

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Abderrahmane Mira de Béjaïa
Faculté des Sciences Exactes
Département Mathématiques



Mémoire de fin de cycle

En vue de l'obtention du diplôme Master en Mathématiques

Option : Probabilités Statistiques et Applications

Thème

Estimation non paramétrique fonctionnelle du mode conditionnel

Réalisé par :

M^{lle} KHABIL Nawal

Devant le Jury composé de

<u>Président :</u>	M. Samir. RAHMANI	M.C.A	Université de Béjaïa.
<u>Promotrice :</u>	M ^{me} Hadjila. TABTI	M.C.B	Université de Béjaïa.
<u>Examinatrice :</u>	M ^{me} Kheira. MECHERI	M.C.B	Université de Béjaïa.
<u>Examinatrice :</u>	M ^{me} Kenza. TAMITI	DOCTEUR	Université de Béjaïa.

Juillet 2019

Dédicaces

Je dédie ce modeste travail :

A

Ma mère adorée aucune dédicace ne saurait être assez élégante pour exprimer
ce qu'elle mérites.

A

Mon père qui m'a indiqué la bonne voie, qui m'a encouragé durant toutes mes
années d'études.

Que Dieu les protège et leurs donne tout le bonheur du monde.

A

Mon frère Mustapha. À qui je souhaite un avenir radieux et plein de réussite.

Ma chère Hamida et son mari.

A

Mes chers cousins et cousines : Fafa, Nada, Dina, Zaki.

A

Mes chères : Riham, Rihab, Amir, Ines.

A

Ceux qui m'ont créé un milieu d'ambiance et de travail, mes amis (es) :

Djamila, Katia, Nouara, Saida, Sylia, Nawal, Chanez, Yasmine.

A

Tous mes enseignants.

A

Tous ceux qui me sont chers

Nawal

Remerciements

Tout d'abord, je remercie Dieu le tout-puissant qui m'a donné le courage, la volonté pour mener à bien ce travail.

J'ai l'honneur et le plaisir d'exprimer mes profondes gratitudee à ma promotrice **M^{me}. TABTI Hadjila** pour avoir accepté de m'encadrer, pour ses remarques, ses orientations, et ses conseils qu'elle m'a prodigué tout au long de mon travail.

Les membres qui ont accepté de siéger sur le jury de ce mémoire doivent aussi trouver ici l'expression de ma reconnaissance, soient le président **M^r. RAHMANI Samir** les examinateurs, **M^{me}. MECHERI Kheira**, et le docteur **M^{me}. TAMITI Kenza**.

Tout particulièrement je remercie **M^{me}. SAADI Noura**, pour son aide, et ses orientations.

Également je tiens à remercier toute ma famille, surtout ma mère, mon père, mon frère qui m'ont épaulés, soutenus et suivis tout au long de ce projet.

Je remerciée mes amies pour leurs aide et conseils qu'elles m'ont donné tout au long de mon travail.

Enfin, je remerciée toutes les personnes ayant contribué de proche ou de loin à la réalisation de ce travail.

TABLE DES MATIÈRES

Table des matières	2
listes des abréviations	3
Table des figures	4
Introduction générale	6
1 Préliminaires	9
1.1 Variables aléatoires fonctionnelles	9
1.1.1 Les données fonctionnelles	10
1.2 Méthode de noyau	10
1.3 Noyaux	11
1.3.1 Noyaux usuels	11
1.3.2 type de noyau	13
1.4 Choix de la semi-métrique	13
1.5 Convergence presque complète	14
1.6 Inégalités exponentielles	16
1.7 Probabilité des petites boules	19
2 Estimation non paramétrique du mode conditionnelle	20
2.1 Estimation de la densité conditionnelle	21
2.1.1 Historique	21
2.2 Estimation non paramétrique de la densité conditionnelle	21

2.2.1	Présentation du modèle	21
2.3	La convergence presque complète de la densité conditionnelle	23
2.3.1	Hypothèses	23
2.4	Estimation non paramétrique du mode conditionnel	36
2.4.1	Le mode conditionnel	36
2.4.2	Présentation du modèle	36
2.5	La convergence presque complète du mode conditionnel	37
3	Simulation	40
3.1	Choix du langage de programmation	40
3.2	Simulation des données fonctionnelles	41
3.2.1	Résultats de simulation	41
3.2.1.1	Résultats de plusieurs courbes simulées	41
3.2.1.2	Résultat d'une variable aléatoire fonctionnelle	42
3.3	plan d'estimation du mode conditionnel	42
3.3.1	Algorithme	43
	Conclusion générale	47
	Bibliographie	48

Notations et abréviations

$B(x, h)$: Boule ouverte du centre x et de rayon h .

(E, d) : Espace semi-métrique de dimension éventuellement finie.

\mathbb{R} : Ensemble des réels.

H : Fonction de noyau intégrée.

K_0 : Fonction de noyau symétrique standard.

i.i.d : Indépendante et identiquement distribué.

$f^x(y)$: La densité conditionnelle de Y sachant X .

$f(x)$: La fonction de densité.

$F^x(y)$: La fonction de répartition conditionnelle de Y sachant X .

$\hat{f}^x(y)$: L'estimateur à noyau de la densité conditionnelle de Y sachant X .

$\hat{F}^x(y)$: L'estimateur à noyau de la fonction de répartition conditionnelle de Y sachant X .

$\mathbf{1}_{[0,1]}$: La fonction indicatrice sur l'ensemble $[0,1]$.

K : Le noyau.

φ_x : Mesure de la boule $B(x, h)$ par rapport à la loi de probabilité de X .

x : Observation de la Variable aléatoire X .

y : Observation de la Variable aléatoire Y .

r : Opérateur de régression non linéaire.

h : Paramètre de lissage ou largeur de fenêtre h_n .

g : Paramètre de lissage ou largeur de fenêtre g_n .

$p.co$: Presque complète.

$p.s$: Presque sûre.

$d(\cdot)$: Semi-métrique sur un espace fonctionnel E .

S : Sous-ensemble compact générique de \mathbb{R} .

$v.a.f$: Variable aléatoire fonctionnelle.

X : Variable aléatoire fonctionnelle.

Y : Variable aléatoire réel.

$O_{p.co}$: Vitesse de convergence presque complète.

TABLE DES FIGURES

Figure 1.1 Quelques noyaux symétriques usuels.....	12
Figure 3.1 Courbes de plusieurs v.a.f.....	41
Figure 3.2 Une courbe de v.a.f.....	42
Figure 3.3 Résultats de l'estimation.....	45

INTRODUCTION GÉNÉRALE

La statistique fonctionnelle occupe désormais une place importante dans la recherche en statistique. Il s'agit de la modélisation statistique des données qui sont des courbes supposées observées sur toutes leurs trajectoires. Ceci est pratiquement possible en raison de la précision des appareils de mesures modernes et de l'importante capacité de stockage qu'offrent les systèmes informatiques actuels. Il est facile d'obtenir une discrétisation très fine d'objets mathématiques tels que des courbes, surfaces...ce type de variables se retrouve dans de nombreux domaines, comme la météorologie, la chimie quantitative, la biométrie, l'économétrie ou l'imagerie médicale. Parmi les ouvrages de référence en la matière, on peut citer les monographies de Ramsay et Silverman (1997) pour les aspects appliqués, Bosq (2000) pour les aspects théoriques, Ferraty et Vieu (2006)[8] pour une étude non paramétrique et Ferraty et Romain (2011) pour des développements récents. Dans le même contexte, nous renvoyons à Manteiga et Vieu (2007) ainsi que Ferraty (2010), les premiers résultats en statistique non paramétrique fonctionnelle ont été élaborés par Ferraty et Vieu (2000) et ils concernent l'estimation de la fonction de régression à variable explicative de dimension fractale. Ils ont établi la convergence presque complète d'un estimateur à noyau de ce modèle non paramétrique dans le cas i.i.d(Indépendante et identiquement distribué).

Les modèles conditionnels ont pris leurs places dans la statistique : la fonction de répartition conditionnelle, la densité conditionnelle, le quantile conditionnel, la fonction du hasard conditionnelle et le mode conditionnel. A ces sujets, les premiers résultats ont été réalisés au début des années soixante par Roussas (1968) qui a étudié les propriétés asymptotiques de l'estimateur à noyau de la fonction de répartition conditionnelle où il a montré la convergence en probabilité, un estimateur alternatif pour le même modèle a été élaboré par Stone (1977). Youndjé (1993)[20]

à donné l'estimateur de la fonction de densité conditionnelle avec des données indépendantes et dépendantes. Ainsi l'intérêt que revêt le mode conditionnel dans le domaine de la prévision, plusieurs auteurs s'en sont intéressés, nous pouvons citer , Collomb et al(1987)[6], et aussi Perzen (1962)[15] qui a été l'un des premiers à considérer le problème de l'estimation du mode d'une densité de probabilité univariée. Il a montré que, sous certaines conditions, l'estimateur du mode obtenu en maximisant un estimateur à noyau est convergent et est asymptotiquement normal quand les données sont indépendantes et identiquement distribuées (i.i.d). Samanta et Thavaneswaren (1990) qui ont étudié l'estimation non paramétrique du mode conditionnel et ils ont établi la normalité asymptotique de l'estimateur de ce modèle avec des données indépendantes et identiquement distribuées, le cas α – *mélange* a été étudié par Ould-Said (1993), puis il a étendu ce dernier travail en (1997) pour des données ergodiques. De leur côté, Quintela et Vieu (1997) ont estimé le mode conditionnel comme étant le point annulant la dérivée d'ordre un de l'estimateur de la densité conditionnelle et ils ont établi la convergence presque complète de cet estimateur sous la condition α – *mélangeance*. Berlinet et al. (1998), quant à eux, ils ont présenté des résultats sur la normalité asymptotique des estimateurs convergents du mode conditionnel, indépendamment de la structure de dépendance des données avec une application au cas d'un processus stationnaire α – *mélangeant*. Tandis que Louani et Ould-Said (1998-1999) ont montré la normalité asymptotique de l'estimateur du mode conditionnel avec des données censurées et des données fortement mélangeantes. Nous pouvons, aussi citer le travail mené par Laksaci et Yousfate (2002) et dans lequel ils ont établi , pour un processus markovien stationnaire, la convergence en norme L^p de l'estimateur à noyau de la densité conditionnelle, et aussi Ioannides et Matzner (2002) ont construit un estimateur pour le mode conditionnel, lorsque, les observations sont entachées d'erreurs. Ould-Said et Cai (2005), quant à eux, ils ont établi la convergence uniforme sur un compact. Par ailleurs, dans le cadre de données à valeurs dans un espace de dimension éventuellement finie, les travaux de Ramsay et Silverman (2002 et 2005) constituent un recueil important de méthodes statistiques, principalement du point de vue pratique, mais des développements théoriques peuvent être trouvés dans Bosq (2000) et Ferraty et Vieu (2006)[8]. Donc, L'estimation de la densité conditionnelle en dimension éventuellement finie a connu un grand intérêt en statistique.

Dans ce mémoire, nous nous intéressons à l'estimation non paramétrique de la densité conditionnelle, et les paramètres qui en découlent, comme le mode conditionnel, pour les variables aléatoires fonctionnelles.

Ce mémoire comporte trois chapitres, une conclusion, et une bibliographie.

Dans le premier chapitre, nous avons présenté les variables aléatoires fonctionnelle et les données fonctionnelles, les notions de noyaux et les types de quelques noyaux usuels, ainsi que inégalités exponentielles, et aussi nous avons défini la notion de la convergence presque complète

Dans le deuxième chapitre, nous avons traité des notions de l'estimateur de la densité conditionnelle et le mode conditionnel quand les variables explicative est fonctionnelle, et nous avons présenté ensuite certaines hypothèses de base de la statistique non paramétrique et puis nous avons établi la convergence presque complète pour les deux modèles(La densité conditionnelle, Le mode conditionnel).

Dans le troisième chapitre, nous avons illustré les résultats théoriques par une simulation, à l'aide du logiciel R.

Introduction

La théorie de l'estimation non paramétrique s'est développée considérablement ces deux dernières décennies, ce développement a exigé la modernisation des méthodes statistiques comme outils d'analyse et de contrôle. Ainsi, une nouvelle branche de la statistique, dénommée statistique fonctionnelle, s'est développée pour traiter des observations comme les variables aléatoires fonctionnelles.

L'objectif de ce chapitre, est de présenter, quelques définitions de base sur les variables fonctionnelles et sur le choix de la semi-métrique. Puis, nous donnerons quelques théorèmes et propriétés de convergence presque complète, et nous terminerons par les notions des probabilités des petites boules et par la présentation générale de la méthode de noyau et nous citons quelques exemples des noyaux usuels.

1.1 Variables aléatoires fonctionnelles

Il y'a en réalité un nombre croissant de situations venant des champs différents de sciences appliquées (environnements, chimétries, la biométrie, la médecine, l'économétrie, ...) dans lequel les données rassemblées sont des courbes. En effet, le progrès des outils de calcul, tant en termes de mémoire que des capacités informatiques, nous permet de traiter les grands ensembles de données. Particulièrement pour un seul phénomène, nous pouvons observer un très grand ensemble de variables. Par exemple, regardez la situation habituelle suivant où une certaine

variable aléatoire peut être observée à plusieurs moments différents dans la gamme $(t_{\min}, t_{\max}) \in T$. Une observation peut être exprimée par la famille aléatoire $X(t_j)_{j=1\dots J} J \in N$.

Dans les statistiques moderne, quand le grille devient plus en plus petite, les instants consécutifs sont de plus en plus proche. Une façon de prendre en compte est de considérer les données comme une observation de la famille continue $X = X(t), t \in (t_{\min}, t_{\max})$. [8]

Définition 1.1.1

Une variable aléatoire X est dite fonctionnelle si ses valeurs sont dans un espace fonctionnel. [15]

Définition 1.1.2

On appelle modèle fonctionnel, tout modèle prenant en compte au moins une variable aléatoire fonctionnelle (v.a.f). [9]

1.1.1 Les données fonctionnelles

Depuis le milieu des années quatre-vingt-dix, le nombre croissant de situations ,où des variables fonctionnelles peuvent être observées, a entraîné des développements statistiques différents que nous pourrions à la suite nommer Statistique pour des Variables Fonctionnelles (ou des Données fonctionnelles) [8].

Définition 1.1.1.1

Une donnée fonctionnelle x est une réalisation de la variable fonctionnelle X .

1.2 Méthode de noyau

Le concept de noyau à d'abord été décrit en (1951), dans un rapport non publié par Fix et Hodges. Un peu plus tard, la première forme de l'estimateur à noyau était introduite dans les travaux de Rosenblatt (1956) [16], et Parzen (1962) [14]. La méthode à noyau (ou méthode de Parzen-Rosenblatt) est une généralisation de la méthode d'estimation par histogramme étudiée par des statisticiens et des probabilistes, cette dernière méthode est une méthode non paramétrique d'estimation de la densité d'une variable aléatoire. L'estimateur de la densité par

la méthode du noyau est le plus répandu aujourd'hui, car il répond au problème du choix des différents paramètres et possède des propriétés.[20]

1.3 Noyaux

Nous définissons la notion de noyau et nous donnons ses différents types

Définition 1.3.1

Un noyau $K : \mathbb{R} \rightarrow \mathbb{R}^+$ est une fonction de pondération utilisée dans des techniques d'estimation non paramétrique, qui doit vérifier les deux conditions suivantes :

- $\int_{\mathbb{R}} K(u) du = 1$.
- $K(u) = K(-u)$ pour toutes les valeurs de u .

1.3.1 Noyaux usuels

On donne ici une brève présentation de quelques noyaux usuels[1]

- Noyau rectangulaire(Uniforme)

$$K(u) = \begin{cases} \frac{1}{2} & \text{si } |u| \leq 1 \\ 0 & \text{si } |u| > 1. \end{cases}$$

- Noyau triangulaire

$$K(u) = \begin{cases} 1 - |u| & \text{si } |u| \leq 1 \\ 0 & \text{si } |u| > 1. \end{cases}$$

- Noyau d'épanechnikov ou parabolique

$$K(u) = \begin{cases} \frac{3}{4}(1 - u^2) & \text{si } u \in [-1, 1] \\ 0 & \text{sinon.} \end{cases}$$

- Noyau quadratique

$$K(u) = \begin{cases} \frac{15}{16}(1 - u^2)^2 & \text{si } u \in [-1, 1] \\ 0 & \text{sinon.} \end{cases}$$

- Noyau cubique

$$K(u) = \begin{cases} \frac{35}{32}(1 - u^2)^3 & \text{si } u \in [-1, 1] \\ 0 & \text{sinon.} \end{cases}$$

- Noyau gaussien

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right), u \in \mathbb{R}.$$

- Noyau sinus

$$K(u) = \begin{cases} \frac{1}{2\pi} \left(\frac{\sin(\frac{\pi}{2}u)}{\frac{\pi}{2}}\right)^2 & \text{si } u \neq 0 \\ \frac{1}{2\pi} & \text{si } u = 0. \end{cases}$$

- Noyau cosinus

$$K(u) = \begin{cases} \frac{\pi}{4} \cos\left(\frac{\pi u}{2}\right) & \text{si } u \in [-1, 1] \\ 0 & \text{sinon.} \end{cases}$$

- Noyau de silverman

$$K(u) = \frac{1}{2} \exp\left(-\frac{|u|}{\sqrt{2}}\right) \sin\left(\frac{-|u|}{\sqrt{2} + \frac{\pi}{4}}\right), u \in \mathbb{R}.$$

La figure (1.1) représente quelques noyaux symétriques usuels

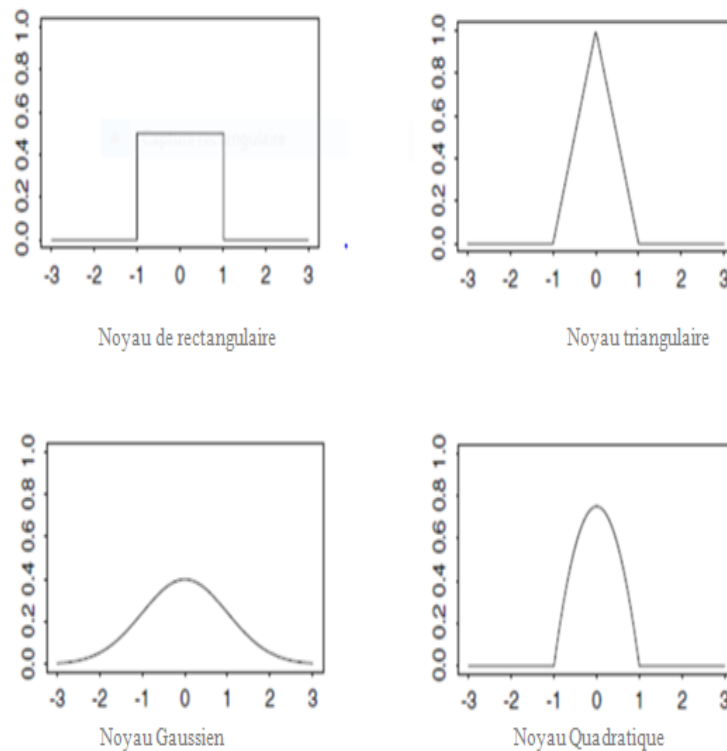


FIGURE 1.1 – Quelques noyaux symétriques usuels

1.3.2 type de noyau

Définition 1.3.2.1 [8]

Le noyau K est dit de type I si il existe deux constantes réelles $0 < C_1 < C_2 < \infty$ telle que

$$C_1 1_{[0,1]} \leq K \leq C_2 1_{[0,1]}.$$

Définition 1.3.2.2 [8]

Le noyau K est dit de type II si son support est $[0, 1]$ et si sa dérivé K' existe sur l'intervalle $[0, 1]$ et cette dérivée vérifie la condition suivante : s'il existe deux constantes réelles C_1 et C_2 telle que $-\infty < C_2 < C_1 < 0$ alors

$$C_2 \leq K' \leq C_1.$$

Définition 1.3.2.3 [8]

Une fonction $K : \mathbb{R} \rightarrow \mathbb{R}^+$ telle que $\int K(u)du = 1$ sur un support compact $[-1, 1]$ et $\forall u \in (0, 1), K(u) > 0$ est appelé noyau de type 0.

1.4 Choix de la semi-métrie

La question du choix optimal de la semi-métrie a une importance capitale dans la pratique. A notre connaissance il n'existe pas encore de méthode pratique pour le choix optimal de la semi-métrie. plus crucial pour l'étude des variables aléatoires fonctionnelles. Nous présentons ici quelques pistes permettant de construire une semi-métrie [8]. En fait, nous présentons, dans ce qui suit, seulement trois familles de semi-métries mais, naturellement, beaucoup d'autres peuvent être construites.

1. Semi-métrie PCA (méthode d'analyse en composantes principales fonctionnelles)

$$d_q^{PCA}(X_i, x) = \sqrt{\sum_{k=1}^q \left(\int [X_i(t) - x(t)] v_k(t) dt \right)^2}.$$

Et $t \in T$

Avec $v_k = v_1, v_2, \dots, v_q$ les fonctions propres orthonormales de l'opérateur de covariance.

Et q n'est pas vraiment un paramètre de lissage, mais plutôt un paramètre de réglage indiquant le niveau de résolution auquel le problème est considéré. On peut approximer l'intégrale de la manière suivante :

$$\int [X_i(t) - x(t)] v_k(t) dt \approx \sum_{j=1}^J w_j (X_i(t_j) - x_{i'}(t_j)) v_k(t_j)$$

Où les poids $w_j = t_j - t_{j-1}$

2. Semi-métrie PLS

$$d_q^{PLS}(x_i, x_{i'}) = \sqrt{\sum_{k=1}^p \left(\sum_{j=1}^J w_j (X_i(t_j) - x_{i'}(t_j)) [v_k^q]_j \right)^2}$$

3. Semi-métriques basées sur des dérivés

$$d_q^{deriv}(x_i, x_{i'}) = \sqrt{\int (x_i^{(q)}(t) - x_{i'}^{(q)}(t))^2 dt.}$$

Avec $x_i, x_{i'}$ deux courbes observées, ou $x^{(q)}$ désigne la dérivée d'ordre q de x .

1.5 Convergence presque complète

Dans la théorie des probabilités, il existe différentes notions de convergence, parmi elles (la convergence presque sûr, convergence presque complète, la convergence en probabilité, etc..).

Dans cette section nous présentons la convergence presque complète qui à été introduit par Hsu et Robbins (1947)[11].

Definition 1.5.1 [3]

Une suite de variables aléatoires réelles $(Z_n)_{n \in \mathbb{N}}$ converge presque complètement vers une variable aléatoire Z lorsque $n \rightarrow \infty$ (est on écrit $\lim_{n \rightarrow \infty} Z_n = Z$ p.co) si et seulement si :

$$\forall \epsilon > 0, \sum_{n \in \mathbb{N}} \mathbb{P}[|Z_n - Z| > \epsilon] < \infty.$$

Definition 1.5.2 [3]

On dit que la vitesse de convergence presque complète de la suite de variables aléatoires réelles $(Z_n)_{n \in \mathbb{N}}$ vers Z est d'ordre (U_n) (avec U_n est une suite numérique déterministe), et on note

$$Z_n = O(U_n)_{p.co}$$

si seulement si :

$$\exists \epsilon_0 > 0, \sum_{n \in \mathbb{N}} \mathbb{P}[|Z_n - Z| > \epsilon_0 U_n] < \infty.$$

Notons que la convergence presque complète entraîne à la fois la convergence presque sûre et la convergence en probabilité.

Proposition 1.5.3 [10]

Si $\lim_{n \rightarrow \infty} Z_n = Z$ p.co, alors Z_n converge en probabilité et presque sûrement vers Z .

Preuve de proposition

La convergence en probabilité se déduit facilement de la convergence de la série suivant

$$\sum_{n \in \mathbb{N}} \mathbb{P}[|Z_n - Z| > \epsilon] < \infty$$

($\mathbb{P}[|Z_n - Z| > \epsilon]$, est le terme général d'une série convergente). Le lemme de Borel contelli implique que :

$$\forall \epsilon > 0, \mathbb{P}\left[\lim_{n \rightarrow \infty} |Z_n - Z| > \epsilon\right] = 0 \quad p.s$$

De plus, $\lim_{n \rightarrow \infty} Z_n(\omega) \neq Z(\omega)$, implique l'existence de $\epsilon > 0$, telque

$$\lim_{n \rightarrow \infty} \sup |Z_n(\omega) - Z(\omega)| > \epsilon \quad p.s$$

On a alors $\mathbb{P}\left(\lim_{n \rightarrow \infty} \sup |Z_n(\omega) - Z(\omega)| > \epsilon\right) = 1$, c'est à dire $Z_n \rightarrow Z$, p.s.

1.6 Inégalités exponentielles

Dans la section 1.6 nous introduisons quelques inégalités exponentielles de type Bernstein (dans le cas des variables aléatoires indépendantes), qui nous seront utiles à la démonstration des propriétés de la convergence presque complète.

Nous supposons que Z_1, Z_2, \dots, Z_n est une suite de variables aléatoires réelles indépendantes et centrées. Pour démontrer la convergence presque complète, nous avons besoin de trouver des bornes supérieures pour certaines probabilités concernant des sommes de variables aléatoires réelles telles que :

$$\mathbb{P} \left(\left| \sum_{i=1}^n Z_i \right| > \epsilon \right).$$

Où le réel positif ϵ diminue avec n .

Dans ce contexte, il existe de puissants outils probabilistes, appelés génériquement inégalités exponentielles.

La littérature contient diverses versions d'inégalités exponentielles. Ces inégalités diffèrent selon les différentes hypothèses vérifiées par les variables Z_i .

Nous en présentons ici celles qu'on appelle inégalités de type Bernstein qui est la plus facile pour les développements théoriques sur les statistiques fonctionnelles[8].

Proposition 1.6.1[13]

on a l'inégalité de Cramer suivante : Si

$$\forall m \geq 2, \quad |\mathbb{E}(Z_i^m)| \leq \left(\frac{m!}{2}\right)(a_i)^2 b^{m-2}$$

Où

$(a_i)_{1 \leq i \leq n}$ sont des réels positifs, $b \in \mathbb{R}^+$ et $A_n^2 = a_1^2 + a_2^2 + \dots + a_n^2$ Alors

$$\forall \epsilon \geq 0, \quad \mathbb{P} \left[\sum_{i=1}^n |Z_i| > \epsilon A_n \right] \leq 2 \exp \frac{-\epsilon^2}{2(1 + \frac{b\epsilon}{A_n})}.$$

Preuve de proposition voir [8]**Corollaire 1.6.2 [8]**

a) $\forall m \geq 2, \exists C_m > 0$ et une constante a positive, tels que :

$$\mathbb{E} | Z_1^m | \leq C_m a^{2(m-1)}.$$

On a

$$\forall \epsilon > 0, \mathbb{P} \left[\left| \sum_{i=1}^n Z_i \right| > \epsilon n \right] \leq 2 \exp \frac{-\epsilon^2 n}{2a^2(1+\epsilon)}.$$

b) Supposons que les $(Z_i)_{1 \leq i \leq n}$ dépendent de n (c'est à dire que $Z_i = Z_{i,n}$). Si $\forall m \geq 2$, il existe un réel C_m strictement positif et une suite (a_n) de réels positifs, tels que :

$$\mathbb{E} | Z_1^m | \leq C_m a_n^{2(m-1)}.$$

Et si $U_n = n^{-1} a_n^2 \log n$, vérifie $\lim_{n \rightarrow \infty} U_n = 0$, alors nous avons

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{(U_n)}) p.co.$$

Démonstration du Corollaire 1.6.2

1. En remplace $b = a^2$ et $A_n = a\sqrt{n}$ dans la proposition précédente, on aboutit à **a**).
2. En posant $\epsilon = \epsilon_0 \sqrt{U_n}$ dans a) et comme U_n tend vers zéro, pour une certaine constante C' nous avons :

$$\mathbb{P} \left[\frac{1}{n} \left| \sum_{i=1}^n Z_i \right| > \epsilon_0 U_n \right] \leq 2 \exp \frac{-\epsilon_0^2 \log n}{2(1 + \epsilon_0 \sqrt{U_n})} \quad (1.1)$$

$$\leq 2n^{-C'} \epsilon_0^2. \quad (1.2)$$

D'où, pour un choix convenable de ϵ_0 nous déduisons que

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{U_n}).$$

Corollaire 1.6.3[13]

a) S'il existe $M < \infty$, telle que :

$$|Z_1| \leq M$$

Alors nous avons

$$\forall \epsilon > 0, \mathbb{P} \left[\left| \sum_{i=1}^n Z_1 \right| > \epsilon n \right] \leq 2 \exp \frac{-\epsilon^2 n}{2\sigma^2(1 + \frac{M\epsilon}{\sigma^2})}$$

Avec

$$\sigma^2 = \mathbb{E}Z_1^2.$$

b) Supposons que les $(Z_i)_{1 \leq i \leq n}$ dépendent de n (c'est à dire $Z_i = Z_{i,n}$) et que $\sigma^2 = \mathbb{E}Z_i^2$, s'il existe $M < \infty$ telle que :

$$|Z_1| \leq M$$

Et

$$\frac{M}{\sigma_n^2} \leq C < \infty$$

Et si

$U_n = n^{-1}\sigma_n^2 \log n$, vérifie $\lim_{n \rightarrow \infty} U_n = 0$, alors nous avons,

$$\frac{1}{n} \sum_{i=1}^n Z_i = O(\sqrt{U_n}).p.co$$

Démonstration du Corollaire 1.6.3

* En appliquant la proposition (1.6.2) à $a_i^2 = \sigma^2$, $A_n^2 = M$, $b = M$ nous aboutissons à a).

** Comme $\frac{MU_n}{\sigma_n^2}$ tend vers zéro, il suffit de reprendre le résultat a) pour $\epsilon = \epsilon_0 \sqrt{U_n}$, nous arrivons donc à l'existence d'une constante C' telle que :

$$\begin{aligned} \mathbb{P} \left[\frac{1}{n} \left| \sum_{i=1}^n Z_i \right| > \epsilon_0 U_n \right] &\leq 2 \exp \frac{-\epsilon_0^2 \log n}{2(1 + \epsilon_0 \sqrt{\frac{MU_n}{\sigma_n^2}})} \\ &\leq 2n^{-C'} \epsilon_0^2. \end{aligned}$$

1.7 Probabilité des petites boules

Soit X une v.a.f (variable aléatoire fonctionnelle) dans E , x un élément fixe dans E , et pour mieux fixer les idées, on utilise un noyau asymétrique simple de type I. La relation qui lie la pondération locale et la notion de probabilité des petite boule est donnée comme suit :

$$\mathbb{E} \left(1_{[0,1]} \left(\frac{d(x, X)}{h} \right) \right) = \mathbb{E} \left(1_{B(x,h)}(X) \right) = \mathbb{P}(X \in B(x, h)).$$

tel que

$$B(x, h) = \{x' \in E, d(x, x') \leq h\}.$$

Qui est la boule, par rapport à la topologie induite par d , centrée à x et de rayon h (où d est une semi-métrique sur E). Nous disons des probabilité de petite boule, car le paramètre de lissage h (également appelé fenêtre) diminue avec la taille de l'échantillon des variables fonctionnelles (plus précisément, h tend à zéro lorsque n a tendance à 1). Ainsi, lorsque nous prenons une très grande quantité, h est proche de zéro, puis $B(x, h)$ est considérée comme une petite boule et $\mathbb{P}(X \in B(x, h))$ comme une petite probabilité de la petite boule[8].

CHAPITRE 2

ESTIMATION NON PARAMÉTRIQUE DU MODE

CONDITIONNEL

Introduction

L'étude de modèles non paramétriques liés à la distribution conditionnelle a été largement considérée en statistique non paramétrique. Historiquement, il y a plusieurs auteurs qui sont intéressés, par exemple : Perzen (1962)[15] qui a été l'un des premiers à s'intéresser au problème de l'estimation du mode θ d'une densité de probabilité univariée, et L'estimation du mode conditionnelle fut traitée pour la première fois par Collomb et al (1987)[6]. Ces auteurs ont montré la convergence uniforme de l'estimateur à noyau de ce modèle conditionnelle tel que en (1989), Samanta a étudié la normalité asymptotique de l'estimateur à noyau lorsque les observations sont indépendantes et identiquement distribuées et Thavaneswaran en (1990), ont obtenu la même propriété asymptotique de cet estimateur dans le cadre i.i.d. Alors que Ould-said (1997), a étudié l'estimateur à noyau dans le cas des observations ergodiques. Quintela et Vieu (1997) ont traité le mode conditionnel comme étant le point annulant la dérivée d'ordre 1 de la densité conditionnelle et ont construit un estimateur pour ce modèle en utilisant l'estimateur à noyau de la dérivée de la densité conditionnelle.[12]

Le but de ce chapitre est d'étudier les notions de l'estimation du mode conditionnelle, et de la densité conditionnelle par la méthode du noyau, et puis nous établissons la convergence presque complète pour les deux modèles.

2.1 Estimation de la densité conditionnelle

2.1.1 Historique

L'estimation de la fonction de densité conditionnelle et ses dérivées, en statistique fonctionnelle, a été introduite par Ferraty et al[8]. Ces auteurs ont obtenu la convergence presque complète dans le cas i.i.d (independant et identiquement distribué). Alternativement, Ezzahrioui et Ould-said (2005,2006) ont estime le mode conditionnel par le point qui annule la dérivée de l'estimateur à noyau de la densité conditionnelle. Ces derniers se sont concentrés sur la normalité asymptotique de l'estimateur proposé dans les deux cas (i.i.d. et mixing). La précision des termes dominants de l'erreur quadratique de l'estimateur à noyau de la densité conditionnelle à été obtenue par Laksaci (2007). Nous renvoyons à Laksaci et al.(2010) pour la question du choix du paramètre de lissage dans l'estimation de la densité conditionnelle.[17]

2.2 Estimation non paramétrique de la densité conditionnelle

2.2.1 Présentation du modèle

Soit $(X_1, Y_1), \dots, (X_n, Y_n)$ un échantillon aléatoire, indépendant, identiquement distribué, du couple (X, Y) qui est à valeurs dans $E \times \mathbb{R}$, où E est une espace fonctionnelle.

On a f^x la densité conditionnelle de Y sachant $X = x$ donnée par :

$$f^x(y) = \frac{f(x, y)}{f(x)} \quad (2.1)$$

Avec $f(x, y)$ (respectivement $f(x)$) désigne la densité conjointe (respectivement la densité marginale).

Pour tout $y \in \mathbb{R}$. notons $F^x(y) = \mathbb{P}(Y \leq y \mid X = x)$ la fonction de répartition conditionnelle de Y sachant X . Où cette fonction est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R} . Notons aussi par : $B(x, h) = \{x' \in E \mid d(x', x) < h\}$ la boule de centre x et de rayon h , N_x un voisinage de x , et d est une semi-métrique dans E .

L'estimateur à noyau de la fonction de répartition conditionnelle F^x est défini par :

$$\hat{F}^x(y) = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i))H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(x, X_i))}. \forall y \in \mathbb{R} \quad (2.2)$$

avec la convention $\frac{0}{0} = 0$.

De $\hat{F}^x(y)$, nous concluons que un estimateur de la densité conditionnelle, noté $\hat{f}^x(y)$, défini par :

$$\hat{f}^x(y) = \frac{\partial \hat{F}^x}{\partial y}(y) \quad (2.3)$$

En supposant que H est différentiable, défini par

$$\forall u \in \mathbb{R}, H(u) = \int_{-\infty}^u K_0(v)dv$$

nous avons

$$\frac{\partial \hat{F}^x}{\partial y}(y) = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i)) \frac{\partial}{\partial y} H(g^{-1}(y - Y_i))}{g \sum_{i=1}^n K(h^{-1}d(x, X_i))}.$$

d'où

$$\hat{f}^x(y) = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i))H'(g^{-1}(y - Y_i))}{g \sum_{i=1}^n K(h^{-1}d(x, X_i))}.$$

donc l'estimateur de la densité conditionnelle est donnée par :

$$\hat{f}^x(y) = \frac{\sum_{i=1}^n K(h^{-1}d(x, X_i))HK_0(g^{-1}(y - Y_i))}{g \sum_{i=1}^n K(h^{-1}d(x, X_i))}. \quad (2.4)$$

et d est une semi-métrie.

Avec K est un noyau de type I ou de type II et la fonction K_0 est un noyau symétrique standard, et $h = h(n)$ (resp $g = g(n)$) est une suite de nombres réels positifs qui tend vers 0 lorsque n tend vers l'infinie (il est aussi appelé le paramètre de lissage).

2.3 La convergence presque complète de la densité conditionnelle

Afin d'établir la convergence presque complète de notre estimateur, il est nécessaire d'introduire les hypothèses de régularité sur la fonction de densité conditionnelle, et des conditions sur la concentration de la variable fonctionnelle X dans les petites boules.

2.3.1 Hypothèses

Maintenant, nous introduisons les hypothèse de base permettant de donner un théorème général sur la convergence presque complète.

Soient x (resp. y) un élément de E , $N_x \in E$ un voisinage de x et S un sous espace compact de \mathbb{R} .

(H.1) La probabilité de concentration de la variable fonctionnelle X dans les petites boules est non nulle.

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

(H.2) Nous supposons que K satisfait la condition de Hölder suivante :

$$\exists C < \infty, \forall (u, u') \in \mathbb{R} \times \mathbb{R}, |K_0(u) - K_0(u')| \leq C |u - u'|.$$

Et

$$\exists \zeta > 0, \lim_{n \rightarrow \infty} gn^\zeta = \infty \text{ et } \lim_{n \rightarrow \infty} \frac{\log n}{ng\varphi_x(h)} = 0.$$

(H.3) La condition sur le paramètre de lissage h

- h est un paramètre positive.

Avec

$$\lim_{n \rightarrow +\infty} h = 0$$

et

$$\lim_{n \rightarrow +\infty} \frac{\log n}{n\varphi_x(h)} = 0$$

- K est une noyau de type I ou de type II.

Et la fonction des probabilités conditionnelles des petites boules $\varphi_x(\cdot)$, vérifie :

$$\exists C > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > C \in \varphi_x(\epsilon).$$

(H.4) Le paramètre de lissage g est tel que :

- g est un paramètre positive.

Avec

$$\lim_{n \rightarrow +\infty} g = 0.$$

- K_0 est un noyau de type 0.

(H.5) Nous supposons que f satisfait les conditions de la continuité et de la monotoné suivante :

$$C_{E,\mathbb{R}}^0 = \left\{ \begin{array}{l} f : E \times \mathbb{R} \longrightarrow \mathbb{R} \\ \forall x' \in N_x, \lim_{d(x,x') \rightarrow 0} f(x', y) = f(x, y) \\ \text{et } \forall y' \in \mathbb{R}, \lim_{|y'-y|} f(x, y') = f(x, y). \end{array} \right\}$$

et

$$S_{dens}^x = \left\{ \begin{array}{l} f : E \times \mathbb{R} \longrightarrow \mathbb{R} \\ \exists \xi > 0, \exists ! y_0 \in S, f^x \text{ est strictement décroissante sur } (y_0 - \xi, y_0) \\ \text{et strictement croissante sur } (y_0, y_0 + \xi). \end{array} \right\}$$

Alors

$$f^x(y) \in C_{E,\mathbb{R}}^0 \cap S_{dens}^x$$

(H.6) Nous supposons que f satisfait la condition de lipschitz suivante :

$$Lip_{E \times \mathbb{R}}, \beta = \left\{ \begin{array}{l} f : E \times \mathbb{R} \longrightarrow \mathbb{R} \\ \exists C_x > 0 \text{ tel que } \forall (y_1, y_2) \in S^2, \forall (x_1, x_2) \in N_x \times N_x, \\ |f(y_1, x_1) - f(y_2, x_2)| \leq C_x (d^{\beta_1}(x_1, x_2) + |y_1 - y_2|^{\beta_2}), \beta_1 > 0, \beta_2 > 0. \end{array} \right\}$$

(H.7)

$$\begin{cases} f^x(j) \text{ fois continuellement différentiable} \\ \lim_{n \rightarrow \infty} \frac{\log n}{ng^{2j-1}\varphi_x(h)} = 0 \end{cases}$$

(H.8)

$$\begin{cases} f^{x(l)}(\theta(x)) = 0 \quad \forall l = 1, \dots, j-1 \\ f^{x(l)}(\theta(x)) \neq 0 \end{cases}$$

Le théorème suivant donne la convergence presque complète de l'estimateur $\hat{f}^x(y)$.

Théorème 2.3.1 (convergence presque complète sous les conditions de la continuité) [8]

Sous les conditions (H.1), (H.2), (H.3), (H.4) et (H.5), nous avons pour tout sous-ensemble compact $S \subset \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |f^x(y) - \hat{f}^x(y)| = 0, \quad p.co. \quad (2.5)$$

Preuve

La preuve est basée sur la décomposition suivante :

$$\hat{f}^x(y) - f^x(y) = \frac{\{(\hat{r}_4(x, y) - \mathbb{E}\hat{r}_4(x, y)) - (f^x(y) - \mathbb{E}\hat{r}_4(x, y))\}}{\hat{r}_1(x)} - \frac{f^x(y)}{\hat{r}_1(x)} \{\hat{r}_1(x) - 1\} \quad (2.6)$$

Où \hat{r}_1 est défini par :

$$\hat{r}_1(x) = \frac{1}{n} \sum_{i=1}^n \Delta_i.$$

Avec

$$\Delta_i = \frac{K(h^{-1}d(x, X_i))}{\mathbb{E}K(h^{-1}d(x, X_1))}$$

Et

$$r_4 = \hat{r}_1(x) \hat{f}^x(y) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Omega_i(y)$$

Avec

$$\Omega_i(y) = g^{-1} K_0(g^{-1}(y - Y_i)).$$

Les lemmes et la proposition suivants seront utiles dans la suite de cette preuve :

Lemme2.3.1

Sous les conditions **(H.1)** et **(H.3)**, on a :

$$\hat{r}_1 - 1 = O_{p.co} \left(\sqrt{\frac{\log n}{n\varphi_x(h)}} \right). \quad (2.7)$$

Preuve (Voir [8])

Lemme2.3.2

Si K est un noyau de type I, il existe des constantes réelles C et C' non négatives tel que :

$$C\varphi_x(h) \leq \mathbb{E}K \left(\frac{d(x, X)}{h} \right) \leq C'\varphi_x(h). \quad (2.8)$$

Preuve (Voir [8])

Proposition2.3.3

On suppose que $\lim_{n \rightarrow \infty} U_n = 0$, $X_n = O_{p.co}(U_n)$, $\lim_{n \rightarrow \infty} Y_n = l_Y$ où l_Y est un nombre réel déterministe, nous avons :

$$X_n Y_n = O_{p.co}(U_n). \quad (2.9)$$

Lemme2.3.4

Si K est un noyau de type II et la vérifie que :

$$\exists C_3 > 0, \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > C_3 \epsilon \varphi_x(\epsilon)$$

Alors, il existe des constantes réelles positifs C et C' telle que, pour h suffisamment petit

$$C\varphi_x(h) \leq \mathbb{E}K \left(\frac{d(x, X)}{h} \right) \leq C'\varphi_x(h). \quad (2.10)$$

Preuve (Voir [8])

Notons d'abord que les dénominateurs de (2.6) sont directement traité en utilisant le lemme 2.3.1 et la proposition 2.3.3. On a

$$\hat{r}_1(x) - 1 = O_{p.co} \left(\sqrt{\frac{\log n}{n\varphi_x(h)}} \right)$$

Avec

$$U_n = \sqrt{\frac{\log n}{n\varphi_x(h)}}$$

On suppose que :

$$X_n = \hat{r}_1(x) - 1 = O_{p.co}(U_n)$$

Alors

$$\frac{f^x(y)}{\hat{r}_1} \{\hat{r}_1 - 1\} = O_{p.co}(U_n).$$

On remarque aussi que le dernier terme est traité en utilisant aussi lemme 2.3.1, et que $f^x(y)$ est uniformément borné sur $y \in S$ (puisque'il est continu sur compact S).

Par conséquent, le résultat (2.5) sera conséquence directe des deux résultats suivants :

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\mathbb{E} \hat{r}_4(x, y) - f^x(y)| = 0, \quad p.co. \quad (2.11)$$

et

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - \mathbb{E} \hat{r}_4(x, y)| = 0, \quad p.co. \quad (2.12)$$

Le resultat de (2.11) est confirmée, parceque $\mathbb{E} \Delta_1 = 1$ et $\int K_0 = 1$, nous avons après integration par substitution.

$$\begin{aligned} \mathbb{E} \hat{r}_4(x, y) - f^x(y) &= \mathbb{E} \Delta_1 \Omega_1(y) - f^x(y) \\ &= \mathbb{E} (\Delta_1 \mathbb{E}(\Omega_1(y)/x) - f^x(y)) \\ &= \mathbb{E} \left(\Delta_1 \left(\int_{\mathbb{R}} g^{-1} K_0(g^{-1}(y-u)) f^x(u) du - f^x(y) \right) \right) \\ &= \mathbb{E} \left(\Delta_1 \left(\int_{\mathbb{R}} g^{-1} K_0(g^{-1}(y-u)) (f^x(u) - f^x(y)) du \right) \right) \\ &= \mathbb{E} \left(\Delta_1 \left(\int_{\mathbb{R}} K_0(v) (f^x(y - vg) g^{-1} - f^x(y)) dv \right) \right) \\ &= \mathbb{E} \left(\mathbf{1}_{B(x,h)}(x) \Delta_1 \int_{\mathbb{R}} K_0(v) (f^x(y - vg) - f^x(y)) dv \right) \end{aligned} \quad (2.13)$$

Nous pouvons écrire

$$|f^x(y - vg) - f^x(y)| \leq |f^x(y - vg) - f^x(y - vg)| + |f^x(y - vg) - f^x(y)|$$

Puisque S est compact, la fonction f^x est uniformément continue sur S , avec K_0 est support compact(car il est du type 0), et parceque l'hypothèse **(H.5)**(modèle de continuité) est verifié, alors on a :

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, +1]} \sup_{y \in S} \mathbf{1}_{B(x, h)} |f^x(y - vg) - f^x(y)| = 0$$

et

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, +1]} \sup_{y \in S} |f^x(y - vg) - f^x(y)| = 0$$

ce qui implique que

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, +1]} \sup_{y \in S} \mathbf{1}_{B(x, h)} |f^x(y - vg) - f^x(y)| = 0 \quad (2.14)$$

En combinant(2.13) et (2.14), avec la positivité de Δ_1 et K_0 , on trouve

$$\sup_{y \in S} |\mathbb{E}\hat{r}_4(x, y) - f^x(y)| = O(1) \quad (2.15)$$

Ceci combine au lemme2.3.1 et la proposition2.3.3, suffit à prouver (2.11) donc, il reste à verifer (2.12) est vrai. En utilisant la compacité de S on peut écrire que $S \subset \bigcup_{k=1}^{z_n} S_k$ où $S_k = (t_k - l_n, t_k + l_n)$ et ou l_n et z_n peuvent être choisis de telle sorte que :

$$l_n = Cz_n^{-1} \sim C_n n^{2\zeta} \quad (2.16)$$

En prenant

$$t_y = \arg \min_{t \in \{t_1, \dots, t_n\}} |y - t|$$

on a

$$\frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - \mathbb{E}\hat{r}_4(x, y)| = A_1 + A_2 + A_3 \quad (2.17)$$

Où

$$A_1 = \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, y) - \hat{r}_4(x, t_y)|$$

$$A_2 = \frac{1}{\hat{r}_1(x)} \sup_{y \in S} |\hat{r}_4(x, t_y) - \mathbb{E}\hat{r}_4(x, t_y)|$$

$$A_3 = \frac{1}{\hat{r}_1(x)} \sup_{y \in S} | \mathbb{E} \hat{r}_4(x, t_y) - \mathbb{E} \hat{r}_4(x, y) |$$

La condition de continuité de Hölder nous permet d'écrire directement

$$\begin{aligned} | \hat{r}_4(x, y) - \hat{r}_4(x, t_y) | &= \frac{1}{n} \sum_{i=1}^n \Delta_i | \Omega_i(y) - \Omega_i(t_y) | \\ &\leq \frac{1}{ng} \sum_{i=1}^n \Delta_i | K_0(g^{-1}(y - Y_i)) - K_0(g^{-1}(t_y - Y_i)) | \\ &\leq \frac{C}{ng} \sum_{i=1}^n \Delta_i \frac{|y - t_y|}{g} \\ &\leq C \hat{r}_1(x) l_n g^{-2} \end{aligned}$$

En utilisant (2.16), on trouve que

$$A_1 \leq \frac{C}{(gn^\zeta)^2}$$

et **(H.2)** implique que :

$$\lim_{n \rightarrow \infty} A_1 = 0 \quad (2.18)$$

Après des arguments similaires, nous pouvons écrire

$$A_3 \leq \frac{C}{\hat{r}_1(x)(gn^\zeta)^2} \quad (2.19)$$

et selon le lemme 2.3.1 et la proposition 2.3.3, nous obtenons

$$\lim_{n \rightarrow \infty} A_3 = 0, p.co \quad (2.20)$$

En regardant maintenant le terme A_2 nous pouvons écrire pour tout $\epsilon > 0$

$$\begin{aligned} \mathbb{P} \left(\sup_{y \in S} | \hat{r}_4(x, t_y) - \mathbb{E} \hat{r}_4(x, t_y) | \right) > \epsilon &= \mathbb{P} \left(\max_{j=1, \dots, z_n} | \hat{r}_4(x, t_j) - \mathbb{E} \hat{r}_4(x, t_j) | > \epsilon \right) \\ &\leq z_n \max_{j=1, \dots, z_n} \mathbb{P} (| \hat{r}_4(x, t_j) - \mathbb{E} \hat{r}_4(x, t_j) | > \epsilon) \\ &\leq z_n \max_{j=1, \dots, z_n} \mathbb{P} \left(\left| \frac{1}{n} \sum_{i=1}^n (U_i - \mathbb{E} U_i) \right| > \epsilon \right) \quad (2.21) \end{aligned}$$

En utilisant le lemme 2.3.2 et le lemme 2.3.4 selon le fait que K est de type I ou II, et parce que $\Omega_i(y) \leq \frac{C}{h}$, nous avons :

$$| U_i | \leq \frac{C}{g\varphi_x(h)} \quad (2.22)$$

D'autre part, nous avons après intégration et par substitution et en utilisant(2.14)

$$\begin{aligned}
\mathbb{E}U_i^2 &= \mathbb{E}\left(\Delta_i^2 \mathbb{E}(\Omega^2(t_j)|X=x)\right) \\
&= \mathbb{E}\left(\Delta_i^2 \left(\int_{\mathbb{R}} g^{-2} K_0^2(g^{-1}(t_j-u)) f^x(u) du\right)\right) \\
&= \mathbb{E}\left(\Delta_i^2 g^{-1} \int_{\mathbb{R}} K_0^2(z) f^x(t_j+zg) dz\right) \\
&\leq C \mathbb{E}\left(\Delta_i^2 g^{-1} \int_{\mathbb{R}} K_0^2(z) dz f^x(t_j)\right)
\end{aligned}$$

Parce que f^x est bornée (car elle est continue sur l'ensemble compact S) et en appliquant le résultat (6.9) donnée dans Ferraty(voir[8])

$$\mathbb{E} | Y_1 |^m \Delta_1^m = O(\varphi_x(h)^{-m+1}). \text{ avec } m = 2$$

On obtient

$$\mathbb{E}U_i^2 \leq \frac{C}{g\varphi_x(h)} \quad (2.23)$$

Parce que les variables U_i est bornées, on est donc en mesure d'appliquer l'inégalité exponentielle du Bernstein donnée par corollaire1.6.2 cette inégalité avec (2.21),(2.22) et (2.23) donne directement

$$\mathbb{P}\left(\sup_{y \in S} |\hat{r}_4(x, t_y) - \mathbb{E}\hat{r}_4(x, t_y)| > \epsilon\right) \leq z_n \exp(-Cn\epsilon^2 g\varphi_x(h))$$

En utilisant (2.16), on obtient

$$\mathbb{P}\left(\sup_{y \in S} |\hat{r}_4(x, t_y) - \mathbb{E}\hat{r}_4(x, t_y)| > \epsilon\right) \leq Cn^{2\zeta} \exp(-Cn\epsilon^2 g\varphi_x(h))$$

Parce que $\left(\frac{\log n}{ng\varphi_x(h)}\right)$ tend vers zéro, on obtient directement

$$\forall \epsilon > 0, \sum_{i=1}^{\infty} \mathbb{P}\left(\sup_{y \in S} |\hat{r}_4(x, t_y) - \mathbb{E}\hat{r}_4(x, t_y)| > \epsilon\right) < \infty \quad (2.24)$$

Le dénominateur de A_2 est traité directement en utilisant à nouveau le lemme2.3.1 avec la proposition2.3.3. C'est suffisant pour obtenir

$$\lim_{n \rightarrow \infty} A_2 = 0, \quad p.co \quad (2.25)$$

Finalement, le résultat(2.12) se déduit (2.17), (2.18), (2.20) et (2.25). La combinaison (2.10), (2.11) et (2.12) nous permet de terminer la preuve de ce théorème.

Le théorème suivant donne la vitesse de convergence presque complète ponctuelle de \hat{f}^x .

Théorème2.3.2 (convergence presque complète ponctuelle sous la condition de lipschitz)[19]

Sous les conditions(H.1),(H.2),(H.3),(H.6) et(2.4), nous avons pour tout nombre réel fixé y :

$$\hat{f}^x(y) - f^x(y) = O(h^{\beta_1}) + O(g^{\beta_2}) + O\left(\sqrt{\frac{\log n}{ng\varphi_x(h)}}\right). \quad (2.26)$$

Preuve

La preuve est basée sur la décomposition qui sont :

$$\hat{f}^x(y) - f^x(y) = \frac{\{(\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y)) - (f^x(y) - \mathbb{E}\hat{r}_3(x, y))\}}{\hat{r}_1(x)} - \frac{f^x(y)}{\hat{r}_1(x)}\{\hat{r}_1(x) - 1\}. \quad (2.27)$$

Où \hat{r}_1 est défini par

$$\hat{r}_1(x) = \frac{1}{n} \sum_{i=1}^n \Delta_i. \quad (2.28)$$

Et où

$$r_3 = \hat{r}_1(x)\hat{f}^x(y) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Omega_i(y) \quad (2.29)$$

Avec

$$\Omega_i(y) = g^{-1}K_0(g^{-1}(y - Y_i)). \quad (2.30)$$

Ainsi, la preuve est une conséquence direct des résultats qui suivant.

Lemme2.3.2

Sous les hypothèses (H.1), (H.6), lorsque n tend vers l'infini, nous avons :

$$\mathbb{E}\hat{r}_3(x, y) - f^x(y) = O(h^{\beta_1}) + O(g^{\beta_2}). \quad (2.31)$$

Preuve

Puisque $\mathbb{E}\Delta_i = 1$ et puisque K_0 est une fonction intégrable, nous avons :

$$\begin{aligned}
\mathbb{E}\hat{r}_3(x, y) - f^x(y) &= \mathbb{E}\Delta_1\Omega_1(y) - f^x(y) \\
&= \mathbb{E}(\Delta_1(\mathbb{E}(\Omega_1(y)|x) - f^x(y))) \\
&= \mathbb{E}\left(\Delta_1 \int_{\mathbb{R}} \frac{1}{g} K_0\left(\frac{y-u}{g}\right) f^x(u) du - f^x(y)\right) \\
&= \mathbb{E}\left(\Delta_1 \int_{\mathbb{R}} \frac{1}{g} K_0(v)(f^x(y - vg) - f^x(y)) dv\right) \\
&= \mathbb{E}\left(\Delta_1 \int_{\mathbb{R}} \frac{1}{g} K_0(v)(f^x(y - vg) - f^x(y)) dv\right) \\
&= \mathbb{E}\left(\mathbf{1}_{B(x,h)} \Delta_1 \int_{\mathbb{R}} \frac{1}{g} K_0(v)(f^x(y - vg) - f^x(y)) dv\right). \tag{2.32}
\end{aligned}$$

Compte tenu du fait que le support $K_0 = [-1, 1]$ et puisque h et g tendent vers zéro, la condition de Hölder **(H.2)** permet d'écrire que :

$$\sup_{v \in [-1, 1]} |f^x(y - vg) - f^x(y)| = O(h^{\beta_1}) + O(g^{\beta_2}). \tag{2.33}$$

Ainsi, le résultat (2.31) découle directement en combinant (2.32), (2.33) et le fait que $\mathbb{E}\Delta_1 = 1$.

Lemme 2.3.3

Sous les hypothèses du théorème, lorsque n tends vers l'infini, nous avons :

$$\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y) = O_{p.co} \left(\sqrt{\frac{\log n}{ng\varphi_x(h)}} \right). \tag{2.34}$$

Preuve

On utilise la décomposition suivante :

$$\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y) = \frac{1}{n} \sum_{i=1}^n Z_i \tag{2.35}$$

Avec

$$Z_i = (T_i - \mathbb{E}T_i).$$

Et

$$T_i = \Delta_i \Omega_i(y).$$

Afin d'appliquer l'inégalité de type-Bernstein, montrons d'abord

$$|T_i| \leq \frac{C}{g\varphi_x(h)}. \quad (2.36)$$

Et

$$ET_i^2 \leq \frac{C}{g\varphi_x(h)}. \quad (2.37)$$

En utilisant le Lemme2.3.2 ou le Lemme2.3.4 et en tenant compte de l'hypothèse **[(H.3)]** et du fait que K est de type I ou de type II, nous avons :

$$C\varphi_x(h) \leq \mathbb{E}K \left(\frac{d(x, X_i)}{h} \right) \leq C'\varphi_x(h).$$

et en utilisant le dernier résultat et puisque K_0 est borné, nous obtenons

$$|T_i| \leq \frac{C}{g\varphi_x(h)}.$$

Le second moment des variables T_i peut être calculé en utilisant l'intégration par changement de variable :

$$\begin{aligned} \mathbb{E}T_i^2 &= \mathbb{E} \left(\Delta_i^2 \frac{1}{g^2} K_0^2 \left(\frac{y - Y_i}{g} \right) \right) \\ &= \mathbb{E} \left(\mathbb{E} \left(\Delta_i^2 \frac{1}{g^2} K_0^2 \left(\frac{y - Y_i}{g} \mid X = x \right) \right) \right) \\ &= \frac{1}{g^2} \mathbb{E} \left(\Delta_i^2 \int_{\mathbb{R}} K_0^2 \left(\frac{y - Y_i}{g} \right) f^x(u) du \right) \\ &\leq \frac{g}{g^2} \mathbb{E} \left(\Delta_i^2 \int_{\mathbb{R}} K_0^2(v) f^x(y - vg) dv \right) \\ &\leq \frac{C}{g} \mathbb{E}\Delta_i^2 \end{aligned}$$

Avec C est une constant.

Puisque $0 < \int K^2 < \infty$, si K est de type I (resp II) alors $\left(\frac{K^2}{\int K^2} \right)$ est aussi de type I (resp II).

Ainsi, en appliquant le Lemme2.3.2 ou le Lemme2.3.4 on trouve

$$C\varphi_x(h) \leq \mathbb{E}K^2 \left(\frac{d(x, X_i)}{h} \right) \leq C'\varphi_x(h).$$

et en utilisant le dernier résultat

$$\frac{C}{\varphi_x(h)} \leq \mathbb{E}\Delta_i \leq \frac{C'}{\varphi_x(h)}$$

ce qui implique

$$\mathbb{E}T_i^2 \leq \frac{C}{g\varphi_x(h)}$$

En tenant compte de (2.36), on peut appliquer l'inégalité de type-Bernstein type donné par le corollaire A.9 - i(voir Ferraty et al[8]) et on obtient :

$$\forall \epsilon \geq 0, \mathbb{P}(|\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y)| > \epsilon) \leq 2 \exp\left(\frac{\epsilon^2 ng\varphi_x(h)}{2C'(1 + \epsilon)}\right). \quad (2.38)$$

Puisque la suite $\left(\frac{\log n}{ng\varphi_x(h)}\right)$ tend vers zéro, en choisissant $\epsilon = \epsilon_0 \sqrt{\frac{\log n}{ng\varphi_x(h)}}$ dans le résultat (2.38) nous obtenons directement

$$\mathbb{P}(|\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y)|) > \epsilon_0 \sqrt{\frac{\log n}{ng\varphi_x(h)}} \leq 2 \exp\frac{\epsilon_0^2 \log n}{2C' \left(1 + \epsilon_0 \sqrt{\frac{\log n}{ng\varphi_x(h)}}\right)} \quad (2.39)$$

$$\leq 2n^{-C\epsilon_0^2} \quad (2.40)$$

et il en résulte que pour ϵ_0 assez large ($\epsilon_0 > p\frac{1}{\sqrt{C}}$) :

$$\sum_{i=1}^{\infty} \mathbb{P}\left(|\hat{r}_3(x, y) - \mathbb{E}\hat{r}_3(x, y)| > \epsilon_0 \sqrt{\frac{\log n}{ng\varphi_x(h)}}\right) < +\infty. \quad (2.41)$$

Ainsi, la preuve de (2.34) est maintenant achevée.

Lemme 2.3.4

Sous les hypothèses du théorème, lorsque n tend vers l'infini, nous avons :

$$\hat{r}_1(x) - 1 = O_{p.co}\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right). \quad (2.42)$$

Notons que les dénominateurs introduits dans la décomposition (2.27) sont directement traités en utilisant les lemmes ci-dessus et la proposition A.6-i(voir Ferraty et al[8]).

Preuve

Ce résultat peut directement se déduire du lemme 2.3.3 en prenant $\Delta_i(y) = 1$.

2.4 Estimation non paramétrique du mode conditionnel

La littérature sur l'estimation du mode conditionnel est très abondante lorsque la variable explicative est à valeurs dans un espace de dimension finie. Citons, par exemple, Quintela del Rio et Vieu, Ould-Saïd, et Louani et Ould-Saïd. Cette littérature est relativement restreinte lorsque la variable explicative est de nature fonctionnelle. Dans ce dernier cas, les premiers résultats ont été obtenus par Ferraty et al[8]. Ils ont étudié la convergence presque complète d'un estimateur à noyau du mode conditionnel pour des observations i.i.d.

Dans cette contribution on se propose d'étudier la convergence presque complète du mode conditionnel dans le cadre non paramétrique pour une variable fonctionnelle dans le cas indépendant.

2.4.1 Le mode conditionnel

Le mode est très populaire en classification, car c'est un outil utile pour représenter des groupes et aussi plus robuste que la moyenne (comme la médiane).

Estimer le mode est souvent une conséquence directe de l'estimation de densité. Son importance est due au fait que c'est une mesure naturelle de tendance centrale, qui n'est pas influencée par les queues des distributions. Le mode est la valeur la plus probable : pour une densité f , c'est la valeur pour laquelle f admet un maximum (global ou local). Pour une distribution symétrique, il coïncide avec deux autres paramètres de position, la moyenne et la médiane[7].

2.4.2 Présentation du modèle

Considérons $(X_i, Y_i)_{i=1, \dots, n}$ un échantillon de couple de variables aléatoire (X, Y) , où Y est à valeurs dans \mathbb{R} et X est à valeurs dans un espace semi-métrique (E, d) qui peut être de dimension éventuellement finie, par définition, Le mode conditionnel est la valeur qui maximise la densité conditionnelle Y sachant $X = x$ [5]. On suppose qu'il existe un sous-ensemble compact $S \in \mathbb{R}$, où le mode est unique noté par $\theta(x)$:

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

Avec

$$f^x(y) = \frac{f(x, y)}{f(x)}$$

Où $f^x(y)$ est la densité conditionnelle de Y sachant $X = x$, et $f(x, y)$ (respectivement $f(x)$) désigne la densité conjoint (respectivement la densité marginale).

L'estimateur de ce mode $\hat{\theta}(x)$ est définie par l'expression :

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

Avec

$$\hat{f}^x(y) = \frac{\sum_{i=1}^n K((h^{-1}d(x, X_i))H'(h^{-1}(y - Y_i)))}{g \sum_{i=1}^n K(h^{-1}d(x, X_i))}.$$

et d est une distance dans E .

2.5 La convergence presque complète du mode conditionnel

Dans cette section on utilise les mêmes hypothèses données dans la section (2.3.1), pour montrer la convergence presque complète de l'estimateur $\hat{\theta}$ dans le cas continu.

Théorème 2.6.1 [8]

Si les conditions (H.1), (H.2), (H.3), (H.4) et (H.5), sont vérifiées alors :

$$\lim_{n \rightarrow \infty} \hat{\theta}(x) = \theta(x). \quad p.co \quad (2.43)$$

Preuve

La condition (H.5) assure que la vraie densité conditionnelle f^x est continue et strictement croissante sur $(\theta(x) - \xi, \theta(x))$.

Alors, la fonction f^{x-1} existe et est continue. La propriété de continuité de f^{x-1} au point $f^x(\theta(x))$ peut être écrite pour tout $\epsilon > 0$:

$$\exists \delta_1(\epsilon) > 0, \forall y \in (\theta(x) - \xi, \theta(x)), |f^x(y) - f^x(\theta(x))| \leq \delta_1(\epsilon) \implies |y - \theta(x)| < \epsilon.$$

Puisque $f^x(y)$ est continue et strictement décroissante sur $(\theta(x), \theta(x) + \xi)$, le même type d'argument peut être utilisé pour arriver à :

$$\exists \delta_2(\epsilon) > 0, \forall y \in (\theta(x), \theta(x) + \xi), |f^x(y) - f^x(\theta(x))| \leq \delta_2(\epsilon) \implies |y - \theta(x)| < \epsilon.$$

En combinant les deux résultats, on arrive

$$\exists \delta(\epsilon) > 0, \forall y \in (\theta(x) - \xi, \theta(x) + \xi), |f^x(y) - f^x(\theta(x))| \leq \delta(\epsilon) \implies |y - \theta(x)| < \epsilon.$$

Parceque par constuction $\hat{\theta}(x) \in (\theta(x) - \xi, \theta(x) + \xi)$, nous avons :

$$\exists \delta(\epsilon) > 0, |f^x(y) - f^x(\theta(x))| \leq \delta(\epsilon) \implies |y - \theta(x)| < \epsilon.$$

Finalement , nous arrivons à :

$$\exists \delta(\epsilon) > 0, \mathbb{P}(|\hat{\theta}(x) - \theta(x)| > \epsilon) \leq \mathbb{P}(|f^x(\hat{\theta}(x)) - f^x(\theta(x))| > \delta(\epsilon)).$$

D'autre part, il découle directement des définition de $\theta(x)$ et $\hat{\theta}(x)$ que :

$$\begin{aligned} |f^x(\theta(x)) - f^x(\hat{\theta}(x))| &= f^x(\theta(x)) - f^x(\hat{\theta}(x)) \\ &= f^x(\theta(x)) - f^x(\hat{\theta}(x)) + \hat{f}^x(\theta(x)) - \hat{f}^x(\theta(x)) \\ &= (f^x(\theta(x)) - \hat{f}^x(\theta(x))) + (\hat{f}^x(\theta(x)) - f^x(\hat{\theta}(x))) \\ &\leq (f^x(\theta(x)) - f^x(\theta(x))) + (\hat{f}^x(\theta(\hat{x})) - f^x(\hat{\theta}(x))) \\ &\leq 2 \sup_{y \in (\theta(x) - \xi, \theta(x) + \xi)} |f^x(y) - \hat{f}^x(y)|. \end{aligned}$$

l'uniforme convergence sur l'ensemble compact $[\theta(x) - \xi, \theta(x) + \xi]$ de l'estimateur de la densité conditionnelle (voir Théorème 2.3.1) peut être utilisée, donnant directement à partir des deux précédentes inégalités : La convergence complète peut être utilisée, donnant directement à partir des deux précédentes inégalités :

$$\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|\hat{\theta}(x) - \theta(x)| > \epsilon) < \infty$$

Finalement le résultat de (2.43) est vérifié

Le théorème suivant donne la convergence presque complète du mode conditionnel dans le cas lipschitzienne.

Théorème 2.6.2 [8]

Sous le modèle de type lipschitz défini par (H.6) et (H.8), et si la variable fonctionnelle X satisfait (H.1), et si l'estimation du noyau satisfait (H.3), (H.4), (H.2) et (H.7) alors, nous avons :

$$\hat{\theta}(x) - \theta(x) = O((h^{b_1} + g^{b_2})^{\frac{1}{j}}) + O_{p.co} \left(\left(\sqrt{\frac{\log n}{ng\varphi_x(h)}} \right)^{\frac{1}{j}} \right).$$

Preuve (Voir [8])

Introduction

Dans ce chapitre, nous avons présenté graphiquement les courbes des variables aléatoires fonctionnelles, et nous avons aussi déterminé l'estimation non paramétrique fonctionnelle avec la méthode du noyau des réponses Y_i à partir d'un échantillon des courbes X_i .

3.1 Choix du langage de programmation

R est à la fois un outil et un langage dédié au traitement des données statistiques (calculs, modélisation, production de graphiques). Il appartient à la famille des logiciels de calcul numérique, par opposition aux logiciels de calcul symbolique. Il est né au début des années 1990 à l'initiative de deux professeurs de l'université d'Auckland en Nouvelle-Zélande, Ross Ihaka et Robert Gentleman, qui s'étaient fixés comme objectif de créer un outil basé sur le langage de programmation Scheme et inspiré du nouveau langage statistique S proposé par Rick Becker et John Chambers[4].

Dans ce mémoire, nous avons utilisé le langage de programmation R version 3.5.3, pour estimer et aussi pour tracer les différents graphes.

Les fonctions et les codes R utilisés dans cette section pour obtenir nos résultats sont disponibles à l'adresse : <http://www.Isp.upstlse.fr/staph/npfda>.

3.2 Simulation des données fonctionnelles

Nous simulons 100 variables X_i , à laide du processus suivant :

$$X(i, j) = \cos(\omega_i + \pi(\frac{2j}{100} - 1)).$$

Avec ω_i distribué selon la loi Normal centré réduite $N(0, 1)$.

3.2.1 Résultats de simulation

3.2.1.1 Résultats de plusieurs courbes simulées

- les courbes X_i sont représentés dans le graphe suivant :

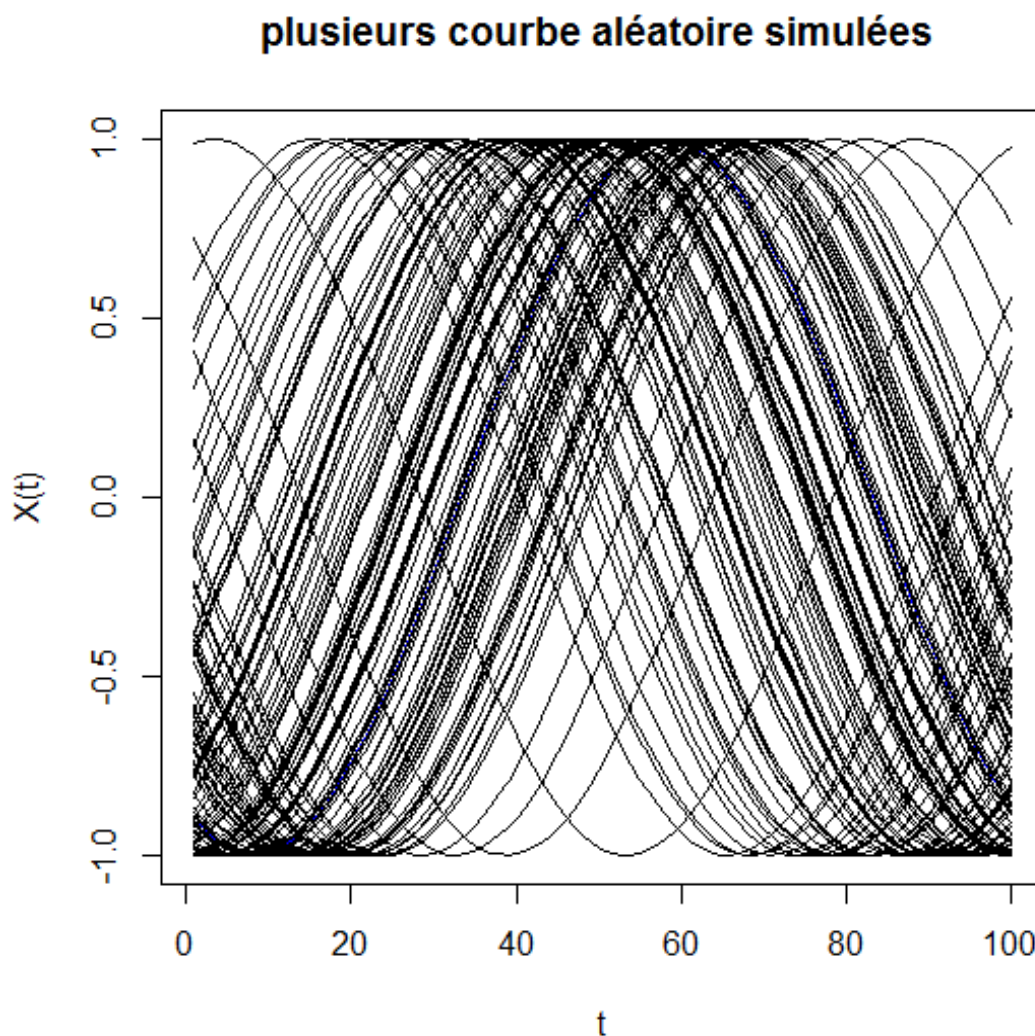


FIGURE 3.1 – Courbes de plusieurs v.a.f

3.2.1.2 Résultat d'une variable aléatoire fonctionnelle

La figure (3.2) représente une courbe aléatoire simulée

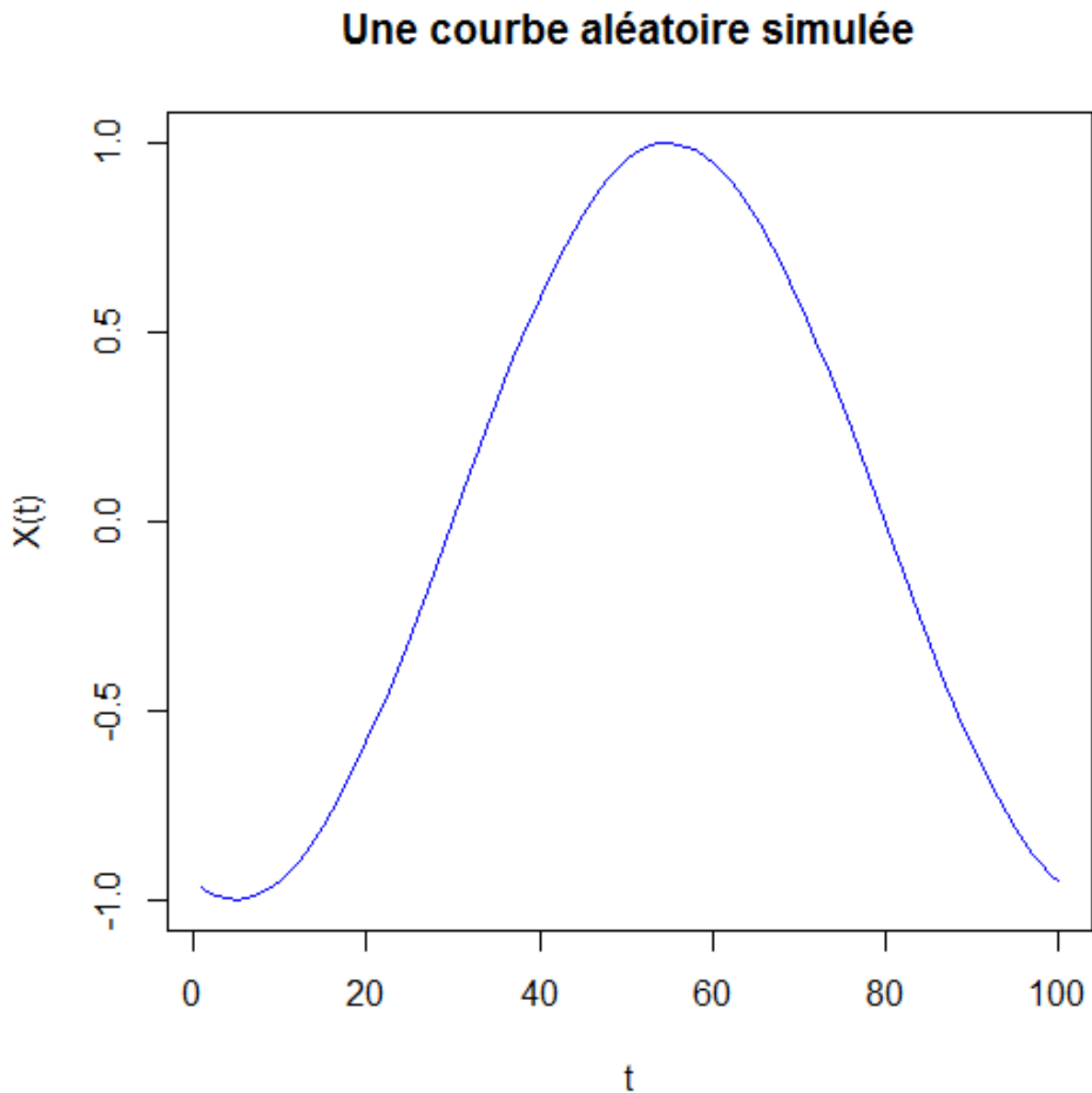


FIGURE 3.2 – Une courbe de v.a.f

3.3 plan d'estimation du mode conditionnel

Pour simuler les variables les réponses $Y_i, i = 1, \dots, 100$, nous donnerons le modèle de régression suivant :

$$Y_i = r(X_i) + \epsilon_i$$

Avec $r(X) = \frac{X^4}{4}$ (X un vecteur qui représente une variable aléatoire fonctionnelle).

Et l'erreur ϵ_i est une variable aléatoire suivant la loi Normal centrée réduite vérifie $\mathbb{E}(\epsilon_i/X) = 0$.

Nous partageons notre échantillon en deux paquets d'observations : un échantillon d'apprentissage $(X_i, Y_i)_{i=1, \dots, 50}$ et un échantillon de test $(X_i, Y_i)_{i=50, \dots, 100}$ (X_i, Y_i) pour ce dernier sous-échantillon (échantillon de test), nous supposons que les valeurs réponses sont inconnues et nous allons les approximer par :

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

Que nous estimons par :

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

Avec

$$\hat{f}^x(y) = \frac{\sum_{i=1}^n K((h^{-1}d(x, X_i))HK_0(g^{-1}(y - Y_i)))}{g \sum_{i=1}^n K(h^{-1}d(x, X_i))}.$$

3.3.1 Algorithmes

- Considérons un échantillon de $n = 100$ courbes X_i .
- Générer $\epsilon_i \rightarrow N(0, 1)$.
- Choisir la fonction de régression

$$r(X_i) = \frac{X_i^4}{4}$$

- K est un noyau gaussien

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right), u \in \mathbb{R}.$$

- H est un noyau quadratique

$$H(u) = \frac{15}{16}(1 - u^2)^2 \text{ si } u \in [-1, 1]$$

- On insère h et g (paramètres de lissage)

Les résultats

L'estimation du mode conditionnel, sous logiciel R se fait moyennant la fonction prédéfinie de R "**funopare.mode.lcv**" développée par F.Ferraty et P.Vieu[8]. disponible à l'adresse :

<http://www.Isp.upstlse.fr/staph/npfda>.

L'expression qui donne les valeurs estimés , les réponses estimés et l'erreur quadratique moyenne (MSE) est :

`funopare.mode.lcv(Y[50 : 100], X[50 : 100,], X[80,], q = 1, nKnot = c(1.100), range.grid = c(1 : 100), Knearest = Null, Kind.Kernel = "quadratic", semimetric = "deriv")`.

Avec

- (X, Y) : Échantillon de test.
- **Knearest** : Vecteur des paramètres de lissage, par défaut le vecteur Knearest contient une séquence de 10 entiers.
- **Kind.Kernel** : La fonction du noyau utilisée pour le calcul de l'estimateur du noyau (ici on a utilisé le noyau quadratique défini par : $K(u) = \frac{15}{16}(1 - u^2)^2$ si $u \in [-1, 1]$).
- **Semimetric** : La semi-métrique utilisée. défini par :

$$d_q^{deriv}(x_i, x_{i'}) = \sqrt{\int (x_i^{(q)}(t) - x_{i'}^{(q)}(t))^2 dt}.$$

Avec $x_i, x_{i'}$ deux courbes observées.

- **q** : L'ordre de dérivation.

La figure (3.3) ci-dessous représente les résultats obtenus lors de notre application :

```

$Estimated.values
 [1] 0.1132507309 0.1694357127 0.0683027455 0.0907767382 0.1581987164
 [6] 0.0795397419 0.0430195037 0.0907767382 0.0008807674 0.0008807674
 [11] 0.0795397419 0.0008807674 0.0121177638 0.0795397419 0.0430195037
 [16] 0.0205455110 0.1132507309 0.0205455110 0.1132507309 0.1132507309
 [21] 0.1132507309 0.1132507309 0.1132507309 0.0683027455 0.1919097054
 [26] 0.1132507309

$Predicted.values
 [1] 0.02054551

$Response.values
 [1] 0.2107457534 0.1744469788 0.1334271986 0.0952567830 0.0617790171
 [6] 0.0344187683 0.0191372246 0.0097395004 0.0036944190 0.0046686760
 [11] 0.0037026041 0.0062733793 0.0009786305 0.0063649307 0.0124477498
 [16] 0.0297617713 0.0558624092 0.0868737704 0.1286948351 0.1674015698
 [21] 0.2058772120 0.2303990310 0.2503698640 0.2533934537 0.2396320385
 [26] 0.2089118254

$Mse
 [1] 0.005469935

```

FIGURE 3.3 – Résultats de l'estimation

Avec

- **Valeurs estimées** : Vecteur contenant les réponses estimées pour chaque courbe.
- **Valeurs prédites** : La valeur prédit par le mode conditionnel.
- **Valeurs de réponse** : Les réponses observées correspondant au deuxième sous-échantillon d'apprentissage.
- **Mse** : Erreur quadratique moyenne entre les valeurs estimées et valeurs observées.

Donc la valeur estimer est :

$$\hat{y} = \hat{\theta} = \arg \sup_y \hat{f}^x(y) = 0.02054551.$$

Interprétation

A partir des résultats obtenus on remarque que l'erreur quadratique moyenne (MSE) est égale à 0.005469935 ce qui montre que les réponses estimées et les réponses observées sont très proches, et la bonne performance de notre méthode d'estimation.

Conclusion

Dans ce travail, nous avons présenté quelques notions de base liées à la statistique non paramétrique fonctionnelle (variables fonctionnelle, donnée fonctionnelle, semi-métrique et la probabilité des petites boules), Ensuite, nous avons fait une présentation générale de la méthode du noyau en citant quelques noyaux usuels. Ainsi que, quelques théorèmes et proposition concernant la convergence presque complète et les inégalités exponentielles.

Nous avons considéré une couple de variables aléatoire (X, Y) où Y est à valeurs dans \mathbb{R} et X est à valeurs dans un espace semi-métrique (E, d) et nous construisons des estimateurs par la méthode du noyau pour la fonction de densité conditionnelle, et le mode conditionnel, ensuite nous avons établi la convergence presque complète pour ces deux derniers modèles (la densité conditionnelle et le mode conditionnel) dans le cas indépendante et identiquement distribué. Ces résultats sont obtenus sous des conditions standards de statistique non paramétrique fonctionnelle.

En fin, nous avons validés nos résultats théoriques par une simulation sous logiciel R.

BIBLIOGRAPHIE

- [1] Alexandre. B. Tsy bakov. (2004). Introduction à l'estimation non paramétrique. Springer-Verlag.
- [2] Belabed. FZ. (2015). Estimation non paramétrique fonctionnelle de la fonction de hasard conditionnelle par la méthode des K plus proches voisins. Thèse de doctorat, Université Djillali Liabes. Sidi Bel Abbès.
- [3] Benchoulak. H.(2012). Bandes de confiance pour les fonction de densité et de régression. thèse du Magistère, Université constantine.
- [4] Bernard. D. (2013). Le livre de R : Apprentissage et référence. Vuibert.
- [5] Chiab. Y. (2013). Estimation de la fonction mode pour des données tronquées et censurées. Thèse de doctorat, Université Annaba
- [6] Collomb.G, W. Härdle and S. Hassani. (1987). A note on prediction via estimation of conditional mode function, J. of Statist. Plan. and Inf, 15, Pages 227-236.
- [7] Ferrani. Y. (2014). Sur l'estimation non paramétrique de la densité et du mode dans les modèle de données incomplète et associées. Thèse de doctorat. Université du Littoral Côte d'Opale
- [8] Ferraty, F and Vieu, P. (2006). Non parametric funtional data analysis. Theory and pratice. Springer-Verlag.
- [9] Fethi. M. (2006). Aspects théoriques pratiques dans l'estimation non paramétrique de la densité conditionnelle pour des données fonctionnelles. Thèse du doctorat, Université de Grenoble. France.

- [10] Gherair. Dj. (2017). Modélisation non paramétrique pour les variables aléatoires fonctionnelles cas de données indépendantes. Mémoire de Master, Université Kasdi Merbah. Ouargla.
- [11] Hsu.p.l and Robbins. herbert. (1947). Complete Convergence and the law of large numbers. Proceedings of the National Academy of Sciences of the United States of America,33,25.
- [12] Kaid. Z. (2012). Sur l'estiméation non paramétrique des modèles conditionls pour variables fonctionnelles spatialement dépendants. Thèse du doctorat, Université Djillali Liabes. Sidi Bel Abbès.
- [13] Kebabi. Kh. (2014). Etsimation non paramétrique de la fonction de régression :cas d'un modèle de censure mixte. Thèse du doctorat, Université Constantine.
- [14] Parzen. E. (1962) On the estimation of a probabily density function and mode. Annals of Mathematical Statistics,33 :1065-1027.
- [15] Preda. C. Méthodes pour l'analyse des données fonctionnelles. Université des Sciences et Technologies de Lille. France.
- [16] Rosenblatt. M. (1956). Remarks on some nonparametric estimator of a density function. Annals of Mathematical Statistics, 27 :832-837.
- [17] TADJ.A. (2011). Sur les modèles non paramétriques conditionnels en statistique fonctionnelle. Thèse du doctorat, Université Toulouse 3.
- [18] Sophie Dobo-Niang and Ali lakasaci. (2008). conditional mode regression :Application to functional time series prediction. Université Lille3. France,arXiv preprint arXiv :0812.4882.
- [19] Youndjé. E. (1993). Estimation non paramétrique de la densité conditionnelle par la méthode du noyau. PhD Thesis from the Rouen University (in French).
- [20] ZOUGAB. N.(2007).Etude comparative des méthodes de sélection du paramètre de lissage dans l'estimation de la densité de probabilité par la méthode du noyau. Thèse de Magistère. Université abderahman Mira.

Résumé

Le but de ce travail, est l'étude d'estimation non paramétrique fonctionnelle du mode conditionnel de la variable aléatoire réelle Y sachant la variable fonctionnelle X dans le cas i.i.d (indépendant identiquement distribuée). nous avons établi la convergence presque complète de l'estimateur non paramétrique de la fonction de densité conditionnelle et le mode conditionnel, et a fin de validé les résultats théoriques obtenus, une simulation a été réalisée a cet effet.

Mots clés

Méthode de noyau, densité conditionnelle, donnée fonctionnelles, convergence presque complète, mode conditionnel.

Abstract

The goal of this work is to study the functional nonparametric estimation of the conditional mode of functional real random variable Y given the functional random X in the i.i.d (identically distributed independent) case. we established the almost complete convergence of the nonparametric estimator for the conditional density function and the conditional mode, a simulation is given to illustrates the good behavior in practice of our methodology.

Keywords

Kernel method, conditional density, functional data, convergence almost complete, Conditional Mode.