



UNIVERSITE KASDI MERBAH
OUARGLA

Faculté des mathématiques et sciences de la
matière

N° d'ordre :
N° de série :

DEPARTEMENT DE MATHEMATIQUES

MASTER

Spécialité : Mathématiques

Option : Probabilité et Statistique

Par : Bellatrach Nadjet

Thème

La convergence presque complète d'estimateur non
paramétrique de densité conditionnelle pour les
variables fonctionnelles

Soutenu publiquement le : 02/06/2016

Devant le jury composé de :

Said ZIBAR	M.A.B université de KASDI Merbah - Ouargla	Président
Fatima MEDDI	M.C.A université de KASDI Merbah - Ouargla	Examineur
Rachid AGOUNE	M.C.A université de KASDI Merbah - Ouargla	Rapporteur

Dédication

*À la plus belle créature que Dieu a créée sur terre ,,
À cet source de tendresse, de patience et de générosité,,
À ma mère!
à mon père celui qui m'a implanté le courage et l'amour du travail.
À mes chers frères et sœurs
À ma chère et ma copine HAYAT qui a toujours était à mes cotés
À tous mes amis et collègues.*

Remerciements

Je voudrais d'abord et avant tout remercier ma encadreur vertueux Professeur AGOUNE RACHID, superviseur de cette thèse, pour l'aide qu'il a fournie sur l'année.

Je tiens à le remercier pour sa patience, et surtout pour sa confiance, et pour sa disponibilité et la précision de ses indications. Sans lui ce projet n'aurait pas été mené à terme !

Je tiens aussi à remercier ma professeur vertueuse *MEDDI FATIMA* pour me aider et de ses efforts depuis de nombreuses années, nous avons beaucoup profité de lui, en particulier dans cette mémoire ... tous les remerciements et sa gratitude à lui.

Avec un grand honneur, j'aimerais présenter mes remerciements et ma gratitude aux mon professeur *ZIBAR SAID* et *MEDDI FATIMA* pour avoir accepté d'évaluer ce travail et pour toutes leurs remarques et critiques.

A tous mes enseignants qui m'ont initié aux valeurs authentiques, en signe d'un profond respect et d'un profond amour !!! Merci à vous tous.

merci ma reconnaissance à monsieur.

Table des matières

1	Quelques rappels sur le modèle non paramétrique	1
1.1	Les variables aléatoires fonctionnelles	1
1.2	Ensembles de données fonctionnels	1
1.3	Modèle non paramétrique	2
1.3.1	Présentation du modèle non paramétrique	2
1.4	Pourquoi utiliser la méthode de noyau pour les données fonctionnelles ?	3
1.4.1	Cas réel	4
1.4.2	Cas multivarié	5
1.4.3	Cas fonctionnelle	7
1.5	Convergence des variables aléatoires	8
1.5.1	Convergence en probabilité :	8
1.5.2	Convergence presque sûre :	8
1.5.3	Convergence presque complète :	8
1.6	Méthode de noyau	9
1.6.1	Introduction :	9
1.6.2	Estimateur à noyau :	9
1.6.3	Quelques propriétés de l'estimateur à noyau :	11
1.6.4	Les conditions sur le noyau K :	11
1.6.5	Choix de paramètre de lissage h :	11
2	Etude asymptotique	12
2.1	Types de noyau	12
2.1.1	Noyau de type I :	12
2.1.2	Noyau de type II :	12

2.1.3	Noyau de type 0 :	13
2.2	Modélisation fonctionnelle non paramétrique pour l'étude de convergence	13
2.2.1	Quelque hypothèses nécessaires :	13
2.2.2	Type de continuité :	14
2.2.3	Type de Lipschitz :	15
2.2.4	Probabilités des petites boules :	15
2.3	Estimateurs du noyau	15
2.3.1	Hypothèses nécessaires :	15
2.3.2	Estimation non paramétrique de la régression	16
2.3.3	Estimation non paramétrique de répartition conditionnelle	16
2.3.4	Estimation non paramétrique de la densité conditionnelle	17
2.4	La convergence presque complète de la densité conditionnelle	18
3	Application et exploration des observations sur des données simulées	25
3.1	Simulée les données fonctionnelles	25
3.1.1	Programme sous R	25
3.2	Résultats de simulation	26
3.2.1	Résultat un variable aléatoire fonctionnelle :	27
3.2.2	Résultat de plusieurs variables aléatoires fonctionnelles :	28
3.3	Simulation les variables de réponse réels	28
3.3.1	Algorithme 1	29
3.3.2	Programme sous R	29
3.4	Simulation l'estimateur de densité conditionnelle	30
3.4.1	Algorithme 2	31
3.4.2	Programme sous R	31
3.5	Comparation	32
3.5.1	Programme sous R	33
3.6	Résultats des simulations	33
3.6.1	Simulation de l'estimateur de densité conditionnelle	33
3.6.2	Simulation de la fonction p	34
3.6.3	La fonction p et l'estimateur f au même plan :	35
3.6.4	Commentaires	36

Table des figures

1	Les données fonctionnelles	x
1.1	Formes des noyaux symétrique	5
1.2	Formes des noyaux	6
3.1	Courbe de un v.a.f	27
3.2	Courbes de plusieurs v.a.f	28
3.3	Comportement graphique de l'estimateur de densité conditionnelle \hat{f}_Y^x .	34
3.4	Courbe de la fonction p	35
3.5	Courbe de la fonction p et l'estimateur de f	36

Abréviations et Notations

- *p.co* : presque complet (convergence) pour la séquence de r.r.v.
- $B(\chi, h)$: boule ouverte du centre χ et rayon h , dans l'espace (E, d)
- *f.d.c.* : fonction de distribution cumulative
- C où C' : constantes réelles positives et finies
- $\mathbb{E}(Y)$: l'espérance de certains v.a.r. Y
- $\mathbb{E}(Y|\mathcal{X} = \chi)$: l'espérance conditionnelles de certains v.a.r. Y étant donné le v.a.f. \mathcal{X}
- v.a.f. : variable aléatoire fonctionnelle
- f : densité marginale du v.a.f.
- $F_Y^{\mathcal{X}}(\chi, y)$ où $F_Y^{\mathcal{X}}(y)$: répartition conditionnelle de certains v.a.r. Y étant donné le v.a.f. \mathcal{X}
- $f_Y^{\mathcal{X}}(\chi, y)$ où $f_Y^{\mathcal{X}}(y)$: densité conditionnelle de certains v.a.r. Y étant donné le v.a.f. \mathcal{X}
- $\phi_{\chi}(h)$: mesure de $B(\chi, h)$ par rapport à la loi de probabilité de χ
- h : notation générique pour le paramètre de lissage $h = h(n)$
- H : notation générique pour une fonction de noyau intégrée
- K : notation générique pour une fonction asymétrique du noyau
- K_0 : notation générique pour une fonction de noyau symétrique standard
- \mathbb{N} : ensemble d'entiers positifs
- $O_{p.co}$: taux de convergence presque complète
- $O_{p.s}$: taux de convergence presque sûre
- (Ω, \mathcal{A}, P) : espace de probabilité sur lequel tous les r.v. sont définis
- r : opérateur de régression non linéaire
- S : sous-ensemble compact de \mathbb{R}

- \mathcal{S} : l'échantillon fonctionnel statistique $\mathcal{S} = \mathcal{X}_1, \dots, \mathcal{X}_n$
- X : variable réelle aléatoire
- x : nombre réel
- \mathbf{X} : Variable aléatoire multivariée générique
- \mathbf{x} : vecteur multivarié générique (non aléatoire)
- $\mathbf{X}_i, i = 1, \dots, n$: échantillon de vecteurs aléatoires
- $\mathbf{x}_i, i = 1, \dots, n$: observations statistiques des vecteurs aléatoires \mathbf{X}_i
- \mathcal{X} : variable aléatoire fonctionnelle
- χ : élément fonctionnel (non réel)
- $\mathcal{X}_i, i = 1, \dots, n$: échantillon de v.a.f.
- $\chi_i, i = 1, \dots, n$ observations statistiques du v.a.f. \mathcal{X}
- *v.a.r* : variable aléatoire réel
- \mathbb{R} : ensemble de nombres réels
- *v.a* : variable aléatoire

Introduction

Au cours des dernières années, nous avons pu observer dans la recherche scientifique, plusieurs développements dans plusieurs domaines. Il existe de nombreuses situations dans lesquelles on peut étudier le lien entre deux variables, dans le but principal de pouvoir prédire de nouvelles valeurs de l'une, d'entre elles compte tenu de l'autre.

Étudier les liens entre deux variables aléatoires est une question très importante en statistique. D'un point de vue historique, ce problème a été abordé pour la première fois dans un contexte géométrique par Galileo Galilei en (1632). L'idée principale de ce dernier est d'ajuster un nuage de point par une droite permettant d'interpréter la relation entre des données contaminées. Une formulation mathématique pour ce problème, connue dans la littérature sous le nom de régression linéaire, a été donnée par Legendre et Gauss indépendamment, en (1805) et (1809), et est basée sur le principe des moindres carrés. En statistique, ce problème peut être modélisé de la manière suivante : supposons qu'on dispose de deux variables aléatoire dépendantes X et Y , la prévision de Y sachant X se fait à travers X par une application r . Autrement dit, on cherche une fonction r telle que $r(X)$ soit une bonne approximation de Y selon un critère donné. Ainsi, le problème devient la minimisation de la fonction de risque.

Dans notre contexte non paramétrique, les premiers résultats ont été obtenus par Tukey (1961). Tandis que l'estimation par la méthode du noyau a été utilisée pour la première fois en 1964 séparément par Nadaraya et Watson. Cette méthode d'estimation a connu un développement continu. En effet, Devroye (1978) a établi la convergence uniforme presque sûre de cet estimateur. Le taux de convergence optimal pour la régression non paramétrique a été donné par Stone (1980, 1982). Collomb (1981, 1983, 1984, 1985) apporte une contribution déterminante sur ce modèle. Ces

travaux se sont focalisés sur l'utilisation de la régression dans la prévision de séries chronologiques.

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.

Notre souhait est d'étudier ce problème lorsque la variable explicative est fonctionnelle et que la réponse est encore réelle. À la fois pour corriger les idées et pour souligner le grand intérêt et l'utilité de ce problème dans de nombreux domaines des sciences appliquées, par exemple nous avons un échantillon statistique (de Taille n) est composé de courbes spectrométriques χ_1, \dots, χ_n (ce sont les données fonctionnelles comme dans la figure suivante) correspondant aux spectres observés pour n pièces de viande hachée finement.

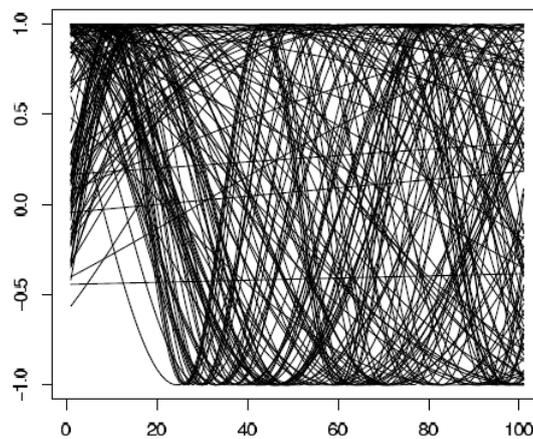


FIGURE 1 – Les données fonctionnelles

En outre, par un procédé chimique analytique, nous avons mesuré la teneur en matière grasse de chaque pièce y_1, \dots, y_n (ce sont les réponses scalaires). Ainsi, nous

recueillons les observations d'une réponse scalaire (le contenu en graisse) et d'une variable fonctionnelle explicative (spectres). Une question est : compte tenu d'un spectre observé d'un morceau de viande, pouvons-nous prédire sa teneur en matières grasses correspondante ? C'est généralement un problème de prédiction fonctionnelle.

La base de ce mémoire est l'estimation de la fonction de densité conditionnelle, en statistique fonctionnelle, a été que introduite par Ferraty et al. (2006) [1]. Ces auteurs ont obtenu la convergence presque complète dans le cas i.i.d. Depuis cet article, une littérature abondante s'est développée sur l'estimation de la densité conditionnelle. Et nous proposons d'étudier la convergence uniforme d'un estimateur à noyau pour la densité d'une variable réelle conditionnée par une variable aléatoire fonctionnelle dans le cas i.i.d des variables aléatoires.

Dans le but de présenter les travaux que nous avons réalisés durant la réalisation de ce mémoire, celle-ci est organisé comme suit :

Le premier chapitre, l'objectif de ce chapitre est d'introduire la notion de variables aléatoires fonctionnelle et les données fonctionnelles ainsi que la statistique non paramétrique, et aussi sur la notion de convergence presque complète, cette concept a été introduit en (1947) par Hsu et Robbins [7], qui ont prouvé que la séquence des moyens arithmétiques de i.i.d. des variables aléatoires convergent complètement vers la valeur attendue des variables. et nous fournissons la méthode de noyau (ou méthode de Parzen-Rozenblatt) et les conditions nécessaires sur le noyau et paramètre de lissage. au travers de laquelle nous allons les étudier.

Dans le deuxième chapitre nous présentons l'étude asymptotique de l'estimateur non paramétrique de la densité conditionnelle par méthode de noyau, qui présentée par Ferraty et vieu [1]. Nous définissons rapidement les notions de types de noyau (type 0, I, II), puis nous présentons ensuite certaines hypothèses qui nous aident dans cette étude. Le résultat est présenté en terme de convergence presque complète. Ce genre de convergence stochastique peut sembler assez inhabituel pour certaines personnes, mais il a été sélectionné car il présente deux avantages importants. Tout d'abord, il est plus fort que la convergence presque sûre et la convergence en probabilité, et deuxièmement, il est plus facile de prouver que la cohérence presque sûre elle-même. Et

pour ce genre d'asymptotique être très familier.

Notre principal contribution est proposé au troisième chapitre, ce chapitre est consacré à les applications et les commentaires, où nous avons confirmée que les variables aléatoires fonctionnelle sont des courbes (sous logiciel R), nous illustrons la convergence presque complète en effectuant des simulations sur l'estimateur de densité conditionnelle par la méthode de noyau pour les variables fonctionnelles (qui présenté par ferraty), et sur autre densité, ensuite nous examinons les résultats. Nous allons comparer les résultats, cet application montre que la convergence presque complète de l'estimateur de densité conditionnelle est confirmée.

Chapitre 1

Quelques rappels sur le modèle non paramétrique

1.1 Les variables aléatoires fonctionnelles

Une variable aléatoire \mathcal{X} est appelée variable fonctionnelle (v.f.) si elle prend des valeurs dans un espace de dimension infini (ou espace fonctionnel).

Une observation χ de \mathcal{X} s'appelle une donnée fonctionnelle.

Notez que, lorsque \mathcal{X} (resp. χ) désigne une courbe aléatoire (ou son observation), nous impliquons implicitement l'identification suivante $\mathcal{X} = \{\mathcal{X}(t); t \in T\}$ (resp. $\chi = \{\chi(t); t \in T\}$).

Dans cette situation, la caractéristique fonctionnelle provient directement des observations.

La situation où la variable est une courbe est associée à un ensemble unidimensionnel $T \subset \mathbb{R}$. Ici, il est important de remarquer que la notion de variable fonctionnelle couvre une plus grande surface que l'analyse des courbes.

1.2 Ensembles de données fonctionnels

Un ensemble de données fonctionnelles χ_1, \dots, χ_n est l'observation de n variables fonctionnelles $\mathcal{X}_1, \dots, \mathcal{X}_n$ est identique à \mathcal{X} .

Cette définition couvre de nombreuses situations, le plus populaire étant les ensembles

de données courbes.

Ici, nous supposons que nous disposons d'un échantillon de données fonctionnelles.

1.3 Modèle non paramétrique

Définition 1.1 Soit \mathbf{X} un vecteur aléatoire évaluée en \mathbb{R}^p et que ϕ est une fonction définie sur \mathbb{R}^p . Un modèle pour l'estimation de ϕ consiste à introduisant une contrainte de la forme

$$\phi \in \mathcal{C}.$$

Le modèle est appelé un modèle paramétrique pour l'estimation de ϕ si \mathcal{C} est indexée par un nombre fini d'éléments de \mathbb{R} .

Dans le cas contraire, le modèle est appelé un modèle non paramétrique.

Cette définition peut être facilement étendue au cadre fonctionnel.

Définition 1.2 Soit \mathcal{Z} une variable aléatoire évaluée dans un espace de dimension infini F et soit ϕ une application défini sur F et selon la distribution de \mathcal{Z} . Un modèle pour l'estimation de ϕ consiste à introduire une certaine contrainte de la forme

$$\phi \in \mathcal{C}.$$

Le modèle s'appelle un modèle paramétrique fonctionnel pour l'estimation de ϕ si \mathcal{C} est indexé par un nombre fini d'éléments de F . Dans le cas contraire, le modèle s'appelle un modèle fonctionnel non paramétrique.

1.3.1 Présentation du modèle non paramétrique

Soit Y est une variable aléatoire réelle i.i.d et \mathcal{X} est une variable aléatoire fonctionnelle. Soit $(\mathcal{X}_i, Y_i)_{i=1, \dots, n}$ n couples indépendantes, identiquement distribuées (i.i.d) comme (\mathcal{X}, Y) , et évaluées dans $E \times \mathbb{R}$, où (E, d) est un espace semi-métrique (c'est-à-dire \mathcal{X} est un v.a.f. et d un semi-métrique), et Soit χ (resp. Y) un élément fixe de E (resp. \mathbb{R}).

On s'intéresse à un modèle de régression non-paramétrique de Y en \mathcal{X} , Le modèle utilisé est de régression, sous forme :

$$Y = r(\mathcal{X}) + \varepsilon \quad (1.1)$$

où ε est le résidu (v.a.r. centré réduite).

On obtient, sous certaines conditions de régularité (essentiellement d'intégrabilité de Y) : $r(\mathcal{X}) = \mathbb{E}(Y/\mathcal{X})$, C'est-à-dire la régression de Y en \mathcal{X} .

Nous proposons de prédire la réponse scalaire Y du prédicteur fonctionnel \mathcal{X} en utilisant diverses méthodes basées sur la distribution conditionnelle de Y donné \mathcal{X} . L'opérateur de régression (non linéaire) r de Y sur \mathcal{X} est défini par :

$$r(\chi) = \mathbb{E}(Y/\mathcal{X} = \chi) \quad (1.2)$$

et la fonction de distribution cumulative conditionnelle (d.c.f.) de Y donnée \mathcal{X} est définie par :

$$\forall y \in \mathbb{R}, F_Y^{\mathcal{X}}(\chi, y) = P(Y \leq y | \mathcal{X} = \chi). \quad (1.3)$$

En outre, si la probabilité de Y sachant \mathcal{X} est absolument continue par rapport à la mesure de Lebesgue, nous notons $f_Y^{\mathcal{X}}(\chi, y)$ la valeur de la fonction de densité correspondante à (χ, y) .

Notez que sous une hypothèse de différentiabilité sur $F_Y^{\mathcal{X}}(\chi, \cdot)$, Cette densité conditionnelle fonctionnelle peut être écrite comme

$$\forall y \in \mathbb{R}, f_Y^{\mathcal{X}}(\chi, y) = \frac{\partial}{\partial y} F_Y^{\mathcal{X}}(\chi, y). \quad (1.4)$$

1.4 Pourquoi utiliser la méthode de noyau pour les données fonctionnelles ?

Les méthodes du noyau sont utilisées de manière intensive par la communauté des non-paramétrés car elles constituent un moyen utile de faire une pondération locale. Nous commençons par rappeler quelle est la pondération locale du noyau dans les cas réels et multivariés avant de l'étendre au contexte fonctionnel.

1.4.1 Cas réel

La pondération locale du noyau est basée sur une fonction de noyau (K) et une bande passante (notée h). Si x est un nombre réel fixe, la pondération locale du noyau transforme n v.a.r. X_1, X_2, \dots, X_n à $\Delta_1, \Delta_2, \dots, \Delta_n$ tel que :

$$\Delta_i = \Delta_i(x, h, K) = \frac{1}{h} K\left(\frac{x - X_i}{h}\right).$$

L'idée principale de la pondération locale autour de x est d'attribuer à chaque v.a.r. X_i un poids prenant en compte la distance entre x et X_i , plus X_i est éloigné de x , plus la pondération est petite.

Laissez-vous rappeler ce qu'est une fonction de noyau exactement dans cette situation la plus simple. Il existe une grande variété de noyaux, toute fonction de densité peut être considérée comme un noyau.

Pour simplifier notre but, nous considérons à ce stade que les noyaux positifs et symétriques les plus classiques.

La figure (1.1) affiche diverses fonctions du noyau qui sont définies analytiquement comme suit :

1. Noyau rectangulaire : $K(u) = \frac{1}{2}1_{[-1,1]}(u)$
2. Noyau triangulaire : $K(u) = (1 - |u|)1_{[-1,1]}(u)$
3. Noyau d'Epanechnikov : $K(u) = \frac{3}{4}(1 - u^2)1_{[-1,1]}(u)$
4. Noyau gaussien : $K(u) = \frac{1}{\sqrt{2\pi}} \exp^{-u^2/2}$

Pour préciser la notion de pondération locale du noyau, considérons le noyau rectangulaire et réécrivez les Δ_i comme suit :

$$\Delta_i = \frac{1}{h} 1_{[x-h, x+h]}(X_i)$$

Dans cette situation, la caractéristique locale de la pondération semble évidente puisque le v.a.r. En dehors de la plage $[x - h, x + h]$ sont ignorés. En outre, la normalisation $1/h$ est proportionnelle à la taille de l'ensemble $[x - h, x + h]$ sur lequel les X_i sont pris en compte. Ces points ne sont pas seulement vrais pour le noyau rectangulaire, mais sont partagés par tous les noyaux compacts.

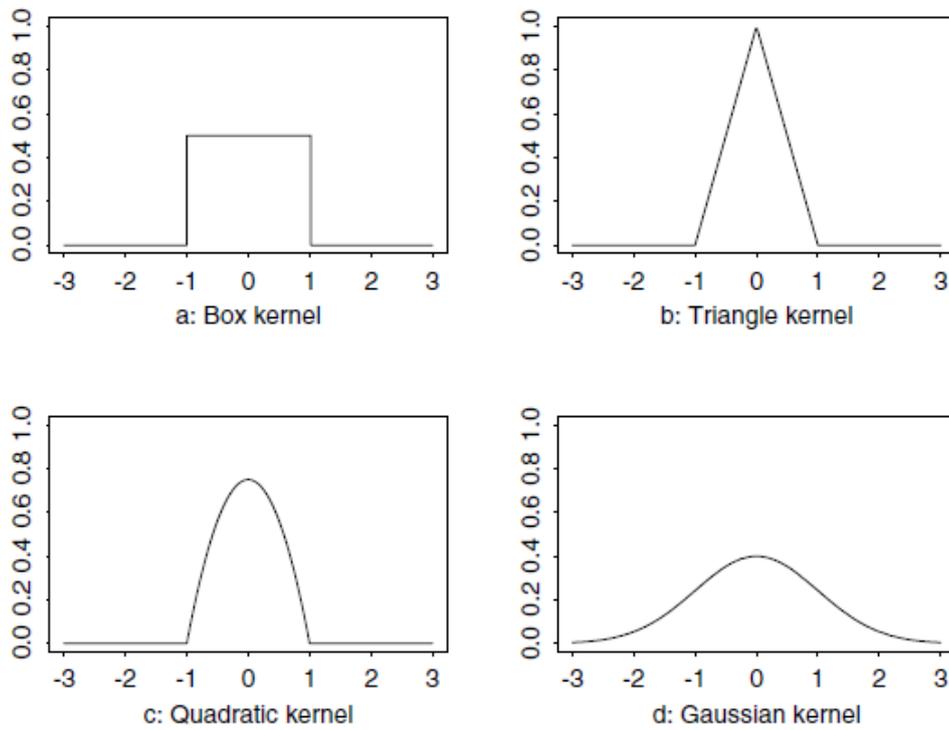


FIGURE 1.1 – Formes des noyaux symétrique

1.4.2 Cas multivarié

Dans des situations multivariées, on observe n vecteurs aléatoires $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ évalué en \mathbb{R}^p . La pondération locale précédente du noyau peut être facilement étendue à cette situation. À cette fin, il suffit de considérer un noyau multivarié K^* qui sera une fonction de \mathbb{R}^p dans \mathbb{R} . Le premier moyen (naturel) de le faire est de définir K^* comme un produit de p fonctions de noyau réel K_1, \dots, K_p :

$$\forall u = {}^t(u_1, \dots, u_p) \in \mathbb{R}^p, K^*(u) = K_1(u_1) \times K_2(u_2) \times \dots \times K_p(u_p).$$

Une deuxième façon consiste à combiner une fonction de noyau réelle H avec une norme (désignée par $\|\cdot\|$) dans \mathbb{R}^p comme suit :

$$\forall u \in \mathbb{R}^p, K^*(u) = K(\|u\|).$$

Notez que si $K_1 = K_2 = \dots = K_p = 1_{[-1,1]}$ et si $\|\cdot\|$ est la norme suprême, les deux approches coïncident en prenant $K = 1_{[-1,1]}$. En outre, parce que $\|u\|$ est toujours une quantité positive, le noyau réel K devrait avoir un support positif (i.e., $\{v \in \mathbb{R} \text{ tel que } K(v) > 0\} \subset \mathbb{R}^p$.) Cela conduit à utiliser des fonctions asymétriques pour le noyau K . Les exemples rapportés dans la figure (2.2) sont les versions asymétriques de celles illustrées à la figure de noyaux symétriques usuels.

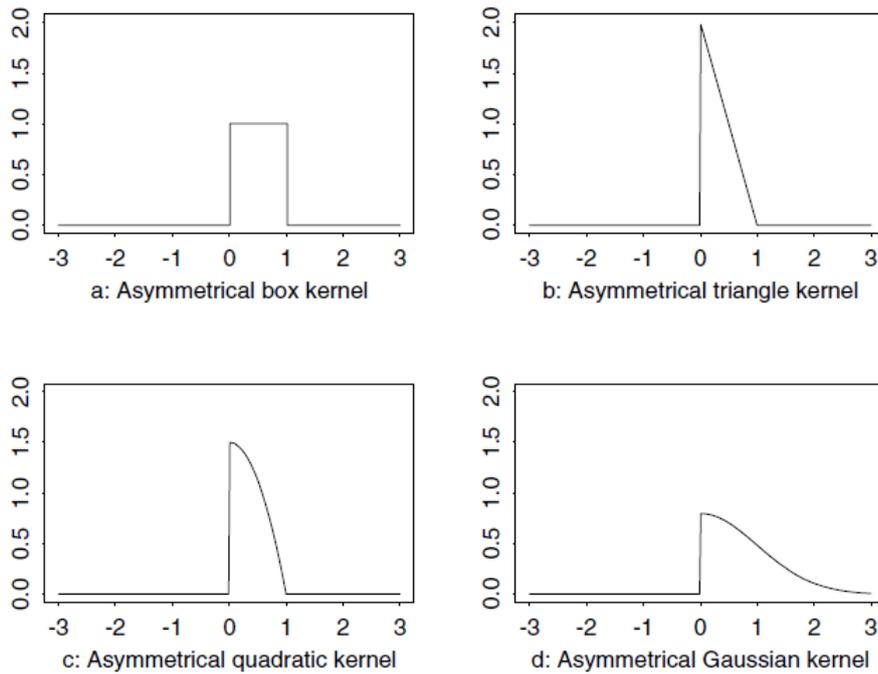


FIGURE 1.2 – Formes des noyaux

Maintenant, nous discuterons de la façon dont cela peut être interprété en termes de pondération locale. Soit x un vecteur fixe de \mathbb{R}^p . La pondération locale du noyau multivarié consiste à transformer les n vecteurs aléatoires $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ à n variables $\Delta_1, \Delta_2, \dots, \Delta_n$ tel que :

$$\Delta_i = \frac{1}{h^p} K^* \left(\frac{x - \mathbf{X}_i}{h} \right).$$

Si l'on considère les noyaux compatibles de manière compacte, il apparaît clairement que les Δ_i sont des transformations pondérées localement des variables \mathbf{X}_i , puisque

$\Delta_i = 0$ tant que le \mathbf{X}_i correspondant est hors de certains voisins de x .

1.4.3 Cas fonctionnelle

Soit $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$ n v.a.f. évalué dans E et soit χ un élément fixe de E . une extension fonctionnelle des idées de pondération locale du noyau multivarié serait transforme le n v.a.f. $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$, les n quantités

$$\frac{1}{V(h)} K \left(\frac{d(\chi, \mathcal{X}_i)}{h} \right),$$

Où d est semi-métrique sur E , K est un noyau réel (asymétrique). Dans cette expression $V(h)$ serait le volume de

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

Qui est la boule, par rapport à la topologie induite par d , centrée à χ et au rayon h . Cependant, cette approche naïve demande de définir $V(h)$.

il faut avoir une mesure sur E . C'est la principale différence avec les cas réels et multivariés pour lesquels la mesure de Lebesgue est implicitement utilisée alors que dans l'espace fonctionnel E .

Par conséquent, afin de se libérer d'un choix d'une mesure particulière, nous construisons la normalisation en utilisant directement la distribution de probabilité de la v.a.f. les variables pondérées locales du noyau fonctionnel sont définies par :

$$\Delta_i = \frac{K \left(\frac{d(\chi, \mathcal{X}_i)}{h} \right)}{\mathbb{E} \left(K \left(\frac{d(\chi, \mathcal{X}_i)}{h} \right) \right)} \quad (1.5)$$

Si nous retournons au cas multivarié, nous avons, pour un C constant selon K et sur la norme $\|\cdot\|$ Utilisé dans R_p ,

$$\mathbb{E}K(\|x - \mathbf{X}_i\|/h) \sim Cf(x)h^p \quad (1.6)$$

Tant que \mathbf{X}_i a une densité par rapport à la mesure de Lebesgue qui est continue et tel que $f(x) > 0$ (ce type de résultat est connu dans la littérature comme théorème

de type de Bochner ([6] donne une large portée sur tel résultats). Ainsi, il est clair maintenant que (1.5) est une extension de la pondération locale du noyau multivarié dans le cadre fonctionnel.

1.5 Convergence des variables aléatoires

1.5.1 Convergence en probabilité :

Définition 1.3 Soit $(X_n)_n$ une suite de variables aléatoires réelles définies sur un même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que X_n converge vers X en probabilité si $\forall \varepsilon > 0 \lim_{n \rightarrow \infty} (\mathbb{P}(|X_n - X| \geq \varepsilon)) = 0$.

On note parfois : $X_n \xrightarrow{p} X$

1.5.2 Convergence presque sûre :

Définition 1.4 On dit que X_n converge presque sûrement vers X si :

$$\mathbb{P}(\lim_{n \rightarrow \infty} (X_n = X)) = 1.$$

Ou de manière équivalente, s'il existe un sous-ensemble \mathbb{P} -négligeable N dans ω tel que : $\forall \omega \in \Omega/N, (X_n)_\omega \xrightarrow{p.s} X_\omega$

On parle aussi de convergence presque partout ou avec probabilité 1 ou forte, et on écrit : $X_n \xrightarrow{p.s} X$

1.5.3 Convergence presque complète :

Le concept de convergence presque complète a été introduit par Hsu et Robbins (1947). Elle implique la convergence presque sûre et se prête bien aux calculs faisant intervenir des sommes de variables aléatoires. Malgré cela, elle ne commence à devenir populaire dans la communauté statistique que dans les années 1980 après les travaux de Collomb. Elle est utilisée surtout en statistique non-paramétrique.

On dit que la suite de variables aléatoires réelles $(X_n)_{n \in \mathbb{N}}$ converge presque complètement vers une variable aléatoire X lorsque $n \rightarrow \infty$ si :

$$\forall \varepsilon > 0, \sum_{n \in \mathbb{N}} \mathbb{P}[|X_n - X| > \varepsilon] < \infty.$$

et on dit que la vitesse de convergence presque complète de la suite de variables aléatoires réelles $(X_n)_{n \in \mathbb{N}}$ vers X est d'ordre (u_n) si :

$$\exists \varepsilon_0 > 0, \sum_{n \in \mathbb{N}} \mathbb{P}[|X_n - X| > \varepsilon_0 u_n] < \infty.$$

Cette définition du taux a été introduite par Ferraty et Vieu (2006). Elle a l'avantage théorique d'impliquer les deux vitesses de convergence classiques en probabilité et presque sûre, et l'avantage pratique d'être souvent plus facile à démontrer.

Dans les quinze dernières années, ce mode de convergence a été très utilisé dans des travaux concernant la statistique non-paramétrique des données fonctionnelles.

Remarque 1.1 *La convergence presque complète est plus forte que la convergence presque sûre et la convergence en probabilité.*

1.6 Méthode de noyau

1.6.1 Introduction :

L'estimation par la méthode noyau (ou méthode de Parzen-Rozenblatt) est une méthode non paramétrique d'estimation de la densité d'une variable aléatoire. Cette méthode permet d'obtenir une densité continue et constitue en ce sens une généralisation de la méthode de l'histogramme. En effet, la fonction indicatrice utilisée pour l'histogramme est ici remplacée par une fonction continue (le noyau) et une somme de fonctions continues reste continue.

1.6.2 Estimateur à noyau :

On va s'intéresser au modèle de régression, où

$$Y_i = f(x_i) + \varepsilon_i, \quad i = 1, \dots, n \tag{1.7}$$

Ici, x est un v.a et les x_i sont n -échantillons de x , les ε_i sont i.i.d., et on cherche à estimer f . Noter aussi que les Y_i sont des variables aléatoires indépendantes mais pas

identiquement distribuées.

On suppose que f est continue, telle que :

$$\forall t \in [0, 1] \quad F(t) = \int_t f(x)dx.$$

On peut estimer la fonction de répartition F de manière empirique par \hat{F} donnée par :

$$\forall t \in [0, 1] \quad \hat{F}(t) = \frac{1}{n} \sum_{i=1}^n Y_i 1_{x_i \leq t},$$

puis par \hat{f} un taux d'accroissement

$$\hat{f}(x) = (\hat{F}(x+h) - \hat{F}(x-h))/2h.$$

On a alors

$$\hat{f}(x) = \frac{1}{2nh} \sum_{i=1}^n Y_i 1_{x-h < x_i \leq x+h}.$$

$\hat{f}(x)$ est encore la moyenne locale des Y_i mais sur une fenêtre "glissante" centrée sur x .

Si on pose $K(x) = \frac{1}{2} 1_{-1 < x \leq 1}$ on a :

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n Y_i \frac{1}{h} K\left(\frac{x_i - x}{h}\right). \quad (1.8)$$

Remarquons que :

$$\sum_{i=1}^n \frac{1}{h} K\left(\frac{x_i - x}{h}\right) = \frac{1}{2} \sum_{i=1}^n 1_{x-h < x_i \leq x+h} \approx nh$$

et on peut choisir

$$\hat{f}(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} \quad (1.9)$$

Le paramètre $h > 0$ est appelé fenêtre, c'est un paramètre de lissage.

Dans les deux cas, il s'agit d'un estimateur linéaire tel que $\hat{f}(x) = \sum_{i=1}^n W_{n_i, x} Y_i$ avec $W_{n_i, x} = \frac{1}{nh} K\left(\frac{x_i - x}{h}\right)$ si on choisit (1.2), ou $W_{n_i, x} = K\left(\frac{x_i - x}{h}\right) / \sum_{k=1}^n K\left(\frac{x_k - x}{h}\right)$ si on choisit (1.3).

Pour généraliser la "moyenne locale avec fenêtre glissante", on peut choisir d'autres types de noyaux, i.e. de fonctions K , que l'on choisit souvent positives, à support compact, paires, d'intégrale 1.

1.6.3 Quelques propriétés de l'estimateur à noyau :

Si k est une densité de probabilité, alors \hat{f} est aussi une densité de probabilité. \hat{f} a les mêmes propriétés de continuité et de différentiabilité que k :

-Si k est continue, sera une fonction continue.

-Si k est différentiable, \hat{f} sera une fonction différentiable.

-Si k peut prendre des valeurs négatives, alors \hat{f} pourra aussi prendre des valeurs négatives.

1.6.4 Les conditions sur le noyau K :

On suppose que K vérifie les 4 conditions suivantes :

1. $\int_{\mathbb{R}} K = 1$
2. K est une fonction paire ou, plus généralement, $\int_{\mathbb{R}} uK(u)du = 0$
3. $\int_{\mathbb{R}} u^2K(u) < \infty$
4. $\int_{\mathbb{R}} K(u^2) < \infty$.

1.6.5 Choix de paramètre de lissage h :

Le choix de h est une étape importante lors de l'estimation par noyau dans le sens où h optimal nécessite de connaître la densité que l'on cherche à estimer, ce qui n'est en pratique pas le cas. Par ailleurs, si h est trop petit, le biais de l'estimateur devient petit devant sa variance et l'estimateur trop fluctuant.

On obtient un phénomène de sous-lissage. Dans le cas contraire, lorsque h est trop grand, le biais prend l'ascendant sur la variance et l'estimateur varie peu : on obtient un phénomène de sur-lissage.

En pratique, on utilise souvent la méthode de la validation croisée pour choisir automatiquement h .

Chapitre 2

Etude asymptotique

Nous rappelons, dans ce paragraphe et dans un premier temps quelques notations et hypothèses qui paraissent importantes pour la suite de ce travail de thèse. Ensuite, les résultats obtenus par Ferraty et al. (2006) sur l'estimation de la densité conditionnelle.

2.1 Types de noyau

2.1.1 Noyau de type I :

Une fonction K de \mathbb{R} dans \mathbb{R}^+ est appelé un noyau de type **I**, s'il existe deux constantes réelles $0 < C_1 < C_2 < \infty$, tel que :

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

2.1.2 Noyau de type II :

Une fonction K de \mathbb{R} dans \mathbb{R}^+ est appelé un noyau de type **II**, Si son support est $[0, 1]$ et si sa dérivée K' existe sur $[0, 1]$ et vérifie pour deux constantes réelles $-\infty < C_1 < C_2 < 0$:

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

2.1.3 Noyau de type 0 :

Une fonction K de \mathbb{R} dans \mathbb{R}^+ est appelé un noyau de type **0**, si $\int K(u)du = 1$ avec support compact $[-1, 1]$ et telle que :

$$\forall u \in [0, 1], K(u) > 0.$$

2.2 Modélisation fonctionnelle non paramétrique pour l'étude de convergence

2.2.1 Quelques hypothèses nécessaires :

Considérons le couple de variables aléatoire (\mathcal{X}, Y) où Y est à valeurs dans \mathbb{R} et \mathcal{X} est à valeurs dans \mathbb{R}^P .

On a l'opérateur de régression r (qui est un opérateur non linéaire de E dans \mathbb{R}) et nous considérerons les modèles suivants :

$$r \in C_E^0, \tag{2.1}$$

où :

$$C_E^0 = \left\{ f : E \longrightarrow \mathbb{R}, \lim_{d(\mathcal{X}, \mathcal{X}') \rightarrow 0} f(\mathcal{X}') = f(\mathcal{X}) \right\} \tag{2.2}$$

et $\exists \beta > 0$ tel que

$$r \in Lip_{E, \beta} \tag{2.3}$$

où

$$Lip_{E, \beta} = \{ f : E \longrightarrow \mathbb{R}, \exists C \in \mathbb{R}_*^+, \forall \mathcal{X}' \in E, |f(\mathcal{X}) - f(\mathcal{X}')| < Cd(\mathcal{X}, \mathcal{X}')^\beta \}.$$

On a l'opérateur $F_Y^{\mathcal{X}}$ (qui est un opérateur non linéaire de $E \times \mathbb{R}$ dans \mathbb{R}) peut être défini comme suit :

$$F_Y^{\mathcal{X}} \in C_{E \times \mathbb{R}}^0 \cap \mathcal{S}_{cdf}^{\mathcal{X}} \tag{2.4}$$

lorsque $\mathcal{S}_{cdf}^{\mathcal{X}}$ est :

$$\mathcal{S}_{cdf}^{\mathcal{X}} = \{ f : E \times \mathbb{R} \longrightarrow \mathbb{R}, f(\mathcal{X}, \cdot), \text{ est un c.d.f strictement croissant} \}.$$

et

$$C_{E \times \mathbb{R}}^0 = \begin{cases} f : E \times \mathbb{R} \longrightarrow \mathbb{R}, \forall \chi' \in \mathcal{N}_\chi \\ \lim_{d(\chi, \chi') \rightarrow 0} f(\chi', y) = f(\chi, y) \\ \text{et } \forall y' \in \mathbb{R}, \lim_{|y - y'| \rightarrow 0} f(\chi, y') = f(\chi, y) \end{cases}$$

et on a :

$$Lip_{E \times \mathbb{R}, \beta} = \begin{cases} f : E \times \mathbb{R} \longrightarrow \mathbb{R}, \\ \forall (\chi_1, \chi_2) \in \mathcal{N}_\chi^2, \forall (y_1, y_2) \in S^2, \\ |f(\chi_1, y_1) - f(\chi_2, y_2)| \leq C(d(\chi_1, \chi_2)^\beta + |y_1 - y_2|^\beta) \end{cases}$$

Où, il existe $\beta > 0$ tel que :

$$F_Y^\chi \in Lip_{E \times \mathbb{R}, \beta} \cap \mathcal{S}_{cdf}^\chi \quad (2.5)$$

On a l'opérateur f_Y^χ (qui est un opérateur non linéaire de $E \times \mathbb{R}$ dans \mathbb{R})

Rappelons que S est un compact fixe $S \in \mathbb{R}$. Nous commençons par introduire l'ensemble suivant de contraintes :

$$\mathcal{S}_{dens}^\chi = \begin{cases} f : E \times \mathbb{R} \longrightarrow \mathbb{R}, \\ \exists \xi > 0, \exists ! y_0 \in S, f(\chi, \cdot) \text{ est stricte croissante sur } (y_0 - \xi, y_0) \\ \text{et stricte décroissante sur } (y_0, y_0 + \xi) \end{cases}$$

alors on a :

$$f_Y^\chi \in C_{E \times \mathbb{R}}^0 \cap \mathcal{S}_{dens}^\chi \quad (2.6)$$

où, il existe $\beta > 0$ tel que :

$$f_Y^\chi \in Lip_{E \times \mathbb{R}, \beta} \cap \mathcal{S}_{dens}^\chi \quad (2.7)$$

2.2.2 Type de continuité :

Définition 2.1 *Les modèles fonctionnels non paramétriques (2.1) et (2.4) et (2.6) seront appelés modèles non paramétriques fonctionnels de type de continuité, parce que la propriété de continuité est leur caractéristique fonctionnelle commune principale.*

2.2.3 Type de Lipschitz :

Définition 2.2 *Les modèles fonctionnels non paramétriques (2.3), (2.5) et (2.7) seront appelés modèles non paramétriques fonctionnels de type de Lipschitz pour la même raison.*

2.2.4 Probabilités des petites boules :

Soit \mathcal{X} un v.a.f. a valeur dans E et soit χ un élément fixe de E . Nous pouvons écrire :

$$\mathbb{E} \left(1_{[0,1]} \left(\frac{d(\chi, \mathcal{X})}{h} \right) \right) = \mathbb{E}(1_{B(\chi, h)}(\mathcal{X})) = P(\mathcal{X} \in B(\chi, h)). \quad (2.8)$$

tel que

$$B(\chi, h) = \left\{ \chi' \in E, d(\chi, \chi') \leq h \right\}$$

qui est la boule, par rapport à la topologie induite par d , centrée à χ et de rayon h (où d est une semi-métrique sur E).

nous disons des probabilités 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 ∞). Ainsi, lorsque nous prenons une très grande quantité, h est proche de zéro, puis $B(\chi, h)$ est considéré comme une petite boule et $P(\mathcal{X} \in B(\chi, h))$ comme une petite probabilité de boule.

Nous utiliserons pour tout χ en E et pour tout réel réel h , la notation :

$$\varphi_\chi(h) = P(\mathcal{X} \in B(\chi, h)) \quad (2.9)$$

Où :

$$\forall \epsilon > 0, P(\mathcal{X} \in B(\chi, \epsilon)) = \varphi_\chi(\epsilon) \quad (2.10)$$

2.3 Estimateurs du noyau

2.3.1 Hypothèses nécessaires :

Dans le cadre non paramétrique multivarié, on suppose que la densité de la variable explicative multivariée est strictement positive et que l'hypothèse (2.10) est

une extension d'une telle notion. En outre les paramètres impliqués dans l'estimateur, c'est-à-dire le paramètre de lissage et la fonction du noyau, doivent satisfaire :

$$\left\{ \begin{array}{l} h \text{ est une séquence positive telle que} \\ \lim_{n \rightarrow \infty} h = 0 \text{ et } \lim_{n \rightarrow \infty} \frac{\log n}{n\varphi_x(h)} = 0 \\ K \text{ est un noyau de type I} \\ \text{ou} \\ K \text{ est un noyau de type II et } \varphi_x \text{ vrieifie que} \\ \exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u) du > C_3 \epsilon \varphi_x(\epsilon) \end{array} \right. \quad (2.11)$$

et on a

$$\left\{ \begin{array}{l} g \text{ est une séquence positive telle que : } \lim_{n \rightarrow \infty} g = 0 \\ K_0 \text{ est du type 0.} \end{array} \right. \quad (2.12)$$

2.3.2 Estimation non paramétrique de la régression

Nous proposons à l'opérateur non linéaire r l'estimateur de régression fonctionnel du noyau suivant :

$$\hat{r}(\chi) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi - \mathcal{X}_i))Y_i}{\sum_{i=1}^n K(h^{-1}d(\chi - \mathcal{X}_i))} \quad (2.13)$$

Où K est un noyau asymétrique et h (selon n) est un réel strictement positif.

Le changement principal provient du d semi-métrique qui mesure la proximité entre les objets fonctionnels.

2.3.3 Estimation non paramétrique de répartition conditionnelle

Nous concentrons maintenant sur l'estimateur $\hat{F}_Y^{\mathcal{X}}$ du conditionnelle fonction de répartition $F_Y^{\mathcal{X}}$, mais nous allons d'abord expliquer comment nous pouvons étendre l'idée précédemment utilisée pour la construction de l'estimateur par la régression du noyau. clairement :

$$F_Y^{\mathcal{X}}(\chi, y) = P(Y \leq y | \mathcal{X} = \chi) \quad (2.14)$$

peut être exprimé en termes d'attente conditionnelle :

$$F_Y^{\mathcal{X}}(\chi, y) = E(1_{]-\infty, y]}(Y) | \mathcal{X} = \chi) \quad (2.15)$$

soit K_0 un noyau symétrique habituel, On définit H comme :

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

et définir l'estimateur c.d.f. conditionnelle de noyau comme suit :

$$\widehat{F}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi - \mathcal{X}_i))H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi - \mathcal{X}_i))} \quad (2.16)$$

Où g est un nombre réel strictement positif (selon n) (défini en l'équation (2.12)).

Considérons K_0 comme un noyau de type $\mathbf{0}$, et nous pouvons écrire :

$$H(g^{-1}(y - Y_i)) = \begin{cases} 0 & \Leftrightarrow y \leq Y_i - g, \\ 1 & \Leftrightarrow y \geq Y_i + g. \end{cases} \quad (2.17)$$

2.3.4 Estimation non paramétrique de la densité conditionnelle

On sait que la fonction de densité conditionnelle peut être obtenue en dérivant le conditionnel c.d.f. (Voir 1.9). Comme nous avons l'estimateur $\widehat{F}_Y^{\mathcal{X}}(\chi, y)$ de $F_Y^{\mathcal{X}}(\chi, y)$, nous proposons l'estimation suivante :

$$\widehat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\partial}{\partial y} \widehat{F}_Y^{\mathcal{X}}(\chi, y) \quad (2.18)$$

en supposant la différentiabilité de H , nous avons

$$\frac{\partial}{\partial y} \widehat{F}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i)) \frac{\partial}{\partial y} H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))}$$

alors

$$\widehat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i)) \frac{1}{g} H'(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))} \quad (2.19)$$

plus généralement, nous pouvons affirmer pour n'importe quel noyau K_0 la définition suivante :

$$\widehat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i)) \frac{1}{g} H K_0(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))} \quad (2.20)$$

présentée par Ferraty (voir [1]).

2.4 La convergence presque complète de la densité conditionnelle

La condition de continuité de Hölder :

La condition de continuité de Hölder suivante doit être ajoutée pour l'estimation de densité conditionnelle du noyau $\widehat{f}_Y^{\mathcal{X}}(x, y)$ définie par (2.20) :

$$\left\{ \begin{array}{l} \exists C < \infty, \forall (x, x') \in \mathbb{R} \times \mathbb{R}, |K_0(x) - K_0(x')| \leq C|x - x'|, \\ \lim_{n \rightarrow \infty} \frac{\log n}{n g \varphi_{\mathcal{X}}(h)} = 0 \text{ et } \exists \zeta > 0, \lim_{n \rightarrow \infty} gn^{\zeta} = \infty \end{array} \right. \quad (2.21)$$

Proposition 2.1 *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 :*

$$X_n Y_n = O_{p.co.}(u_n).$$

Lemme 2.1 *Selon les hypothèses (2.10) et (2.11), on a :*

$$\widehat{r}_1(\mathcal{X}) - 1 = O_{p.co.} \left(\sqrt{\frac{\log n}{n \varphi_{\mathcal{X}}(h)}} \right).$$

Preuve. Voir [1] ■

Lemme 2.2 *Si K est un noyau de type I, il existe des constantes réelles C et C' positifs tel que :*

$$C \varphi_{\mathcal{X}}(h) \leq \mathbb{E}K \left(\frac{d(\mathcal{X}, \mathcal{X})}{h} \right) \leq C' \varphi_{\mathcal{X}}(h) \quad (2.22)$$

Preuve. Voir [1] ■

Lemme 2.3 *Si K est un noyau de type II et $\varphi_{\mathcal{X}}$ vérifie que :*

$$\exists \epsilon_0, \forall \epsilon < \epsilon_0, \int_0^{\epsilon} \varphi_{\mathcal{X}}(u) du > C_3 \epsilon \varphi_{\mathcal{X}}(\epsilon) \quad (2.23)$$

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

$$C \varphi_{\mathcal{X}}(h) \leq \mathbb{E}K \left(\frac{d(\mathcal{X}, \mathcal{X})}{h} \right) \leq C' \varphi_{\mathcal{X}}(h) \quad (2.24)$$

Preuve. Voir [1] ■

Corollaire 1 (*L'équation de Bernstein*)

Si $\exists M < \infty$, et on note $\sigma^2 = \mathbb{E}Z_1^2$, tel que les Z_i sont des v.a.r. indépendants, nous avons

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

Théorème 2.1 *Sous le modèle de type-continuité défini par (2.6) et (2.10), et si l'estimation du noyau satisfait (2.11), (2.12) et la condition de continuité de Hölder(2.21), nous avons pour tout sous-ensemble compact $S \subset \mathbb{R}$:*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |f_Y^X(y) - \hat{f}_Y^X(y)| = 0, \quad p.co. \quad (2.25)$$

qui a été présenté par Ferraty (voir[1])

Preuve. La preuve est basée sur la décomposition suivante

$$\begin{aligned} \hat{f}_Y^X(y) - f_Y^X(y) &= \frac{\{(\hat{r}_4(\chi, y) - \mathbb{E}\hat{r}_4(\chi, y)) - (f_Y^X(y) - \mathbb{E}\hat{r}_4(\chi, y))\}}{\hat{r}_1(\chi)} \\ &\quad - \frac{f_Y^X(y)}{\hat{r}_1(\chi)} \{\hat{r}_1(\chi) - 1\} \end{aligned} \quad (2.26)$$

Où

$$\hat{r}_1 = \frac{1}{n} \sum_{i=1}^n \Delta_i, \quad \text{avec} \quad \Delta_i = \frac{K(h^{-1}d(\chi - \mathcal{X}_i))}{\mathbb{E}K(h^{-1}d(\chi - \mathcal{X}_1))}$$

et

$$\hat{r}_4(\chi, y) = \hat{r}_1(\chi) f_Y^X(y) = \frac{1}{n} \sum_{i=1}^n \Delta_i \Omega_i(y)$$

avec

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

Le côté droit de (2.26). Notez d'abord que les dénominateurs sont directement traités en utilisant le lemme (2.1) avec la proposition (2.1).

on a

$$\hat{r}_1(\chi) - 1 = O_{p.co} \left(\sqrt{\frac{\log n}{n\varphi_\chi}(h)} \right) \text{ avec } u_n = \sqrt{\frac{\log n}{n\varphi_\chi}(h)}$$

on suppose que :

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

on applique la proposition (2.1) alors :

$$\frac{f_Y^\chi(y)}{\hat{r}_1(\chi)} \{\hat{r}_1(\chi) - 1\} = O_{p.co}(u_n) \text{ avec } u_n = \sqrt{\frac{\log n}{n\varphi_\chi}(h)}$$

Notez également que le dernier terme est traité en utilisant également lemme (2.1) et que $f_Y^\chi(\cdot)$ est uniformément continue sur $y \in S$ (puisqu'il est continu sur l'ensemble compact S). Par conséquent, le résultat (2.25) sera une conséquence directe des deux assertions suivantes :

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\mathbb{E}\hat{r}_4(\chi, y) - f_Y^\chi(y)| = 0, \quad p.co., \quad (2.27)$$

et

$$\lim_{n \rightarrow \infty} \frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\hat{r}_4(\chi, y) - \mathbb{E}\hat{r}_4(\chi, y)| = 0, \quad p.co., \quad (2.28)$$

• Le résultat (2.27) est confirmée, parce que $\mathbb{E}\Delta_1 = 1$ et parce que K_0 est intégré jusqu'à un, nous avons, après intégration par substitution :

$$\begin{aligned} \mathbb{E}\hat{r}_4(\chi, y) - f_Y^\chi(y) &= \mathbb{E}\Delta_1\Omega_1(y) - f_Y^\chi(y) \\ &= \mathbb{E}(\Delta_1(\mathbb{E}(\Omega_1(y)|\mathcal{X}) - f_Y^\chi(y))) \\ &= \mathbb{E}\left(\Delta_1\left(\int_{\mathbb{R}} g^{-1}K_0(g^{-1}(y-u))f_Y^\chi(u)du - f_Y^\chi(y)\right)\right) \\ &= \mathbb{E}\left(\Delta_1\int_{\mathbb{R}} g^{-1}K_0(g^{-1}(y-u))(f_Y^\chi(u) - f_Y^\chi(y))du\right) \\ &= \mathbb{E}\left(\Delta_1\int_{\mathbb{R}} K_0(v)(f_Y^\chi(y-vg) - f_Y^\chi(y))dv\right) \\ &= \mathbb{E}\left(1_{B(\chi, h)}(\mathcal{X})\Delta_1\int_{\mathbb{R}} K_0(v) \times (f_Y^\chi(y-vg) - f_Y^\chi(y))dv\right). \end{aligned} \quad (2.29)$$

En outre, nous pouvons écrire :

$$|f_Y^\chi(y-vg) - f_Y^\chi(y)| \leq |f_Y^\chi(y-vg) - f_Y^\chi(y-vg)| + |f_Y^\chi(y-vg) - f_Y^\chi(y)|.$$

Puisque S est compact, la fonction f_Y^χ est uniformément continue sur S . Étant donné que K_0 est de support compacte (car il est du type 0), et parce que le modèle de continuité (6.2), nous avons

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

et

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

ce qui implique que

$$\lim_{n \rightarrow \infty} \sup_{v \in [-1, +1]} \sup_{y \in S} 1_{B(\chi, h)}(\mathcal{X}) |f_Y^\chi(y - vg) - f_Y^\chi(y)| = 0 \quad (2.30)$$

En combinant (2.29) et (2.30), avec la positivité de Δ_1 et K_0 , on obtient :

$$\sup_{y \in S} |\mathbb{E}\hat{r}_4(\chi, y) - f_Y^\chi(y)| = o(1) \quad (2.31)$$

Ceci, combiné au lemme (2.1) et à la Proposition (2.1), suffit à prouver (2.27).

• Il reste à vérifier que (2.28) est vrai. En utilisant la compacité de S , on peut écrire que $S \subset \cup_{k=1}^{z_n} S_k$ où $S_k = (t_k - l_n, t_k + l_n)$ et où l_n et z_n peuvent être choisis de telle sorte que :

$$l_n = Cz_n^{-1} \sim Cn^{-2\zeta} \quad (2.32)$$

prise $t_y = \arg \min_{t \in \{t_1, t_2, \dots, t_{z_n}\}} |y - t|$,

$$\frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\hat{r}_4(\chi, y) - \mathbb{E}\hat{r}_4(\chi, y)| = A_1 + A_2 + A_3, \quad (2.33)$$

où

$$\begin{aligned} A_1 &= \frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\hat{r}_4(\chi, y) - \hat{r}_4(\chi, t_y)|, \\ A_2 &= \frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\hat{r}_4(\chi, t_y) - \mathbb{E}\hat{r}_4(\chi, t_y)|, \\ A_3 &= \frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\mathbb{E}\hat{r}_4(\chi, t_y) - \mathbb{E}\hat{r}_4(\chi, y)|. \end{aligned}$$

La condition de continuité de Hölder (2.21) nous permet d'écrire directement :

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

En utilisant (2.21), on a :

$$\begin{aligned}
|\hat{r}_4(\chi, y) - \hat{r}_4(\chi, t_y)| &\leq C \hat{r}_1(\chi) l_n g^{-2} \\
&\leq C \hat{r}_1(\chi) (gn^\zeta)^{-2}
\end{aligned}$$

alors

$$\frac{1}{\hat{r}_1(\chi)} \sup_{y \in S} |\hat{r}_4(\chi, y) - \hat{r}_4(\chi, t_y)| \leq \frac{C}{(gn^\zeta)^2}$$

il vient que $A_1 \leq C/(gn^\zeta)^2$, et (6.31) implique que :

$$\lim_{n \rightarrow \infty} A_1 = 0. \tag{2.34}$$

Suite à des arguments similaires, nous pouvons écrire :

$$A_3 \leq \frac{C}{\hat{r}_1(\chi) (gn^\zeta)^2}, \tag{2.35}$$

et selon le lemme (2.1) et la proposition (2.1), nous obtenons :

$$\lim_{n \rightarrow \infty} A_3 = 0, \text{ p.co.} \tag{2.36}$$

En regardant maintenant le terme A_2 , nous pouvons écrire pour quelque soit $\epsilon > 0$:

$$\begin{aligned}
P(\sup_{y \in S} |\hat{r}_4(\chi, t_y) - \mathbb{E} \hat{r}_4(\chi, t_y)| > \epsilon) &= P(\max_{j=1, \dots, z_n} |\hat{r}_4(\chi, t_j) - \mathbb{E} \hat{r}_4(\chi, t_j)| > \epsilon) \\
&\leq z_n \max_{j=1, \dots, z_n} P(|\hat{r}_4(\chi, t_j) - \mathbb{E} \hat{r}_4(\chi, t_j)| > \epsilon) \\
&\leq z_n \max_{j=1, \dots, z_n} P\left(\left| \frac{1}{n} \sum_{i=1}^n (U_i - \mathbb{E} U_i) \right| > \epsilon\right),
\end{aligned} \tag{2.37}$$

où

$$U_i = \Delta_i \Omega_{i(t_j)}$$

En utilisant lemme (2.2) ou lemme (2.3) selon le fait que K est de type I ou II , et parce que $\Omega_{i(y)} \leq C/g$, nous avons :

$$|U_i| \leq C/(g\varphi_\chi(h)). \quad (2.38)$$

D'autre part, nous avons, après intégration et par substitution, et en utilisant (2.32) :

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

parce que f_Y^χ est bornée (car il est continu sur l'ensemble compact S) et en appliquant cet résultat qui présenté par Ferraty (voir [1]) :

$$\mathbb{E}|Y_1|^m \Delta_1^m = O(\varphi_\chi(h)^{-m+1})$$

en appliquant avec U_i et $m = 2$, on a :

$$\mathbb{E}|U_i|^2 \Delta_i^2 = O(\varphi_\chi(h)^{-1})$$

alors en obtient :

$$\mathbb{E}U_i^2 \leq C/(g\varphi_\chi(h)). \quad (2.39)$$

Parce que les variables U_i sont bornées, on est donc en mesure d'appliquer l'inégalité exponentielle du Bernstein donnée par corollaire 1. Cette inégalité avec (2.37), (2.38) et (2.39) donne directement :

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

En utilisant (2.32), on obtient :

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

Parce que $\log n/(ng\varphi_\chi(h))$ tend vers zéro, on obtient directement :

$$\forall \epsilon > 0, \sum_{i=1}^{\infty} P(\sup_{y \in S} |\hat{r}_4(\chi, t_y) - \mathbb{E}\hat{r}_4(\chi, t_y)| > \epsilon) < \infty. \quad (2.40)$$

Le dénominateur de A_2 est traité directement en utilisant à nouveau le lemme (2.1) avec la proposition (2.1). C'est suffisant pour obtenir :

$$\lim_{n \rightarrow \infty} A_2 = 0, \text{ p.co.} \quad (2.41)$$

Enfin, le résultat réclamé (2.28) découle de (2.33), (2.34), (2.36) et (2.41).

La combinaison de (6.35), (6.36) et (6.37) nous permet de terminer la preuve de ce lemme. ■

Chapitre 3

Application et exploration des observations sur des données simulées

La densité conditionnelle est un outil fondamental pour décrire la relation entre deux variables aléatoires.

Dans ce chapitre, nous allons déterminer ce lien par la méthode d'estimation non paramétrique de noyau. L'objectif principal est de montrer à l'aide de données simulées l'applicabilité de cette méthode dans le cadre fonctionnel. Dans un premier temps, nous illustrons les variables aléatoires fonctionnelle sont des courbes. Ensuite, nous proposons une implémentation de l'estimateur de la densité conditionnelle proposé dans le chapitre 2.

3.1 Simulée les données fonctionnelles

Nous récoltons des observations fonctionnelles générées à l'aide du processus suivant :

$$\mathcal{X}_i(t) = \cos(w_i + \pi(\frac{2t}{100} - 1))$$

où w_i est distribué selon la loi uniforme $U(0, \frac{\pi}{4})$.

Nous effectuerons une étude de simulation à l'aide du logiciel statistique *R*

3.1.1 Programme sous R

```
x <- matrix(nrow = 100, ncol=100, byrow=TRUE)
xx <- (1 :100)
```

```

w= runif(100,0,pi/4) # échantillon de loi uniforme
for (i in 1 :100) {
for (j in 1 :100) {
x[i,j]=cos(w[i] + pi * ((2 * j/100) - 1))^4
}
}
# boucle pour les observations de v.a.f
yy <- x[1, ]      #Vecteur de premier v.a.f
plot(yy,type="l",col="blue",xlab="t", ylab="X(t)", main="Un courbe aléatoire si-
mulée")
for (j in 1 :100) {
lines(xx,x[j,], col="black")
}      # Tracer un courbe qui représente un v.a.f
plot(xx,yy,type="l",col="blue",xlab="t", ylab="X(t)", main="plusieurs courbes aléa-
toires simulées")
for (j in 1 :100) {
lines(xx,x[j,], col="black")
}      # Tracer les courbes qui représente les v.a.f.

```

3.2 Résultats de simulation

On s'appuie sur le programme ci dessus, on obtient la courbe illustrée dans figure (3.1) qui représente la v.a.f. Comme la figure (3.2) représente la courbes simulées

3.2.1 Résultat un variable aléatoire fonctionnelle :

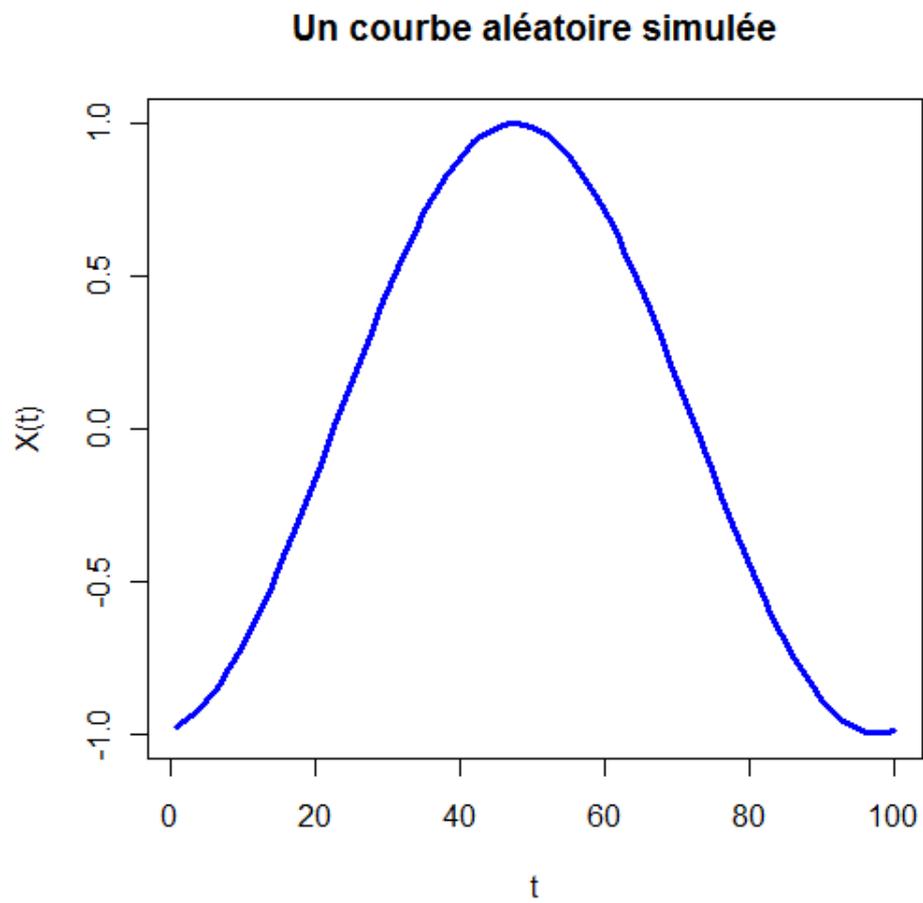


FIGURE 3.1 – Courbe de un v.a.f

3.2.2 Résultat de plusieurs variables aléatoires fonctionnelles :

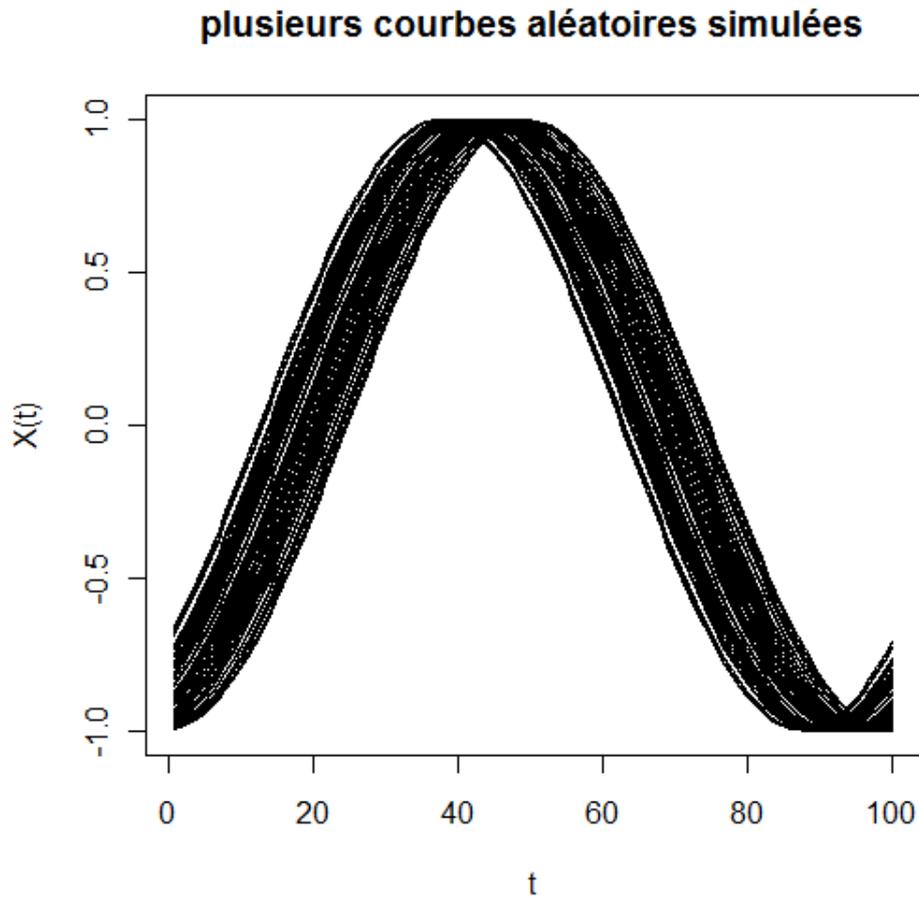


FIGURE 3.2 – Courbes de plusieurs v.a.f

3.3 Simulation les variables de réponse réels

Pour trouver les variables de réponses réels que nous donnerons le terme de régression multiple suivante :

$$Y_i = r(\mathcal{X}_i) + \varepsilon_i$$

tell que les \mathcal{X}_i est une matrice, et pour chaque i nous avons un vecteur qui présente une variables aléatoire fonctionnelle.

Et l'erreur ε_i est centrée réduit. On utilise les variables aléatoires fonctionnelles qui sont utilisé dans la section précédente.

On suit les étapes de l'algorithme suivant :

3.3.1 Algorithme 1

1. Générer un échantillon de taille n d'une variable normale centrée réduit qui présent l'erreur ε_i ($\varepsilon = rnorm(n)$)
2. Nous choisirons cette fonction pour représenter la régression :

$$r(\mathcal{X}_i) = \mathcal{X}_i^2 + 2$$

3. Nous prendrons les valeurs de variables de réponse.

3.3.2 Programme sous R

```
n<-100
err<-rnorm(100,0,1) # échantillon de loi N(0, 1)
f<-function(x)
for (i in 1 : 100) {
f[i]<-(x[i]^2 + 2)
}
yr<-f[i]+err[i]
for (i in 1 : 100) {
yr[i]=f[i]+err[i]
}
plot(f,type="l",col="blue",xlab="x", ylab="f(x)", main="courbe de f")
for (i in 1 :100){
lines(f[i], col="red")
}
```

3.4 Simulation l'estimateur de densité conditionnelle

Dans cette section, nous calculons et représentons graphiquement l'estimateur de densité conditionnelle. Il s'agit de l'estimateur proposé par Ferraty et Vieu (2006) et présenté au chapitre 2.

Cette estimation pour le couple (χ, y) de (\mathcal{X}, Y) est donnée par :

$$\widehat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i)) \frac{1}{g} HK_0(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))}$$

Afin d'estimer doit choisir K un type de noyaux présentée dans le chapitre 1, nous choisirons le noyau gaussien sous forme :

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp^{-u^2/2}$$

et le noyau K_0 est un noyau symétrique, nous choisirons le noyau rectangulaire :

$$K_0(u) = \frac{1}{2} 1_{[-1,1]}(u)$$

Pour estimer plus facilement nous calculerons les noyaux composés $HK_0(y - Y_i)$, tel que :

$$\begin{aligned} \forall \mu \in \mathbb{R}, H(\mu) &= \int_{-\infty}^{\mu} K_0(v) dv \\ &= \int_{-1}^{\mu} \frac{1}{2} 1_{[-1,1]}(v) dv \\ &= \frac{\mu + 1}{2} \end{aligned}$$

et aussi :

$$\begin{aligned} K_0(y - Y_i) &= \frac{1}{2} 1_{[-1,1]}(y - Y_i) \\ &= \begin{cases} \frac{1}{2} & \text{si } -1 \leq y - Y_i \leq 1 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

alors les noyaux composés $HK_0(y - Y_i)$ est sous forme :

$$\begin{aligned}
H(K_0(y - Y_i)) &= \frac{1}{2}(K_0(y - Y_i)) + \frac{1}{2} \\
&= \begin{cases} \frac{3}{4} & \text{si } -1 \leq y - Y_i \leq 1 \\ \frac{1}{2} & \text{sinon} \end{cases}
\end{aligned}$$

On suit les étapes de l'algorithme suivant pour simuler l'estimateur $\widehat{f}_Y^{\mathcal{X}}$ à couple (χ, y) (on fixe (χ_1, y_1)) :

3.4.1 Algorithme 2

1. On choisit la valeur de paramètre de lissage h et g .
2. Calcule $y - Y_i$
3. On insère $K(h^{-1}d(\chi, \mathcal{X}_i))$
4. On insère $H(K_0(y - Y_i))$
5. On doit calculer l'estimateur $\widehat{f}_Y^{\mathcal{X}}$
6. On présente $\widehat{f}_Y^{\mathcal{X}}$

3.4.2 Programme sous R

Le programme sous R est donné par :

```

h <- 0.44 # paramètre de lissage
z <- yr[1]-yr[i]
zz <- (x[1,]-x[i,])/h
for (i in 1 :100){
z[i]=yr[1]-yr[i] zz[i]=(x[1,]-x[i,])/h
}
g <- 0.34 # paramètre de lissage
kz <- function(z)
for (i in 1 :100) {
kz[i]=(1/sqrt(2*pi))*exp(-(1/2)*((z[i])^2))
}

```

```

hk <- function(z) # noyau composée  $HK_0$ 
for (i in 1 :100){
  if(z[i]<=1 z[i]>=-1)
hk[i]<- 3/4
else hk[i]<- 1/2
}
s<-function(kz,hk) for (i in 1 :100) {
s[i]<-(sum((1/g)*(kz[i]*hk[i])))
}
ss<-function(kz)
for(i in 1 :100)
ss[i]<-sum(kz[i])
}
fchp <- numeric(n) # l'estimateur  $\widehat{f}_Y^X$ 
for (i in 1 :100){ fchp[i]<-sum(s[i]*ss[i])/(sum(ss[i])) }
plot(fchp,type="l",col="black",xlab="t", ylab="f,fchp",xlim=c(0,100) ,ylim=c(0,2),
lwd="2",main="Courbe de f et fchp" )
for(i in 1 :100) {
lines(fchp, col="black",lwd="2")
} # présentée  $\widehat{f}_Y^X$ 

```

3.5 Comparation

Afin de faire la comparaison, nous devons proposer une autre fonction, qui doit vérifier les conditions de la fonction de densité (i.e $\int_{-\infty}^{+\infty} f = 1$ et f continue,)

On choisit la fonction suivante :

$$p(\chi, y) = 2 \exp -(\chi + y) 1_{\{0 \leq \chi \leq y\}}$$

On présent programme de p sous R :

3.5.1 Programme sous R

```
p<-function(x)
for (i in 1 :100) {
if(x[i]<=yr[i] x[i]>=0)
p[i]<- ( 2*exp(-(x[i]+yr[i])) )
else p[i]<- 0
}
plot(p,col="blue")
plot(fchp,type="l",col="black",xlab="t", ylab="f,fchp",xlim=c(0,100)
,ylim=c(0,5),main="Courbe de f et fchp" )
for (i in 1 :100){
lines(p, col="blue",lwd="2")
}
```

3.6 Résultats des simulations

3.6.1 Simulation de l'estimateur de densité conditionnelle

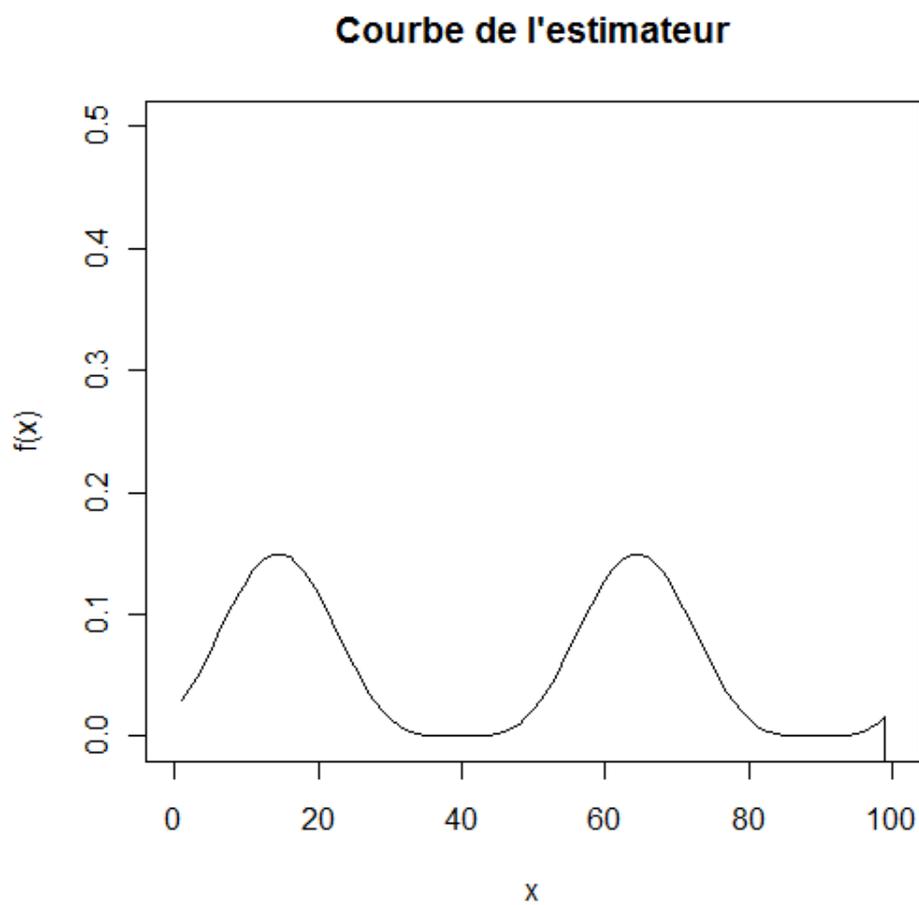


FIGURE 3.3 – Comportement graphique de l'estimateur de densité conditionnelle \hat{f}_Y^x

3.6.2 Simulation de la fonction p

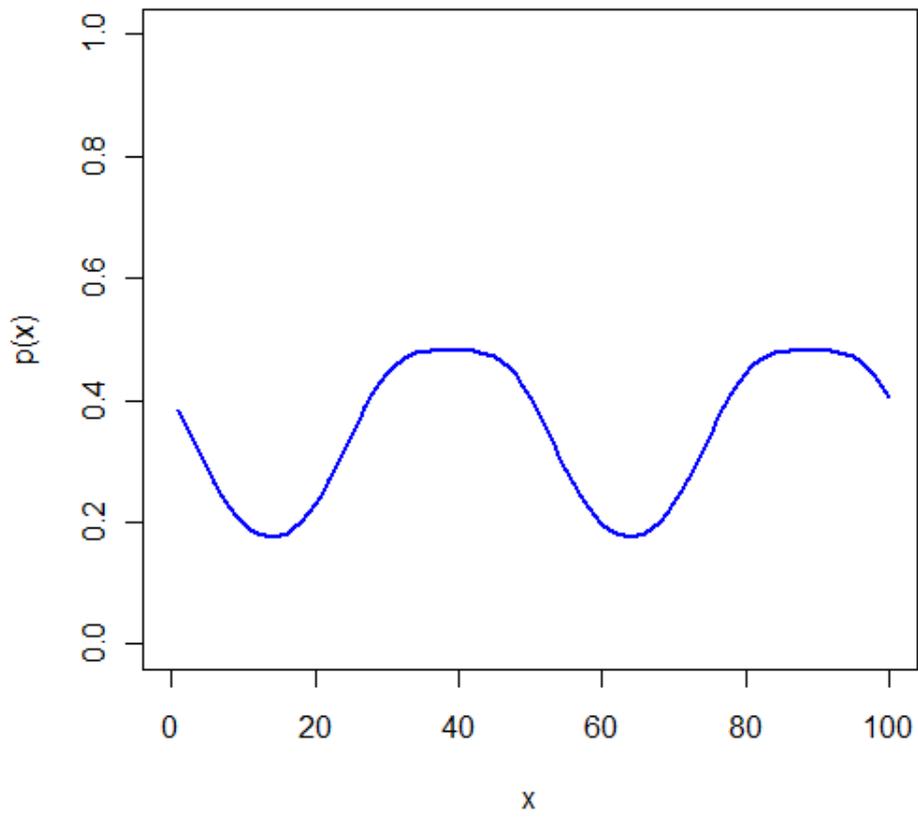


FIGURE 3.4 – Courbe de la fonction p

3.6.3 La fonction p et l'estimateur f au même plan :

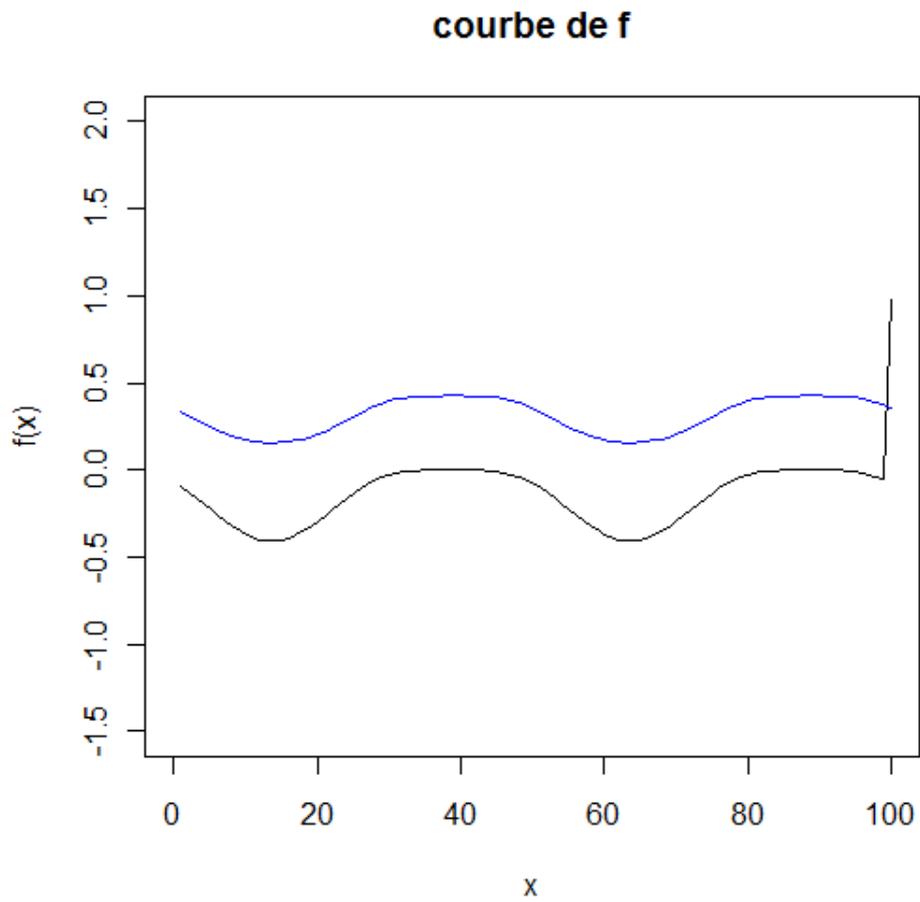


FIGURE 3.5 – Courbe de la fonction p et l'estimateur de f

3.6.4 Commentaires

Après les résultats des simulations, nous avons obtenu les deux courbes sont proches, donc les approximations sont bonnes.

Conclusion

Dans ce travail nous avons présenté d'une part des outils mathématique existants dans la littérature, et qui sont en particulier nécessaire pour la résultat du problème posé par cette dimension infinie.

Nous avons étudié l'estimation non paramétrique de la densité conditionnelle par la méthode de noyau dans le cadre fonctionnelle, où les variables sont dans le cas i.i.d. où nous avons conclu après l'étude théorique et pratique que nous avons fait, que l'estimateur de densité conditionnelle converge presque complète vers les valeurs simulées.

Cet mémoire montre que la convergence presque complète de l'estimateur de densité conditionnelle qui présentée par Ferraty et vieu (2006) est confirmée, et c'est la meilleure approximation.

Bibliographie

- [1] F. FERRATY et P. VIEU : Nonparametric functional data analysis : Theory and practice. Springer, 2006.
- [2] G. COLLOMB, W. HÄRDLE et S. HASSANI : A note on prediction via estimation of the conditional mode function. *Journal of Statistical Planning and Inference*, 15 :227-236, 1987.
- [3] G. COLLOMB, S. HASSANI, P. SARDA et P. VIEU : Estimation non paramétrique de la fonction de hasard pour des observations dépendantes. *Statistique et analyse des données*, 10(3) :42-49, 1985.
- [4] Bashtannyk, D.M. and Hyndman, R.J. (2001). Bandwidth selection for kernel conditional density estimation. *Comput. Statist. Data Anal.*, 36, Pages 279-298.
- [5] Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *Pub. Inst. Stat. Univ. Paris*, 3, Pages 27-42.
- [6] Collomb, G. Estimation nonparamétrique de la régression (in french). PhD Université Paul Sabatier, Toulouse (1976).
- [7] Hsu, P.L. And Robbins, H. (1974). complete Convergence And The Law Of Large Numbers. *proc. Nat. Acad. Sci. Usa* 33 25-31.
- [8] Lagodowski, Z.A. And Rychlik, Z. (1985). complete Convergence And Convergence Rates For Randomly Indexed Sums Of Random Variables With Multidimensional Indices. *bull. Acad. Sci. Polon. Ser. Math.* 32 219-223.
- [9] Szynal, D. (1972). on Almost Complete Convergence For The Sum Of A Random Number Of Independent Random Variables. *bull. Acad. Sci. Polon. Ser. Math. Astronom. Phys.* 20 571-574.

- [10] Hsu, P.L. And Robbins, H. (1974). complete Convergence And The Law Of Large Numbers. *proc. Nat. Acad. Sci. Usa* 33 25-31.
- [11] Gut, A. (1993). complete Convergence And Ces Summation For I.I.D. Random Variables. *probab. Th. Rel. Fields* 97 169-178.
- [12] Erdos, P. (1950). remark On My Paper "On A Theorem Of Hsu And Robbins". *ann. Math. Statist.* 21 138.
- [13] Erdős, P. (1949). on A Theorem Of Hsu And Robbins. *ann. Math. Statist.* 20 286-291.
- [14] Asmussen, S. And Kurtz, T.G. (1980). necessary And Sufficient Conditions For Complete Convergence In The Law Of Large Numbers. *ann. Probab.* 8 176-182.
- [15] Gut, A. (1993). complete Convergence And Ces Summation For I.I.D. Random Variables. *probab. Th. Rel. Fields* 97 169-178.
- [16] Attouch, M., Laksaci, A., Ould-Saïd, E. (2009). Asymptotic Distribution of Robust Estimator for Functional Nonparametric Models. *Communications in Statistics : Theory and Methods.* 38, 1317-1335.
- [17] Bosq, D., Lecoutre, J. P.(1987). *Théorie de l'estimation fonctionnelle.* ECONOMICA, Paris.
- [18] Burba, F., Ferraty, F., Vieu, P. (2008). Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique. *C. R. Acad. Sci. Paris.* 346, 339-342.
- [19] Carbon, M., Francq, C., Tran, L. T. (2007). Kernel regression estimation for random fields. *J. Statist. Plann. Inference.* 137, 778-798.
- [20] Cardot, H., Crambes, C., Sarda, P. (2004). Spline estimation of conditional quantiles for functional covariates, *C. R. Math. Acad. Sci. Paris.* 339, 141-144.
- [21] Collomb, G. (1983). *Méthodes non paramétriques en régression, analyse de séries temporelles, prédiction et discrimination.* Doctorat d'état, Toulouse
- [22] Collomb, G. (1984). Propriétés de convergence presque complète du prédicteur à noyau. *Z. W. Gebiete.* 66, 441-460.
- [23] Collomb, G. (1985). Nonparametric regression : an up to date bibliography *Statistics.* 16, 309-324.

- [24] Deheuvels, P., Mason, D. (2004). General asymptotic confidence bands based on kernel type function estimators. *Statistical Inference for Stochastic Processes*, 7, 225-277.
- [25] Delsol, L. (2007). Régression non paramétrique fonctionnelle : expression asymptotique des moments, *Ann. I.S.U.P. Vol LI*, 3, 43-67.
- [26] Devroye, L. (1978). The uniform convergence of the Nadaraya-Watson regression function estimate. *Canad. J. Statist.* 6, No.2, 179-191.
- [27] Ezzahrioui, M. (2007). Prédiction dans les modèles conditionnels en dimension infinie. PhD Thesis. Univ. du Littoral Côte d'Opale.
- [28] Ferraty, F., Goia, A., Vieu, P. (2002). Functional Nonparametric Model for Time Series : A Fractal Approach for Dimension Reduction. *TEST*, 11, 317-344.
- [29] Ferraty, F., Laksaci, A., Tadj, A., Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference.* 140, 335-352.
- [30] Ferraty, F., Laksaci, A., Tadj, A., Vieu, P. (2011). Kernel regression with functional response. *Electronic Journal of Statistics.* 5, 159-171.
- [31] Ferraty, F., Laksaci, A., Vieu, P. (2005). Functional time series prediction via conditional mode estimation. *C. R. Math. Acad. Sci. Paris.* 340, 389-392.
- [32] Györfi, L., Härdle, W., Sarda, P. et Vieu, P. (1989). Nonparametric curve estimation for time series, *Lecture Notes in Statistics.* 60, Springer- Verlag.
- [33] Laksaci, A., Madani, F., Rachdi, M. (2010). Kernel conditional density estimation when the regressor is valued in a semi metric space. *International Statistical Review.* (In press).
- [34] Li, W. V., Tran, L. T. (2009). Nonparametric estimation of conditional expectation. *J. Statist. Plann. Inference.* 139, 164-175.

Résumé :

L'objet de mémoire est d'établir une synthèse des résultats limite de l'estimation non paramétrique de la densité conditionnelle pour les variables aléatoires fonctionnelles (c'est-à-dire des variables qui prennent leurs valeurs dans des espaces de dimension infinie), et dans le cas i.i.d (indépendant et identiquement distribuée), et en utilise la méthode de noyau (Parzen-Rozenblatt), qui a été présentée par Ferraty (2006).

Mots Clés : Méthode de noyau, densité conditionnelle, régression, donnée fonctionnelles.

Abstract :

The object of memory is to establish a synthesis of the limit results of the nonparametric estimation of the conditional density for the functional random variables (that is to say variables that take their values in spaces of infinite dimension) , and In the case iid (independent and identically distributed), and uses the kernel method (Parzen-Rozenblatt), Which was presented by Ferraty (2006).

Key words: kernel method, conditional density, regression, functional data.

ملخص:

الهدف من المذكرة هو لإعداد نتيجة تقريبية للمقدر غير المعلمي للكثافة الشرطية من أجل المتغيرات العشوائية الوظيفية (بمعنى متغيرات تأخذ قيمها في فضاء ذو أبعاد غير منتهية) وفي حالة المتغيرات مستقلة ولها نفس التوزيع. و يستخدم طريقة النواة (بريزون-بلات) ، الذي قدم من طرف فيراتي (2006).

الكلمات الرئيسية: طريقة النواة ، الكثافة الشرطية، الانحدار ، معطيات وظيفية.