\mathbb{E}[\mid Z_1^m \mid] \leq C\,(\mathbb{E}[K_1])^{-m+1}.$$

Then, we can apply Lemma 2.5.1 with $a = \sqrt{\mathbb{E}[K_1]^{-1}}$. Thus, under the fact that $\mathbb{E}[K_1] = O(\phi_x(h))$ that

$$\mathbb{P}\left(\left|\widetilde{\Psi}_D(x) - \mathbb{E}\left[\widetilde{\Psi}_D(x)\right]\right| > \varepsilon_0\sqrt{\log n/n\mathbb{E}[K_1]}\right) \leq 2\exp\left(-C\varepsilon_0^2\log n\;\phi_x(h)\right) \leq 2n^{-C\varepsilon_0^2}$$

and in the same manner

$$\mathbb{P}\left(\left|\widetilde{\Psi}_N(x) - \mathbb{E}\left[\widetilde{\Psi}_N(x)\right]\right| > \varepsilon_0\sqrt{\log n/n\mathbb{E}[K_1]}\right) \leq 2\exp\left(-C\varepsilon_0^2\log n\phi_x(h)\right) \leq 2n^{-C\varepsilon_0^2}.$$

Finally, an appropriate choice of ε_0 and the Borel-Cantelli Lemma's use complete the proof.
Proof of Lemma2.3.2. Similarly to the previous Lemma 2.3.1, we give only the proof for the bias of $\widetilde{\Psi}_D(x)$. In fact, to obtain the second equation one may follows the same lines. Now, let $\mathbf{f}(\cdot|\cdot)$ (resp. $\mathbf{f}^*(\cdot|\cdot)$) the conditional density of Y given X (resp. the conditional density of Y given X under the left-truncation condition). So, by the truncation effect, we have :

$$\begin{aligned}\mathbb{E}\tilde{\Psi}_D(x,t)] &= \frac{\tau}{\mathbb{E}[K_1]}\int\int\frac{K\left(\frac{d(x,u)}{h}\right)y^{-2}}{G(y)}\mathbf{f}^*(y|u)\,dy\,d\mathbf{P}_x^X(u)\\ &= \frac{1}{\mathbb{E}[K_1]}\int\int K\left(\frac{d(x,u)}{h}\right)y^{-2}\mathbf{f}(y|u)\,dy\,d\mathbf{P}_x^X(u)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h}\right)Y_1^{-2}\right] \\
&= \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h}\right)\mathbb{I}_{B(x,h)}r_2(X_1)|\right].
\end{aligned}
$$

Then,

$$
\begin{aligned}
\mid \mathbb{E}[\widetilde{\Psi}_D(x)] - r_2(x) \mid &\leq \frac{\tau}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1\left(\frac{d(x,X_i)}{h}\right)1\!\!1_{B(x,h)} \mid r_2(X_1) - r_2(x) \mid\right] \\
&\leq C_1 h^{k_2}
\end{aligned}
$$

and in the same way, we have

$$
\begin{aligned}
\mid \mathbb{E}\widetilde{\Psi}_N(x)] - r_1(x) \mid &\leq \frac{\tau}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1\left(\frac{d(x,X_i)}{h}\right)1\!\!1_{B(x,h)} \mid r_1(X_1) - r_1(x) \mid\right] \\
&\leq C_1 h^{k_1}.
\end{aligned}
$$

The proof of this Lemma is then finished.

Proof of Lemma 2.3.3. We give only the proof of the first case, the second one is similar.

$$
\begin{aligned}
\left|\widehat{\Psi}_D(x) - \widetilde{\Psi}_D(x)\right| &\leq \left[\frac{\mid \tau_n - \tau \mid}{G_n(a_F)} + \tau\frac{\sup_{y\geq a_F} \mid G_n(y) - G(y) \mid}{G_n(a_F)G(a_F)}\right] \\
&\quad \times \frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}.
\end{aligned}
$$

It is shown in Theorem 3.2 by He and Yang(1998) that

$$
\mid \tau_n - \tau \mid = O_{a.s.}(n^{-\frac{1}{2}}),
$$

while Remark 6 in Woodroofe (1985) gives

$$
|G_n(a_F) - G(a_F)| = O_{a.s.}(n^{-\frac{1}{2}}).
$$

Furthermore, by following the same lines as for the proof of Lemma 2.3.1, we show that

$$\frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^{n}\left[K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}-\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}\right]\right]=O_{a.s.}\left(\sqrt{\frac{\log n}{n\phi_x(h)}}\right)$$

and by combining the latter with the fact that

$$\frac{1}{n\mathbb{E}[K_1]}\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}\right]-r_2(x)=o(1)$$

we get

$$\frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{h}\right)Y_i^{-2}=O_{a.s.}(1).$$

So, we have

$$\left|\widehat{\Psi}_D(x)-\widetilde{\Psi}_D(x)\right|=O_{a.s.}(n^{-\frac{1}{2}})$$

and by the same arguments we obtain

$$\left|\widehat{\Psi}_N(x)-\widetilde{\Psi}_N(x)\right|=O_{a.s.}(n^{-\frac{1}{2}})$$

which is the claimed result.

Proof of Lemma 2.3.4. We write

$$\sqrt{n\phi_x(h)}(r_2^2(x)\sigma(x))^{-1}\left[\left(\widetilde{\Psi}_D(x)-\mathbb{E}[\widetilde{\Psi}_D(x)]\right)r_1(x)\right.$$
$$\left.-\left(\widetilde{\Psi}_N(x)-\mathbb{E}[\widetilde{\Psi}_N(x)]\right)r_2(x)\right]=\frac{S_n}{r_2^2(x)\sigma(x)}$$

with

$$S_n=\sum_{i=1}^{n}(L_i(x)-\mathbb{E}[L_i(x)]),$$

where,

$$L_i(x) := \frac{\sqrt{n\phi_x(h)}}{n\mathbb{E}[K_1]}\frac{\tau}{G(Y_i)}K_i\left(r_1(x)Y_i^{-2} - r_2(x)Y_i^{-1}\right). \tag{2.6}$$

So, it suffices to show the asymptotic normality of S_n to achieve this lemma's proof. This last is reached by applying the Lyapounov central limit Theorem on $L_i(x)$ which is based on

$$\frac{\sum_{i=1}^{n}\mathbb{E}\left[|L_i(x) - \mathbb{E}\left[L_i(x)\right]|^{2+\delta}\right]}{\left(var\left(\sum_{i=1}^{n}L_i(x)\right)\right)^{(2+\delta)/2}} \to 0, \quad \text{for some } \delta > 0. \tag{2.7}$$

Clearly,

$$\begin{aligned} var\left(\sum_{i=1}^{n}L_i(x)\right) &= n\phi_x(h)\left(var\left[\widetilde{\Psi}_N(x)\right]r_2^2(x) + var\left[\widetilde{\Psi}_D(x)\right]r_1^2(x)\right. \\ &\left. -2r_1(x)r_2(x)Cov(\widetilde{\Psi}_N(x), \widetilde{\Psi}_D(x))\right). \end{aligned}$$

By a standard arguments we obtain

$$n\phi_x(h)var\left[\widetilde{\Psi}_D(x)\right] \to \tau\frac{g_4(x)\beta_2}{\beta_1^2},$$

and

$$n\phi_x(h)var\left[\widetilde{\Psi}_N(x)\right] \to \tau\frac{g_2(x)\beta_2}{\beta_1^2},$$

and

$$n\phi_x(h)Cov\left(\widetilde{\Psi}_N(x), \widetilde{\Psi}_D(x))\right) \to \tau\frac{g_3(x)\beta_2}{\beta_1^2}.$$

It follows that,

$$var\left(\sum_{i=1}^{n}L_i(x)\right) = r_2^2(x)\sigma(x) + o(1).$$

Therefore, to complete the proof of this Lemma, it suffices to show that the numerator of (2.7) converges toward 0. For this, we use the C_r-inequality (cf. Loève (1963), page 155) to show that :

$$\sum_{i=1}^{n} \mathbb{E}\left[\left|L_i(x) - \mathbb{E}\left[L_i(x)\right]\right|^{2+\delta}\right] \leq C \sum_{i=1}^{n} \mathbb{E}\left[|L_i(x)|^{2+\delta}\right] + C' \sum_{i=1}^{n} \left|\mathbb{E}\left[L_i(x)\right]\right|^{2+\delta}. \tag{2.8}$$

Recall that, for all $j > 0$, $\mathbb{E}\left[K_1^j\right] = O(\phi_x(h))$, then, since assumption (H5), we get

$$\sum_{i=1}^{n} \mathbb{E}\left[|L_i(x)|^{2+\delta}\right] \quad \leq \quad C(n\phi_x(h))^{-\delta/2}\Big(\mathbb{E}\left[K_1^{2+\delta}\right]/\phi_x(h)\Big) \to 0.$$

Similarly, the second term of (2.8) becomes

$$\sum_{i=1}^{n} \left|\mathbb{E}\left[L_i(x)\right]\right|^{2+\delta} \quad \leq \quad Cn^{-\delta/2}\left(\phi_x(h)\right)^{1+\delta/2} \to 0,$$

which completes the proof.

<u>Proof of Lemma 2.3.5.</u> By standard arguments we show that

$$\mathbb{E}\left[\widetilde{r}(x)\right] = \frac{\mathbb{E}\left[\widetilde{\Psi}_N(x)\right]}{\mathbb{E}\left[\widetilde{\Psi}_D(x)\right]} + O\left(\frac{1}{n\phi_x(h)}\right).$$

So, it suffices to evaluate

$$\mathbb{E}\left[\widetilde{\Psi}_N(x)\right] \quad \text{and} \quad \mathbb{E}\left[\widetilde{\Psi}_D(x)\right].$$

Indeed, by the same ideas as in the proof of Lemma 2.3.2, we have

$$\mathbb{E}[\widetilde{\Psi}_D(x)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1\mathbb{E}[Y_1^{-2}|X_1]\right] \quad \text{and} \quad \mathbb{E}[\widetilde{\Psi}_N(x)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1\mathbb{E}[Y_1^{-1}|X_1]\right].$$

Now, by using same arguments as those used by Ferraty et al. (2007), for the regression operator, we show that

$$\mathbb{E}\left[\widetilde{\Psi}_D(x)\right] = r_2(x) + h\Psi_2'(0)\left[\frac{K(1) - \int_0^1 (uK(u))'\chi_x(u)du}{K(1) - \int_0^1 K'(u)\chi_x(u)du}\right] + o(h),$$

and

$$\mathbb{E}\left[\widetilde{\Psi}_N(x)\right] = r_1(x) + h\Psi_1'(0)\left[\frac{K(1) - \int_0^1 (uK(u))'\chi_x(u)du}{K(1) - \int_0^1 K'(u)\chi_x(u)du}\right] + o(h).$$

Finally,

$$A_n = \frac{\mathbb{E}\left[\widetilde{\Psi}_N(x)\right]}{\mathbb{E}\left[\widetilde{\Psi}_D(x)\right]} - r(x) = hB_n(x) + o(h).$$

Proof of Lemma 2.3.6. For the first limit we have, by Lemmas 2.3.5 and 2.3.2's results, that

$$\mathbb{E}\left[\widetilde{\Psi}_D(x) - r_2(x)\right] \to 0 \quad \text{and} \quad var\left[\widetilde{\Psi}_D(x)\right] \to 0.$$

Hence,

$$\widetilde{\Psi}_D(x) - r_2(x) \to 0, \text{ in probability.}$$

Next, for the last needed convergence, we obtain by the same manner

$$\mathbb{E}\left[\left(\frac{n\phi_x(h)}{r_1(x)^2\sigma^2}\right)^{1/2} A_n\left(\widetilde{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]\right)\right] = 0$$

and

$$var\left[\left(\frac{n\phi_x(h)}{r_1(x)^2\sigma^2}\right)^{1/2} A_n\left(\widetilde{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]\right)\right] = O(A_n^2) = O(h^2) \to 0.$$

It follows then that

$$\left(\frac{n\phi_x(h)}{r_1(x)^2\sigma^2}\right)^{1/2} A_n\left(\widetilde{\Psi}_D(x) - \mathbb{E}[\widetilde{\Psi}_D(x)]\right) \to 0, \text{ in probability.}$$

Bibliographie

[1] Attouch, M., Laksaci, A. and Ould-Said, E. (2009). Asymptotic distribution of robust estimator in nonparametric functional models, *Comm. in Statist. Theory and Methods,* **38**, 1317-1335.

[2] Aneiros, G. Bongiorno, E. G. Cao, R. and Vieu, P. (2017) *Functional Statistics and Related Fields,* Springer International Publishing, Cham, Switzerland.

[3] Azzedine, N., Laksaci, A. and Ould-Said, E. (2008). On the robust nonparametric regression estimation for functional regressor, *Statist. Probab. Lett.* **78**, 3216-3221.

[4] Bongiorno, E. G., Goia, A., Salinelli, E. and Vieu, P. (2014) *Contributions in infinite-dimensional statistics and related topics*, Esculapio, Bologna.

[5] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference*, **147**, 1-23.

[6] Demongeot, J., Hamie, A., Laksaci, A., and Rachdi, M. (2016). Relative-error prediction in nonparametric functional statistics : theory and practice. *J. Multivariate Anal.* **146**, 261–268

[7] Derrar, S., Laksaci, A. and Ould-Said, E. (2015). On the nonparametric estimation of the functional ψ-regression for a random left-truncation model. *J. Stat. Theory Pract.* **9** , 823–849

[8] Ferraty, F. and Vieu, P. (2006) *Nonparametric Functional Data Analysis. Theory and Practice,* Springer-Verlag, New York,

[9] Gheriballah, A. Laksaci, A. and Sekkal, S. (2013) Nonparametric M-regression for functional ergodic data. *Statist. Probab. Lett.* , **83**, 902–908.

[10] He, S. and Yang, G. (1998). Estimation of the truncation probability in the random truncation model. *Annals of Statistics*, **26**, 1011–1027.

[11] Helal, N., and Ould-Said, E. (2016). Kernel conditional quantile estimator under left truncation for functional regressors. Opuscula Math. **36**, 25–48.

[12] Horrigue, W. and Ould-Said, E. (2011). Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Oper. Stoch. Equ.*, **19** , 131–156.

[13] Horrigue, W. and Ould-Said, E. (2014). Nonparametric regression quantile estimation for dependant functional data under random censorship : Asymptotic normality. Comm. in Statistics. Theory and Methods., **44**, No. 20, 4307–4332.

[14] Hsing, T. and Eubank, R. (2015). Theoretical foundations of functional data analysis, with an introduction to linear operators. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester.

[15] Ling, N., Liu, Y. and Vieu, P. (2016). Conditional mode estimation for functional stationary ergodic data with responses missing at random. *Statistics* **50**, 991–1013.

[16] Ling, N., Liu, Y. and Vieu, P.(2015). Nonparametric regression estimation for functional stationary ergodic data with missing at random. *J. Statist. Plann. Inference* **162**, 75–87.

[17] Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices of the Royal Astronomical Society*, **155**, 95–118.

[18] Ould-Said, E. and Lemdani, M. (2006). Asymptotic properties of a nonparametric regression function estimator with randomly truncated data. *Ann. Inst. Statist. Math.*, **58**, 357–378.

[19] Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.*, **21**, 146–156.

[20] Wang, J., Liang, H.Y. and Fan, G. (2012a). Asymptotic properties for an M-estimator of the regression function with truncation and dependent data. *Journal of the Korean Statistical Society*, **41**, 35–367.

[21] Wang, J., Liang, H.Y. and Fan, G. (2012b). Local M-estimation of nonparametric regression with left-truncated and dependent data. *Scientia Sinica Mathematica*, **42**, 995–1015.

[22] Woodroofe, M. (1985). Estimating a distribution function with truncated data. *The Annals of Statistics*, **13**, 163–177.

[23] Zhang, J. (2014). *Analysis of variance for functional data.* Monographs on Statistics and Applied Probability, 127. CRC Press, Boca Raton, FL.

CHAPITRE 3

Asymptotic properties of the functional relative regression in the dependent case

La vérité scientifique a pour signe
la cohérence et l'efficacité.
La vérité poétique a pour signe la beauté.
Aimé Cesaire

3.1 The model

Let $(X_1, Y_1), \dots (X_n, Y_n)$ be n strongly mixing process, identically distributed as (X, Y) which is a random pair valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space, d denoting the semi-metric. For any x in $\mathcal{F}$, we define the relative error regression

$$\theta(x) = \frac{\mathbb{E}\left[Y^{-1} \mid X = x\right]}{\mathbb{E}\left[Y^{-2} \mid X = x\right]}. \tag{3.1}$$

Using the kernel estimate of the classical regression operator, we obtain as kernel estimator of $\theta(x)$, denoted by, $\widehat{\theta_x}$

$$\widehat{\theta_x} := \frac{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))Y_i^{-1}}{\sum_{i=1}^{n} K(h^{-1}d(x, X_i))Y_i^{-2}},$$

where K is a kernel function and $h := h_n$ (to simplify the notations) is a sequence of positive real numbers which goes to zero as n goes to infinity.

Our main purpose is to study the almost completely consistency of $\widehat{\theta_x}$ to θ_x in the strong mixing property, defined as follows

Définition 3.1.1. Let $\mathcal{F}_1^k$ the σ-algebra generated by $(X_1, Y_1), \dots, (X_k, Y_k)$ and $\mathcal{F}_{k+n}^{\infty}$ that generated by $(X_{k+n}, Y_{k+n}), \dots$. Set now, for any $n \geq 1$,

$$\alpha(n) = \sup_{A \in \mathcal{F}_1^k} \sup_{B \in \mathcal{F}_{k+n}^{\infty}} \left\{ \left| \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \right| \right\}. \tag{3.2}$$

The process $(X_n, Y_n)_{n \geq 1}$ is said to be strongly mixing if

$$\lim_{n \to \infty} \alpha(n) = 0. \tag{3.3}$$

There exists many processes fulfilling the strong mixing property. We quote, here, the usual ARMA processes which are geometricaly strongly mixing, *i.e.,* there exists $\rho \in (0,1)$ and $a > 0$ such that, for any $n \geq 1$, $\alpha(n) \leq a\rho^n$ (see, *e.g.*, Jones (1978). The threshold models, the EXPAR medols (see, Ozaki (1979)), the simpla ARCH models (see Engle (1982)), their extensions the GARCH models (see Bollerslev (1986)) and the bilinear Markovian models are geometricaly strongly mixing under some general ergodicity conditions.

3.2 Main results

All along the paper, when no confusion will be possible, we will denote by C or/and C' some generic constant in $\mathbb{R}^{*+}$. In the following, we will denote, for all $x \in \mathcal{F}$, for all $i = 1, \ldots, n$, by $K_i(x) = K(h^{-1}d(x, X_i))$, and we pose $\widehat{\Psi}(x,t) = \dfrac{\widehat{\Psi}_N}{\widehat{\Psi}_D(x)}$ with

$$\widehat{\Psi}_D = \frac{1}{n\mathbb{E}[K_1(x)]}\sum_{i=1}^{n} K_i(x)Y_i^{-2} \text{ and } \widehat{\Psi}_N = \frac{1}{n\mathbb{E}[K_1(x)]}\sum_{i=1}^{n} K_i(x)Y_i^{-1}.$$

In order to establish the pointwise almost complete convergence of the estimator $\widehat{\theta_x}$ for all fixed point x in $\mathcal{F}$, we consider the following assumptions :

(H1) $\mathbb{P}(X \in B(x,h)) = \phi_x(h) > 0$,

(H2) The functions $g_l(x) = \mathbb{E}[Y^{-l}|X = x]$ satisfies Hölder's condition

$$\forall\, x_1, x_2 \in \mathcal{N}_x,\ |g_l(x_1) - g_l(x_2)| \leq C' d^k(x_1, x_2), \qquad k > 0.$$

(H3) The conditional inverse moment of the response variables of $k = 1, \ldots 4.$ are finished

(H4) $(X_i, Y_i)_{i\in N}$ is an α-mixing sequence whose coefficients satisfy

$$\exists\, a > 0,\ \exists\, C > 0 :\ \forall n \in N \alpha(n) \leq Cn^{-a}.$$

(H5) $0 < \sup_{i\neq j} \mathbb{P}\left((X_i, X_j) \in B(x,h) \times B(x,h)\right) = O\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right)$.

(H6) K is a continuous function with support $[0,1]$ such that $0 < C < K(t) < C' < \infty$.

(H7) $\lim_{n\to\infty} h = 0$ and $\exists \eta > 0$, such that, $Cn^{\frac{3-a}{a+1}+\eta} \leq \phi_x(h) \leq C' n^{\frac{1}{1-a}}$, with $a > \frac{5+\sqrt{17}}{2}$.

The convergence rate is given in the following theorem :

Théorème 3.2.1. *Under the assumptions (H1)-(H7), we have :*

$$\widehat{\theta}(x) - \theta(x) = O(h^k) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right), \; p.co. \tag{3.4}$$

Proof of the theorem We write

$$\widehat{r}(x) = \frac{\widehat{\Psi}_N}{\widehat{\Psi}_D} \tag{3.5}$$

$$\widehat{\Psi}_N = \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h_n}\right)\right]} \sum_{i=1}^{n} Y_i K\left(\frac{d(x,X_i)}{h}\right)$$

and

$$\widehat{\Psi}_D = \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h_n}\right)\right]} \sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h}\right).$$

We consider the following decomposition :

$$\widehat{r}(x) - \theta(x) = \frac{\widehat{\Psi}_N - \theta(x)}{\widehat{\Psi}_D} + \left(f(x) - \widehat{\Psi}_D\right)\frac{\theta(x)}{\widehat{\Psi}_D}. \tag{3.6}$$

Thus, our Theorem is a direct consequence of the following lemmas

Lemme 3.2.1. *Under the conditions of the Theorem we have*

$$\widehat{\Psi}_N - \mathbb{E}[\widehat{\Psi}_N] = O\sqrt{\frac{\log n}{n\varphi_x(h)}} \tag{3.7}$$

Lemme 3.2.2. *Under the conditions of the Theorem we have :*

$$\mathbb{E}[\widehat{\Psi}_N] - g_1(x) = O\left(h^k\right) \tag{3.8}$$

Lemme 3.2.3. *Under the conditions of the Theorem we have*

$$\mathbb{E}[\widehat{\Psi}_D] - g_2(x) = O\left(h^k\right) \tag{3.9}$$

Lemme 3.2.4. *Under the conditions of the Theorem we have*

$$1 - \widehat{\Psi}_D - g_2(x) = O\left[h^k + \sqrt{\frac{\log n}{n\varphi_x(h)}}\right] \tag{3.10}$$

Lemme 3.2.5. *Under the conditions of the Theorem we have :*

$$\exists \delta > 0, \quad \textit{tel que} \quad \sum \mathbb{P}\left[|\widehat{\Psi}_D| < \delta\right] < \infty \tag{3.11}$$

3.2.1 Proof of the intermediates results

Observe that the bias term does not influenced by the dependence condition, then, its proof is the same as in the first part. Thus, we focus only in the proof of the dispersion term. For the latter, because of the similarity between the proof of the two lemmas we set

$$\widehat{\psi}(x) = \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h}\right)\right]} \sum Y_i^{-l} K\left(\frac{d(x,X_i)}{h-n}\right) \text{ avec } l = 1 \text{ or } 2 = 1.$$

which allows to proof the two lemmas by using together. Indeed, to do that we apply the a Fuk Nagave inequality on the variables

$$\Delta_i = Y_i^l K\left(\frac{d(x,X_i)}{h}\right) - \mathbb{E}\left[Y_i^l K\left(\frac{d(x,X_i)}{h}\right)\right].$$

The latter is based on the asymptotic variance :

$$\begin{aligned}
S_n^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov\left(\Delta_i, \Delta_j\right)| \\
&= \sum_{i=1}^{n}\left[\sum_{i\neq j} Cov(\Delta_i\Delta_j) + Cov(\Delta_i, \Delta_j)\right] \\
&= \sum_{i=1}^{n}\left[\sum_{i\neq j} Cov(\Delta_i, \Delta_j) + Var\Delta_i\right] \\
&= S_n^{2*} + nVar\Delta_1.
\end{aligned}$$

concerning S_n^{2*}, we the techniques of Masry (1986), and we split the sum into subsets

$$E_1 = \Big\{(i,j) \quad \text{such that} \quad 1 \leq |i-j| \leq m_n\Big\}$$

et

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

where $m_n \to \infty,$ as $n \to \infty$. Let $J_{1,n}$ and $J_{2,n}$ and

$$J_{1,n} = \sum_{E_1} \Big|Cov\left(\Delta_i(x), \Delta_j(x)\right)\Big| \leq \sum_{E_1} \Big|E\left[K_i(x)K_j(x)\right] - E^2\left[K_1(x)\right]\Big|.$$

Under (H1), (H6) and (H7) we have

$$J_{1,n} \leq Cnm_n\phi_x(h)\left(\left(\frac{\phi_x(h)}{n}\right)^{1/a} + \phi_x(h)\right).$$

On r E_2, we use the Davydov-Rio's inequality (Rio, 2000, p. 87). Thus, for $i \neq j$, we have

$$|Cov\left(K_i(x), K_j(x)\right)| \leq C\alpha(|i-j|).$$

Therefore,

$$J_{2,n} = \sum_{E_2} |Cov\,(K_i(x), K_j(x))\,| \leq n^2 m_n^{-a}.$$

Taking $m_n = \left(\dfrac{\phi_x(h)}{n}\right)^{-1/a}$ we get

$$S_n^{2*} = J_{1,n} + J_{2,n} = O(n\phi_x(h)). \tag{3.12}$$

Concerning the variance term, we show under (H1), that

$$Var\,[\Delta_1(x)] \leq C\big[\mathbb{E}(K_1^2) + \mathbb{E}^2(K_1)\big] \leq C\big[\phi_x(h) + (\phi_x(h))^2\big]. \tag{3.13}$$

Finally, from (3.12) and (3.13) we obtain

$$S_n^2(x) = O\left(n\phi_x(h)\right). \tag{3.14}$$

Thus, we get from Fuk Nagave inequality

$$\begin{aligned}
\mathbb{P}\left[\left|\mathbb{E}\left[\widehat{\psi}(x)\right] - \widehat{\psi}(x)\right| > \varepsilon\right] &= \mathbb{P}\left[\left|\frac{1}{n\varphi_x(h)}\sum_{i=1}^{n}\Delta_i\right| > \varepsilon\right] \\
&= \mathbb{P}\left[\left|\sum_{i=1}^{n}\Delta_i\right| > \frac{4(\varepsilon n\varphi_x(h))}{4}\right] \\
&\leq 4\left(1 + \frac{\varepsilon^2 n^2 \varphi_x^2(h)}{n\varphi_x(h)16r}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{8r}{\varepsilon n\varphi_x(h)}\right)^{a+1}.
\end{aligned}$$

for $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}$

$$\mathbb{P}\left[\left|\mathbb{E}\left[\widehat{\psi}(x)\right] - \widehat{\psi}(x)\right| > \epsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}\right] \leq 4\left(1 + \frac{\epsilon_0^2(\log n)n^2(\varphi_x)^2(h)}{n^2(\varphi_x(h))^2 16r}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{8r(n\varphi_x(h))^{\frac{1}{2}}}{\epsilon_0 n\varphi_x(h)(\log n)^{\frac{-1}{2}}}\right)$$

$$\leq 4\left(1 + \frac{\epsilon_0^2 \log n}{16r}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{8r}{\epsilon_0}^{a+1}\right)(n\varphi_x(h)\log n)^{-(\frac{a+1}{2})}.$$

It sufices to take $r = C\left(\log(n)\right)^2$, the first term is of order

$$A_1 = \left(1 + \frac{\epsilon_0^2 \log n}{16r}\right)^{\frac{-r}{2}} \leq ce^{\left(\frac{-\epsilon_0^2 \log n}{32}\right)} \tag{3.15}$$

while the second term

$$A_2 \leq c\varepsilon^{-(a+1)} n^{-(\frac{a+1}{2})+1+ab} \left[\varphi_x(h)\right]^{(\frac{a+1}{2})}$$

Under (H8) we have

$$A_2 \leq cn^{-1-\frac{a}{2}(-2b+p\theta)} \tag{3.16}$$

The equations (3.15) and (3.16) imply that there exits ε_0, such that

$$\sum_{n=1} \mathbb{P}\left[\left|\mathbb{E}\left[\widehat{\psi}(x)\right] - \widehat{\psi}(x)\right| > \varepsilon_0 \sqrt{\frac{\log n}{n\varphi_x(h)}}\right] < \infty$$

Bibliographie

[1] Billingsley, Patrick. (1968)*Convergence of probability measures.* John Wiley, New York.

[2] Bosq, D. and Lecroutre, J. P. (1987)*Théorie de l'estimation fonctionnelle.* Economica.

[3] Carbon, M. (1988)*Inégaliés de grandes déviations dans les processus. Application à l'estimation fonctionnelle.* Thèse de Doctorat de l'Université de Paris VI, France.

[4] Chernoff, H. (1964) Estimation of the mode. *Ann. Inst. Statist. Math.*,**16**, 31-41.

[5] Collomb, G., Hardle, W. and Hassani, S. (1987) a note on prediction via estimation of the conditional mode function. *J. Statist. Plann. and Inference.*, **15**, 227-236.

[6] Diebolt, J. and Guégan, D. (1993) Tail behaviour of the stationary density of general nonlinear autoregressive processes of order I. *J. App. Probab.*, **30**, 315-329.

[7] Doob, J. L. (1953) *Stochastics processes.* Wiley, New York.

[8] Eddy, W. F. (1980) Optimum kernel estimators of the mode.*Ann. Statist.*, **8**, 870-882.

[9] Eddy, W. F. (1982) The asymptotic distribution of kernel estimaors of the mode. Z. *Wahrsch. Verw. Gebiete.*, **59**, 279-290.

[10] Engle, R.F. (1982) Autoregressive conditionam heteroscedasticity with estimates of the variance of united Kingdom inflation. *Econometrics,* **50**(40), 987-1007.

[11] Gannoun, A. and Louani, D. (1996) Asymptotic properties of kernel estimators of the mode under censoring. Technical Report no.96-04, University of Montpellier II.

[12] Jones, D. A. (1978) Nonlinear autoregressive processes. *Proc. Roy. Soc. London A,* **360** .

[13] Konakov, V. D. (1974) On the assymptotic normality of the mode of multidimensional distribution. *Theory of Probab. Appl.*, **19**, 794-799 .

[14] Masry, E. (1986), Recursive probability density estimation for weakly dependent stationary processes. *IEEE, Trans. Infor. Theory*, **32**, 254-267 .

[15] Nadaraya, E. N. (1965) On nonparametric estimares of density functions and regression curves. *Theory Probab. Appl.*,**10**, I , 186-190 .

[16] Ould-Said, E. (1993) Estimation non paramétrique due mode conditionnel. Application à la prévision. *C. R. Acad. Sci. Paris,* Série I, t **316**, 943-947.

[17] Ould-Said, E. (1997) A note on ergodic processes prediction via estimation of the conditional mode function. *Scandinavian J. of statist.*, **24**, 231-239 .

[18] Ozaki, T. (1979) Nonlinear time series models for nonlinear random vibration. Technical Report. Univ. of Manchester .

[19] Parzen, E. (1962) On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33,** 10665-1076 .

[20] Rio, E. (1994) Covariate inequalities for strongly mixing process. *Ann. I.H.P.* **23**, 587-597 .

[21] Romano, J. P. (1988) On weak convergence and optimality of kernel density estimates of the mode. *Ann. Statist.*, **16**, 629-647 .

[22] Rouassas, G. G. and Ionnides, D. (1987) Moment inequalities for mixing sequences of random variables.*Stochastic Analysis,* **5**, 61-120 .

[23] Rouassas, G. G.(1990) asymptotic normality of the Kernel estimate under dependence condition : Application to hazard rate. J. *Statist. Plan. Infer.*, **25**, 81-104 .

[24] Samanta. M. (1973) Nonparametric estimation of the mode of a multivariate density. *South African Statist.* **J.**, **7**, 109-117 .

CHAPITRE 4

Numerical study

L'imagination est plus importante
que la connaissance.
Albert Einstein

Sommaire

Ce chapitre fait l'objet d'une publication acceptée par " JOURNAL OF NONPARAMETRIC STATISTICS", dont on va exposer la partie expérimentale.

Application : The main goal of this section is to evaluate the efficiency of our approach in function to the truncation level. Firstly, we compare its finite sample performance to the classical regression, $R(x) = \mathbb{E}[Y|X = x]$, over a three different truncation levels. In particular, we highlight the scale-independency property of the relative error regression (R.E.R). Secondly, in the last part of this numerical study, we illustrate our asymptotic normality result. For all these purposes, we generate our data by the following model

$$Y = r(x) + \sigma(x)\varepsilon$$

where $\varepsilon \hookrightarrow \mathcal{N}(0, .5)$. The operators $r(\cdot)$ is defined by

$$r(X) = 5\left(\frac{1}{1 + \int X^2(t)dt}\right).$$

For the scale operator $\sigma(\cdot)$, we consider three types that are

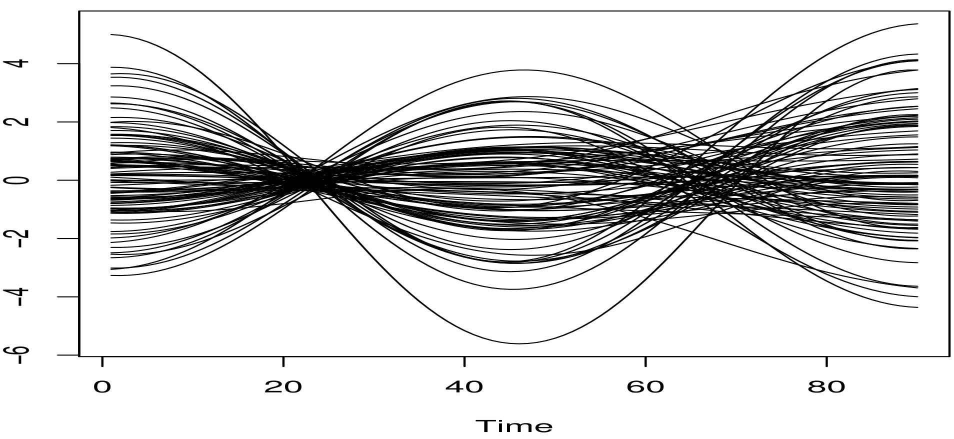
$$\begin{array}{cc} \text{Homoscedastic case} & \sigma(X) \equiv 1 \\ \text{Logarithmic scale} & \sigma(X) = 2\log\left(\int X^2(t)dt\right) \\ \text{Exponential scale} & \sigma(X) = \exp\left(\int X^2(t)dt\right). \end{array}$$

The functional regressors $X_i(t)$ are defined, for $t \in [0, 1]$, by :

$$X_i(t) = 3W_i \sin(2\pi t) + \eta_i t, \text{ where } W_i \hookrightarrow N(0, 0.5) \text{ and } \eta_i \hookrightarrow N(0, 1).$$

A sample of these curves are plotted in the following figure.

Noting that, that unlike to complete and censored data, in this truncated context the explanatory variable X is kept only if the response variable Y is observed. In this simulation



study, we generate the truncating random variable T_i from the Gaussian distribution $N(m, 1)$ with a varied values of m in order to get a different rate of truncation. Precisely, in function of m we consider three truncation levels (weak 5%, mean 30% and strong truncation 65%).

Now, the two estimators

$$\widetilde{r}(x) := \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i)K(h^{-1}d(x, X_i))Y_i^{-1}}{\sum_{i=1}^{n} G_n^{-1}(Y_i)K(h^{-1}d(x, X_i))Y_i^{-2}} \quad \text{and} \quad \widehat{r}(x) := \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i)K(h^{-1}d(x, X_i))Y_i}{\sum_{i=1}^{n} G_n^{-1}(Y_i)K(h^{-1}d(x, X_i))}$$

are computed by using the observation (X_i, Y_i, T_i), $i = 1, \ldots, n$ such that $Y_i \geq T_i$.

For this comparison study, we put $n = 100$, we choose the optimal bandwidth h by the cross-validation method on the k-nearest neighbors and we use the semi-metric defined by the L_2-distance between the first derivatives of the curves. Finally, we point out that, we have simulated by the quadratic kernel K. The results of the comparison study are summarized in the following Table where we give the mean squared errors

$$MSE(RER) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widetilde{r}_i(X_i))^2 \quad \text{and} \quad MSE(CR) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widetilde{r}_i(X_i))^2$$

where $\widetilde{r}_i$ ($\widehat{r}_i$) is the leave-one-out version of $\widetilde{r}$ (resp. $\widehat{r}$) computed by removing the i-th datum from the initial sample.

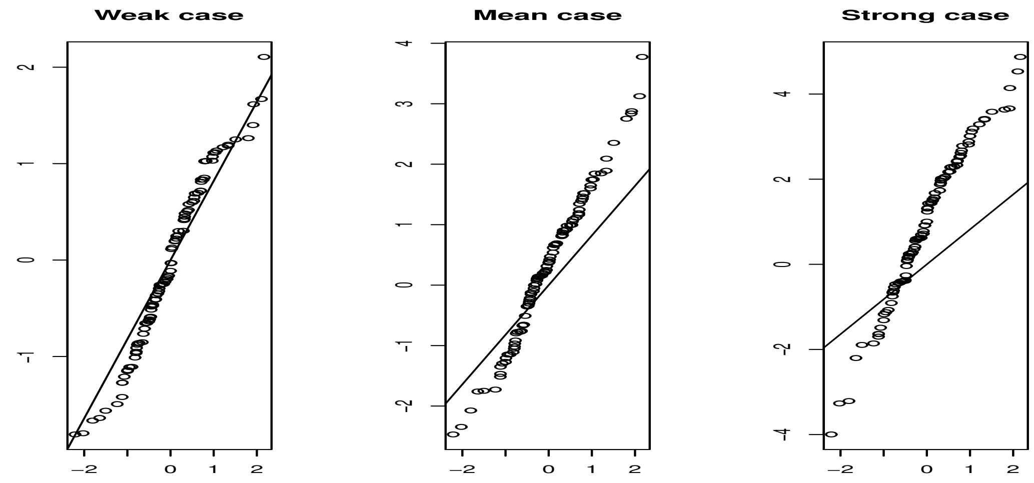
Type of scale	Truncation level	MSE(CR)	MSE(RER)
Homoscedastic case	Weak	0.34	0.65
	Mean	1.90	1.71
	Strong	3.78	2.92
Logarithmic scale	Weak	1.42	1.19
	Mean	3.67	2.21
	Strong	6.94	3.12
Exponential scale	Weak	2.94	2.03
	Mean	10.52	4.17
	Strong	14.11	5.37

Table 1 The MSE-error of the estimates .

From this table, we can find that both estimators have a reasonable finite sample performance for lower truncation level and the two estimators are equivalents in this situation. However, the performance of both estimators are strongly affected by the truncation rate. Moreover, the prediction quality of both models depends on the scale type, but, the R.E.R. model is more stable than the C.R. method in sense that the variability of the MSE-error of the R.E.R. regression is not significantly important compared to the classical one.
Now, to illustrate the asymptotic normality result we fix one curve $\mathcal{X}$ from the previous data, we draw m independent n-samples of the same data and we compute, for each sample, the quantity :

$$\left(\frac{n\phi_{\mathcal{X}}(h)}{\sigma^2(\mathcal{X})}\right)^{1/2}\left(\widetilde{r}(\mathcal{X}) - r(\mathcal{X}) - B_n(\mathcal{X}) - o(h)\right), \quad \text{For a fixed curve } \mathcal{X} \text{ from the previous data.}$$

Of course, the practical use of this asymptotic normality requires a conventional estimation for the normalized deviation $\left(\frac{n\phi_{\mathcal{X}}(h)}{\sigma^2(\mathcal{X})}\right)^{1/2}$ and the bias term $hB_n(\mathcal{X}) + o(h)$. However, for a sake of shortness, we neglect this bias term and we adapt, to the truncated case, the estimator proposed by Demongeot et al. (2016) for $\left(\frac{n\phi_{\mathcal{X}}(h)}{\sigma^2(\mathcal{X})}\right)^{1/2}$. Precisely, we estimate $g_3(\mathcal{X})$ and $g_4(\mathcal{X})$ in the same way as for $g_1(\mathcal{X})$ or $g_2(\mathcal{X})$ and the quantities β_1 and β_2 are



empirically estimated by :

$$\widehat{\beta_1} = \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} K_i G_n^{-1}(Y_i) \quad \text{and} \quad \widehat{\beta_2} = \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} K_i^2 G_n^{-1}(Y_i).$$

Finally, the practical estimator of the normalized deviation is :

$$\left(\frac{\left(\sum_{i=1}^{n} K_i G_n^{-1}(Y_i)\right)^2 \widetilde{g_2}^2(x)}{\left(\sum_{i=1}^{n} K_i^2 G_n^{-1}(Y_i)\right)\left(\widetilde{g_2}(x) - 2\widetilde{r}(x)\widetilde{g_3}(x) + \widetilde{r}^2(x)\widetilde{g_4}(x)\right)} \right)^{1/2}. \tag{4.1}$$

For the computational issue, we fix $m = 100$, we consider the homoscedastic case, we take $\mathcal{X} = X_0$ and we keep the same (h and d) of the first part. In the end, we compare the QQ-Plot against a standard normal distribution for the previous truncation levels and we use the Kolmogorov-Smirnov test to examine the asymptotic normality of the quantity of

$$\widehat{\beta_1} = \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} K_i G_n^{-1}(Y_i) \quad \text{and} \quad \widehat{\beta_2} = \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} K_i^2 G_n^{-1}(Y_i)\left(\widetilde{r}(\mathcal{X}) - r(\mathcal{X})\right)$$

The p.value of the ks.test are respectively 0.62, 0.27 and 0.11. Undoubtedly and similarly to the first illustration, the simulation results reveal that, the behaviour of the asymptotic normality is closely linked to the truncation rate. All in all, the present model has a good asymptotic behaviour and its implementation in practice is easy and fast.

Conclusion et Perspectives

1.Conclusion

On a étudié dans ce modèste travail les propriétés asymptotiques de l'estimateur de la fonction de régression relative où on a montré la convergence presque complète et la normalité asymptotique de notre estimateur dans les différents cas de cette étude.

Sur le modèle non paramétrique : son estimateur possède de bonnes propriétés asymptotiques dans tous les cas de notre étude. Ainsi, l'aspect non paramétrique est bien exploité dans ce travail par les hypothèses précédentes.
Sur les hypothèses sur le modèle, des hypoyhèses sur la variable explicative et des hypothèses techniques.
Sur la corrélation des observations : dans ce travail on a opté pour une vitesse de convergence insensible à la corrélation des observations en ajoutant des hypothèses supplémentaires dans le cas dépendant.
Sur la vitesse de convergence : ses vitesses de convergence sont en deux parties ; partie biais et partie dispersion. On trouvera les conditions sur la dimensionnalité du modèle dans la partie biais. Tandis que la dimensionnalité de la variable explicative est juste dans la partie dispersion.

2.Perspectives

Dans la continuitée de ce travail, plusieurs perspectives sont envisageables. Elles concernent essentiellement les méthodes récursives. Les principales perspectives de cette thèse sont les suivantes :

- Estimation de l'erreur relative d'un modèle aléatoire de données doublement tronquées à gauche.
- Estimation de l'erreur relative d'un modèle aléatoire de données spatio-tronquées.
- Estimation de l'erreur relative d'un modèle aléatoire de données tronquées-censurées.
- Estimation de l'erreur relative d'un modèle aléatoire de données tronqués aléatoirement.
- Estimation de la fonction de regression spatio-fonctionnelle dans le cas des données censurées.

Bibliographie

[1] Abraham, C., Biau, G., Cadre, B. (2004) On the asymptotic properties of a simple estimate of the mode. *ESAIM, Probab. Statist.* **8**, 1-11.

[2] Ahmad, I., (1976) Uniform strong convergence of a generalized failure rate estimate. *Bull. Math. Statist.* **17**, 77-84.

[3] ALIOUM. A., et COMMENGES., D. (1994) A Proportional Hazards Model for Arbitrarily Censored and Truncated Data, *manuscrit Université de Bordeau XII.*

[4] Aurzada, Frank ; Simon, Thomas. (2007) Small ball probabilities for stable convolutions. *ESAIM Probab. Stat.* **11**, 327-343.

[5] Benhenni, K., Hedli-Griche, S., Rachdi, M. (2010) Estimation of the regression operator from functional fixed-design with correlated errors. *J. Multivariate Anal.*, **101**(2), 476-490.

[6] Benhenni, K. ; Hedli-Griche, S. ; Rachdi, M. ; Vieu, P. (2008) Consistency of the regression estimator with functional data under long memory conditions. *Statist. Probab. Lett.* **78,**(8), 1043-1049.

[7] Basu, A.P. (1971) Bivariate failure rate. *J. Amer. Statist. Assoc.* **66**, 103-104.

[8] Bogachev, V.I. (1998) Gaussian measures.*Mathematical Surveys and Monographs 62 American Mathematical Society, Providence, RI.* xii+433 .

[9] Bosq, Denis. (1990) Modelization, nonparametric estimation and prediction for continuous time processes., *Kluwer Acad. Publ., Dordrecht,* **335,** 509-529.

[10] Bosq, Denis. (1991) Prévision d'une fonction aléatoire. Prediction of a random function. *Statistica (Bologna)* **51**(2), 153-164.

[11] Bosq, Denis. (1991) Statistique non paramétrique des processus à longue mémoire : le cas du modèle de régression. Nonparametric statistics of long-memory processes : the case of a regression model Processus longue mémoire. *Publ. Inst. Statist. Univ. Paris,* **36**(1-2), 119-123.

[12] Bosq, D. (1996) Nonparametric statistics for stochastic processes. Estimation and prediction. *Lecture Notes in Statistics Springer-Verlag, New York,* **110,** xii+169.

[13] Bosq, D. (1998) Nonparametric statistics for stochastic processes. Estimaion and prediction. *Second edition. Lecture Notes in Statistics Springer-Verlag, New York.* **110,** xvi+210.

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

[15] Bogachev, V.I. (1999) Gaussian measures. *Math surveys and monographs, Amer. Math.Soc.,* **62**.

[16] Bradley, Richard C. (2007) Introduction to Strong Mixing Conditions, *Kendrick Press, Heber City, UT* **1-2-3,**, 1-8.

[17] Burba, Florent and Ferraty, Frédéric and Vieu, Philippe. (2009) k-nearest neighbour method in functional nonparametric regression, *J. Nonparametr. Stat.*, **21** (4), 453-469.

[18] P. K. Bhattacharya, Herman Chernoff and S. S. Yang. (1983) Nonparametric Estimation of the Slope of a Truncated Regression, *The Annals of Statistics*, **10**.

[19] Billingsley, Patrick. (1995) Convergence of probability Measures,*John Wiley & Sons, Inc., New York.*

[20] Besse, Philippe and Ramsay, J. O. (1986) Principal components analysis of sampled functions, *Psychometrika*, **51** (2), 285-311.

[21] Cardot, Hervé., Crambes, Christophe., Sarda, Pascal. (2005) Quantile regression when the covariates are functions. *J. Nonparametr. Stat.*, **17** (7), 841-856.

[22] Besse, Philippe C., Cardot, Hervé., Faivre, Robert., Goulard, Michel. (2005) Statistical modelling of functional data. *Appl. Stoch. Models Bus. Ind.*, **21,**(2), 165-173.

[23] Cardot, Hervé., Sarda, Pacal. (2005) Estimation in generalized linear models for functional data via penalized likelihood. *J. Multivariate Anal.*,**92,**(1), 24-41.

[24] A. Carbonez and L. Györfi and van der Meulen, Edward C. (1995) Partitioning-estimates of a regression function under random censoring, **Statistics and Risk Modeling**, **18**.

[25] Carbon, M., Francq, C. et Tran, L. T. (2007) Kernel regression estimation for random fields. *J. Statist. Plann. Inference*, **137**, 778-798.

[26] Carbon, M., Francq, C. et Tran, L. T. (2010) Asymptotic normality of frequency polygons for random fields. *J. Statist. Plann. Inference*, **140**, 502-514.

[27] Carbon, M., Hallin, M. et Tran, L. T. (1996) Kernel density estimation for random fields : the L_1 theory *J. Nonparam. Stat.* **6**, 157-170.

[28] Carbon, M., Tran, L. T. et Wu, B. (1997) Kernel density estimation for random fields. *Statist. Probab. Lett.* **36**, 115-125.

[29] Campbell, M. Karen, Donner, Allan. (1989) Classification efficiency of multinomial logistic regression relative to ordinal logistic regression. *J. Amer. Statist. Assoc.* **84**(406), 587-591.

[30] Cohen (A.C.), (1959) Simplified estimators for the normal distribution when samples are singly censored or truncated. Technometrics, **1** (3), 217-237.

[31] Cuevas, Antonio ; Febrero, Manuel ; Fraiman, Ricardo. (2002) Linear functional regression : the case of fixed design and functional response. *Canad. J. Statist.*, **30**(2), 285-300.

[32] Cressie, N. A., Statistics for spatial data. *Wiley, New York*, 1991.

[33] Chen, J. and Zhang, L. (2009) Asymptotic properties of nonparametric M-estimation for mixing functional data, *J. Statist. Plann. and Inference,* **139**, 533-546.

[34] Chernoff, H. (1964) Estimation of the mode. *Ann. Instit. Statist. Math.* **16,** 31-41.

[35] Collomb ,G., Härdle, W. and Hassani, S. (1987) A note on prediction via estimation of the conditional mode function. *J. Statist. Plann. Inference.* **15**, 227-236.

[36] Dabo-Niang, Sophie ; Rhomari, Noureddine. (2003) Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R. Math. Acad. Sci. Paris,***336**(1), 75-80.

[37] Dabo-Niang, Sophie., F. Ferraty and P. Vieu. (2004) Estimation du mode dans un espace vectoriel semi-normé. *C. R. Acad. Sci. Paris, Ser.* **339,** 659-662.

[38] Dabo-Niang, S. et Laksaci, A. (2007) Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *C. R. Math. Acad. Sci. Paris.*, **344**, 49-52.

[39] Dauxois, J. and Pousse, A. and Romain, Y. (1982) Asymptotic theory for the principal component analysis of a vector random function : some applications to statistical inference, *J. Multivariate Anal.*, **12**, 136-154

[40] Dereich, S. (2003) High resolution coding of stochastic processes and small ball probabilities. PhD Thesis.

[41] Deville, Jean-Claude. (1974) Méthodes statistiques et numériques de l'analyse harmonique *Ministère de l'Économie et des Finances. Institut National de la Statistique et des Études Économiques. Annales*, **15**, 3-101.

[42] Deheuvels, Paul. (1980) Nonparametric test of independence, *Springer, Berlin*, **821,** 95-107.

[43] Donsker, M. (1951) An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.***6**.

[44] Doukhan, P. (1994) Mixing : Properties and Examples. *Lecture Notes in Statistics, Springer-Verlag.*

[45] Doukhan, P. et Louhichi, S. (1999) A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84**(2), 313-342.

[46] Eddy, W.F. (1980) Optimum kernel estimators of the mode. *Ann. Statist.* **8**, 870-882.

[47] Eddy, W.F. (1982) The asymptotic distributions of kernel estimators of the mode. *Z. Wahrsh. Verw. Gebiete.* **59**, 279-290.

[48] Estévez-Pérez, G. (2002) On convergence rates for quadratic errors in kernel hazard estimation. *Statist. Probab. Lett.* **57**, 231-241.

[49] Estévez-Pérez, G., Quintela-del-Río, A. et Vieu, P. (2002) Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples. *J. Statist. Plann. Inference.* **104**, 1-30.

[50] Ezzahrioui, M. et Ould-Saïd, E. (2008) Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *Journal of Nonparametric Statistics.* **20** (1), 3-18.

[51] Ferraty, Frédéric., Laksaci, Ali., Vieu, Philippe. (2006) Estimating some characteristics of the conditional distribution in nonparametric functional models. *Stat. Inference Stoch. Process.* **9** (1), 47-76.

[52] Ferraty, F., Vieu, P. (2006) Nonparametric Functional Data Analysis. *Theory and Practice. Springer-Verlag.*

[53] Ferraty, F. et Vieu, P. (2004) Modèle de régression pour variables aléatoires uni, multi et $\infty-$dimensionnées. Cours de DEA.

[54] Ferraty, F. et Vieu, P. (2006) Nonparametric functional data analysis, Theory and Practice. *Springer-Verlag.*

[55] Ferraty, F., Laksaci, A. et Vieu, P. (2005) Functional times seies prediction via conditional mode estimation. *C. R. Math. Acad. Sci. Paris.* **340**, 389-392.

[56] Ferraty, Frédéric and Laksaci, Ali and Vieu, Philippe (2006) Estimating some characteristics of the conditional distribution in nonparametric functional models, *Stat. Inference Stoch. Process.*, **9**(1), 47-76.

[57] Ferraty, F., Rabhi, A. et Vieu, P. (2008) Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle.*Rev. Roumaine Math. Pures Appl,* **53**, 1-18.

[58] Ferraty, F., Tadj, A., Laksaci, A. et Vieu, P. (2010) Rate of uniform consistency for nonparametric estimates with functional variables.*Journal of Statistical Planning and Inference.* **140**(2), 335-352.

[59] Ferraty, Frédéric ; Vieu, Philippe. (2006) Nonparametric functional data analysis. Theory and practice. *Springer Series in Statistics. Springer, New York,* xx+258 pp. ISBN : 0-387-30369-3 ; 978-0387-30369-7.

Ferraty, Frédéric and Vieu, Philippe. (2006) *Physica-Verlag/Springer, Heidelberg*, 112-129.

[60] Frydman, Halina. (1994) Asymptotic inference for the parameters of a discrete-time square-root process. *Math. Finance*, **4,**(2), 169-181.

[61] Frydman, Halina. (1994) A note on nonparametric estimation of the distribution function from interval-censored and truncated observations. *J. Roy. Statist. Soc. Ser.*, **56**(1), 71-74.

[62] Gannoun, A. (2002) Sur quelques problèmes d'estimation fonctionnelle : Théorie, Méthode et applications. H.D.R. Université Montpellier II.

[63] Gao, F. et Li, W.V. (2007) Small ball probabilities for the Slepian Gaussian fields. *Trans. Amer. Math. Soc.,* **359**(3), 1339-1350 (electronic).

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

[65] Giorgi, Roch., Belot, Aurélien., Gaudart, Jean., Launoy, Guy. (2008) The performance of multiple imputation for missing covariate data within the context of regression relative survival analysis. *Stat. Med,***27**(30), 6310-6331.

[66] Gilliom (R.J.), Helsel (D.R.), (1986) Estimation of distributional parameters for censored trace level water quality data. *Estimation techniques. Water Resources Research*, **22** (2), 135-146.

[67] Geffroy, J., (1974) Sur l'estimation d'une densité dans un espace métrique. *C. R. Acad. Sci., Paris, Sér.* A **278**, 1449-1452.

[68] Grenander, U. (1965) Some direct estimates of the mode. *Ann. Math. Stat.* 131-138.

[69] Grund, Birgit., Hall, Peter. (1995) On the minimisation of L^p error in mode estimation. *Ann. Statist.* **23,** (6), 2264-2284.

[70] 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.

[71] Györfi, L., Härdle, W., Sarda, P. et Vieu, P., (1989) Nonparametric curve estimation from time series. Lecture Notes in Statistics.*Springer-Verlag* , **6**.

[72] Konakov, V. D. (1973) The asymptotic normality of the mode of multivariate distributions. *(Russian) Teor. Verojatnost. i Primenen.* **18,** 836-842.

[73] Harter (H.L.), (1967) Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma populationfrom complete and censored samples. *Technometrics,* **9** (1), 159-165

[74] He, Shuyuan and Yang, Grace L. (1987) Estimation of regression parameters with left truncated data, *J. Statist. Plann. Inference*, **117**(1)x.

[75] Horrigue, W. and Ould Sa'id, E. (2015) Nonparametric regression quantile estimation for dependent functional data under random censorship : Asymptotic normality.*Comm. in Statistics. Theory and Methods,* **44,** 4307-4332.

[76] Ioannides, D. A. et Roussas, G. G. (1999) Exponential inequality for associated random variables. *Statist. Probab. Lett.,* **42**(4), 423-431.

[77] Jones, M. C., Park, Heungsun., Shin, key-Il., Vines, S. K., Jeong, Seokoh. (2008) Relative error prediction via kernel regression smoothers. *J. Statist. plann. Inference* **138** (10), 2887-2898.

[78] Laksaci, A. (2007) Erreur quadratique de l'estimateur à noyau de la densité conditionnelle à variable explicative fonctionnelle. *C.R. Math. Acad. Sci. Paris.* **345**, 149-154.

[79] Laksaci, A. et Mechab, B. (2010) Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales. *Rev. Roumaine Math. Pures Appl.* **55** (1), 35-51.

[80] Leclerc, J. et Pierre-Loti-Viau, D. (2000). Vitesse de convergence presque sûre de l'estimateur à noyau du mode. *C.R. Acad. Sci. Paris, Sér. I*, **331**, 637-640.

[81] Lecoutre, J. P. et 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.* **314** (4), 295-300.

[82] Lecoutre, Jean-Pierre. (1987) The histogram with random partition. *New perspectives in theoretical and applied . (Bilbao, 1986) statistics , Wiley Ser. Probab. Math. Statist. Probab. Math. Statist., Wiley, New York,*, 265-276.

[83] Leese (M.N.), (1973) Use of censored data in the estimation of Gumbel distribution parameters for annual maximum flood series. *Water Resources Research,* **9** (6), 15341542.

[84] Li, W. V. and 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.*

[85] Li, J. et Tran, L. T. (2007) Hazard rate estimation on random fields. *J. Multivariate Anal.* **98**, 1337-1355.

[86] Li, J. et Tran, L. T. (2009) Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference*, **139**, 164-175.

[87] Liebscher, E. (1996) Central limit theorems for sums of α-mixing random variables. *Stochastics Stochastics Rep.* **59**(3-4), 241-258.

[88] Liebscher, E. (2001) Central limit theorems for α-mixing triangular arrays with applications to nonparametric statistics. *Math. Methods Statist.* **10** (2), 194-214.

[89] Lifshits, M.A. et Simon, T. (2005) Small deviations for fractional stable processes. *Ann. Inst. H. Poincaré Probab. Statist.*, **41**, 725-752.

[90] Lifshits, Mikhail ; Linde, Werner ; Shi, Zhan. (2006) Small deviations of Gaussian random fields in L_q-spaces. *Electron. J. Probab.* **11**(46), 1204-1233.

[91] Loève, M. (1977) Probability theory I. *Springer-Verlag. New-York.*

[92] Louani, Djamal ; Ould-Sa'́d, Elias. (1999) Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametr. Statist* **11**(4), 413-442.

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

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

[95] Mokkadde m.A and M. Pelletier. (2005) Moderate deviations for the kernel mode estimator and some applications. *J. Statist. Plann. Inference.* **135,** 276-299.

[96] Nadaraya, E. (1964) On estimating regression. *Theor. Prob. Appl.,* **9,** 141-142.

[97] Nadaraya. E.N. (1965) On nonparametric estimates of density functions and regression curves. *Theory Probab. Appl.* **10,** 186-190.

[98] Nadaraya, E. (1989) Nonparametric estimation of probability densities and regression curves. *Transl, from Russian by Samnuel Kotz.*, Kluwer Academic Publisher, Dordrecht.

[99] Ouassou, Idir, Rachdi, Mustapha. (2010) Stein type estimation of the regression operator for functional data. *Adv. Appl. Stat. Sci.*, **1**(2), 233-250.

[100] Ould-Saíid, E. (1993). Estimation non paramétrique du mode conditionnel. *Application à la prévision. C. R. Acad. Sci. Paris. Ser.* **316,** 943-947.

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

[102] Parzen, Emanuel. (1962) Spectral analysis of asymptotically stationary time series. *Bull. Inst. Internat. Statist.* **39**(2), 87-103.

[103] Parzen, Emanuel. (1962) Extraction and detection problems and reproducing kernel Hilbert spaces.*J. SIAM control Ser.***1,** 35-62.

[104] Parzen, (1662) Emanuel. On estimation of a probability density function and mode.*Ann. Math. Statist.***33,** 1065-1076

[105] Phien (H.N.), Fang (T.-S.E.) (1989) Maximum-likelihood estimation of the parameters and quantiles of the general extreme-value distribution from censored samples. *Journal of Eydrology,* **105,** 139-155.

[106] Prakasa Rao, B. L. S. (1983) Nonparametric functional estimation. Probability and Mathematical Statistics. *Academic Press, Inc., New York.*

[107] Prakasa Rao, B. L. S. (1983) Estimation of the drift for diffusion process. *6th international summer school on problems of model choice and parameter estimation in regression analysis (Sellin)*, Seminarberichte, Humboldt Univ., Berlin., **51**, 1-16,

[108] Park, Heungsun. Stefanski , L, A. (1998) Relative error prediction. *Stattist. probab. Lett***40**(3), 227-236.

[109] Quintela-Del-Río, A, Vieu, P. (1997). A nonparametric conditional mode estimate. *J. Nonparametric Statist.* **8,** 253-266.

[110] Rachdi, Mustapha., Vieu, Philippe. (2005) Sélection automatique du paramètre de lissage pour l'estimation non paramétrique de la régression pour des données fonctionnelles. *C. R. Math. Acad. Sci. Paris*, **341**(6), 365-368.

[111] Rachdi, Mustapha., Philippe Vieu. (2007) Nonparametric regression for functional data : Automatic smoothing parameter selection. *Journal of Statistical Planning and Inference.,* **137,**(9), 2784-2801.

[112] Rachdi, M., Vieu, P.(2008). Nonparametric regression for functional data : automatic smoothing parameter selection. *J. Statist. Plann. and Inference,* **1370**, 2784-2801.

[113] Rachdi, M., El Methni, M. (2011) Local weighted average estimation of the regression operator for functional data. *Comm. Stat., Theory and Methods,* (Sous presse).

[114] Rao, J. V. Ranga. (1958) Solution of simultaneous differential equations using analog computers. *J. Sci. Engrg. Res.*, **2,** 215-220.

[115] Ramsay, J. O. (1982) When the data are functions. *Psychometrika,* **47,**4, 379-396.

[116] Ramsay, J. O. (1982) Some statistical approaches to multidimensional scaling data. With discussion and a reply by the author. *J. Roy. Statist. Soc. Ser.*, **145**(3), 285-312.

[117] Ramsay, John Robert. (1987) EXTENSIONS OF LJUSTERNIK-SCHNIRELMANN CATEGORY THEORY TO RELATIVE, EQUIVARIANT AND ISOVARIANT THEORIES. Thesis (Ph.D.)-The University of Wisconsin - Madison. *ProQuest LLC, Ann Arbor, MI.*

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

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

[120] Ramsay, J. O. and Silverman, B. W. (2002) Applied functional data analysis, *Springer-Verlag, New York*, x+190.

[121] Ramsay, J. et Silverman, B. W. (2005) Functional Data Analysis. *2nd Ed. Springer Series in Statistics. Springer, New York* .

[122] Rio, E. (2000) Théorie asymptotique des processus aléatoires faiblements dépendant (in french). Mathématiques & Applications. *Springer Verlag,* **31**.

[123] Romano J. P. (1988) On weak convergence and optimality of kernel density estimates of the mode. *Ann. Statist.* **16,** 629-647.

[124] Roussas, G., (1989) Hazard rate estimation under dependence conditions. *J. Statist. Plann. Inference.* **22**, 81-93.

[125] Roussas, G., (1989) Asymptotic normality of the kernel estimate under dependence conditions : application to hazard rate. *J. Statist. Plann. Inference.* **25**, 81-104.

[126] Rosa, A. C. (1993) Prévision robuste sous une hypothèse ergodique. *Thèse de Doctorat de l'Université de Toulouse I.*

[127] Rosenblatt, Joan Raup. (1956) ON A CLASS OF NON-PARAMETRIC TESTS, Thesis (Ph.D.)-The University of North Carolina at Chapel Hill, *ProQuest LLC, Ann Arbor, MI.*

[128] Rosenblatt, M. (1956) A central limit theorem and a strong mixing condition. *Proc.Nat. Acad. Sci. U.S.A.* **42,** 43-47.

[129] Rio, E. (2000) Théeorie asymptotique des processus aléatoires faiblement dépendants, of Mathématiques et Applications (Berlin) [Mathematics and Applications]. *Springer-Verlag, Berlin*, **31** .

[130] Groeneboom, Piet ; Wellner, Jon A. (1992) Information bounds and nonparametric maximum likelihood estimation. *Birkhäuser Verlag, Basel,* **19** (4), 7643-2794.

[131] Ruiz-Velasco, S. (1991) Asymptotic efficiency of logistic regression relative to linear discriminant analysis. *Biometrica*, **78**(2), 235-243.

[132] Samanta. M. (1973) Nonparametric estimation of the mode of a multivariate density. *South African Statist. J.* **7,** 109-117.

[133] Samanta, M., and Thavaneswaran, A. (1990) Nonparametric estimation of the conditional mode. Comm. *Statist. Theory and Methods.* **16,** 4515-4524.

[134] Sager, T. (1975) Consistency in nonparametric estimation of the mode. *Ann. Stat.* **3,** 698-706.

[135] Shmileva, Elena. (2006) Small ball probabilities for jump Lévy processes from the Wiener domain of attraction. *Statist. Probab. Lett.* **76**(17), 1873-1881.

[136] Stute, W. (1993) Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.*, **21,** 146-156.

[137] Stone, C. (1981) Optimal rates of convergence for nonparametric regression. *Ann. Statist.,* **9,** 1348-1360.

[138] Stone, Charles J. (1982) Optimal global rates of convergence for nonparametric regression.*Ann. Statist.* **10**(4), 1040-1053.

[139] Stone, A. P., (1982) Existence of conservation laws for several vector 1-forms, *The Tensor Society. Tensor. New Series*, **39**, 10-14.

[140] Silverman, B. W., (1986) Density estimation for statistics and data analysis, *Chapman & Hall, London*, x+175.

[141] Singpurwalla, N. et Wong, M.Y. (1983) Estimation of the failure rate. A survey of nonparametric models. Part I : Non-Bayesian methods. *Comm. Statist.* **12**, 559-588.

[142] Tore Dalenius. (1965) The Mode-A Neglected Statistical Parameter *Journal of the Royal Statistical Society.* **1** (128), 110-117.

[143] Tucker, Don Harrell. (1958) ABSOLUTELY CONTINUOUS SOLUTIONS OF $U(,XY) = F(X,Y,U(,X),U(,Y))$. Thesis (Ph.D.)-The University of Texas at Austin. *ProQuest LLC, Ann Arbor,* **45**.

[144] Tran, L. T., (1990) Kernel density estimation on random fields. *J. Multivariate Anal.* **34,** 37-53.

[145] Trunbull, B. W. (1974) Nonparametric estimation of a survivorship function with doubly censored data. *J. Am. Stat. Assoc.,* **69,** 169-173.

[146] Turnbull B.W. (1976) The Empirical Distribution Function with Arbitrary Grouped Censored and Truncated Data. *J.Roy.Statist.SocSer. B.*,290-295.

[147] Van Ryzin, J. (1969) On strong consistency of density estimates. *Ann. Math. Statist.* **40,** 1765-1772.

[148] Venter, J. (1967) On estimation of the mode. *Ann. Math. Stat.* **38,** 1446-1455.

[149] Vieu, P.(1991) Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.*, **39,** 324-347.

[150] Vieu, Philippe.(1996) A note on density mode estimation. *Statist. Probab. Lett.* **26** (4), 297-307.

[151] Watson, Geoffrey S. (1964) Smooth regression analysis. *Sankhya Ser.* **26**, 359-372.

[152] Woodroofe, M. (1985) Estimating a distribution function with truncated data. *The Annals of Statistics,* **13,** 163-177.

[153] Wang, J., Liang, H.Y. and Fan, G. (2002b) Local M-estimation of non-parametric regression with left-truncated and dependent data. *Scientia Sinica Mathematica,* **42,** 995-1015.

[154] Wang, Zhanfeng ; Chen, Zhuojian ; Chen, Zimu. (2018) H-relative error estimation for multiplicative regression model with random effect. *Comput. Statist.* **33,**(2), 623-638.

[155] Wegman, E. (1971). A note on the estimation of the mode.*Ann. Math. Stat.* **42,** 1909-1915.

[156] Yamato, H. (1971). Sequential estimation of a continuous propability density function and mode. *Bull. Math. Statist. Jap.*, **14,** 1-12.

[157] Youndjé, É. (1993) Estimation non paramétrique de la densité conditionnelle par la méthode du noyau. Thèse de Doctorat. Université de Rouen.

[158] Yoshihara, Ken-ichi. (1975) Billingsley's theorems on empirical processes of strong mixing sequences, *Yokohama Math. J.*, **23** (1-2), 77-83.

[159] Yoshihara, K. (2004) Weakly dependent stochastic sequences and their applications. Recent topics on weak and strong limit theorems. *Sanseido Co., Ltd., Chiyodaku.* **XIV.** vi+410.

[160] Younes, Hassan. (2005) Estimation du taux de mortalite sous contraintes d'ordre pour des donnees censurees ou tronquees, Thesis (Ph.D.),Universite du Quebec a Montreal (Canada), *ProQuest LLC, Ann Arbor, MI*, 107.