République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2017/2018

L'estimation du quantile conditionnel

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité : Analyse Stochastique,Statistique des processus et Applications

par

Abdelaziz **M**essadi [1]

Sous la direction de

Dr.N. Hachemi

Soutenu le 19/06/2018 devant le jury composé de

Mr.M Laouni	Université de Saida - Dr Moulay Tahar	Président
Dr.N Hachemi	Université de Saida - Dr Moulay Tahar	Encadrice
Dr.S Rahmani	Université de Saida - Dr Moulay Tahar	Examinatrice
Dr.R Rouane	Université de Saida - Dr Moulay Tahar	Examinatrice

1. e-mail : messadiabdelaziz26@gmail.com

Dédicace

Rien n'est plus dur pour tout étudiant que de trouver là où je suis. Je dédie ce modeste travail à mes chers parents qui m'ont donnée la possibilité de poursuivre mes études,pour leurs guides affectueux ,pour l'espoir qu'ils m6e donnent ,pour les conseils dans la vie et leurs soutiens durant mes années d'études avec patience,courage et j'espère que je puisse leurs rend la maximum de bonheur qu'ils m'ont offert et que dieu les bénisse.

Ames soeurs.
Ames frères.
Ames cousins et cousines
A tout mes amis ainis qu'a tous mes camarades.
A tous mes chers proffeseurs durant toutes mes années d'études.
Finalement, à tous ceux qui me sont chères, qu'ils trouvent ici l'expression de mon Profond respect et ma gratitude.

A tous ce qui m'aiment.
A tous ceux qui j'aime.
Je dédie ce mémoire.

Remerciements

Je tiens à remercie tout d'abord mon encadreur pour sa patience, ses remarques, ses conseils, sa disponibilité et sa bienveillance durant la réalisation de ce travail.
Je remercie également les membres de jury d'avoir accepter de lire et d'évaluer ce travail.
Un grand merci adressé au personnel et enseignants du département de mathémathique.
Puis, je remercie ceux qui m'aident de proche ou de loin pour ce travail.

Résumé

Dans cette thèse, nous nous intéressons au problème de la prévision en considérant des modèles complètes et incomplètes non paramétrique pour le quantile conditionnel dont la variable explicative est réel et fonctionnel. Plus précisément, les points étudiés entre une variable réponse réelle et une variable réel et fonctionnelle est le quantile conditionnel. Nous établissons la convergence presque complète, la normalité asymptotique d'un estimateur à noyau. Ces propriétés asymptotiques sont obtenus sous des conditions assez générales telles, l'hypothèse de mélange forte et l'hypothèse de concentration de la mesure de probabilité de la variable explicative fonctionnelle. Le modèle des quantiles conditionnels est abordé dans la deuxième partie est traité comme fonction inverse de la fonction de répartition conditionnelle qui est estimée par un estimateur tronquée à gauche. Sous les mêmes conditions que celles du modèle précédent, nous donnons l'expression de la vitesse de convergence et nous démontrons la normalité asymptotique de l'estimateur construit.

Table des matières

Introduction

L'estimation des quantiles conditionnels joue un rôle cruciale. elle est utilisée comme outils de prévision, pour la détermination des intervalles prédictive. la détermination des courbes des références ou à la détermination d'événements rares. la modélisation via l'estimation des quantiles conditionnels est mieux sensible à ce type d'erreurs qui proviennent d'une mauvaise lecture ou d'un mauvaise enregistrement. L'objectif de ce mémoire est d'étudier cette alternative de prévision non paramétrique dans le cas des données réels et fonctionelles par des données complètes et incomplètes.
Dans le cas des données complètes, de nombreux résultats ont été énoncés ; Roussas (1969, [21]) a montré la convergence et la normalité aymptotique de l'estimateur du quantile conditionnel sous des conditions de Markov. Gannoun (1989, [**?**]) a étudie les propriétés de l'estimateur de Collomb dans le cas de données indépendantes et identiquement distribuées puis $\alpha-$ mélangentes. Stone ([22], 1977) a étudié l'estimation des quantiles conditionnelles. Il a obtenu la convergence en probabilité d'un estimateur basé sur l'estimation empirique de la fonction de répartition conditionnelles. En 1989, Samanta ([23]) a établit la normalité asymptotique et la convergence uniforme de l'estimateur à noyau de quantiles quantiles conditionnels dans le cas iid. ([2], 1991). Berlinet , Gannoun et Matzner (2001,[2]) étandent ces résultats aux données non indépendantes. ils considèrent le procesus (X_i, Y_i) stationnaire et $\alpha-$ mélangeant. ils énoncent des théorèmes qui établissent qu'un estimateur convergent du quantile construit à partir d'un estimateur convenables de F(./x) est asymptotiquement normal.
Dans le cas fonctionnel, les premiers résultats ont été obtenue par Cardot et al. ([5], 2004), ils ont construit un estimateur des quantiles conditionnels par la méthode de fonctions splines en considérant des observations iid la convergence presque complète et ont été généralisé au cas α-mélange par Ferraty et al ([**?**], 2005). Ezzahrioui et Ould-said ([6], 2005) ont étudié la normalité asymptotique quand la variable explicative est dans un espace normé vérifiant une condition de type fractale généralisé. Niang et Laksaci ([11], 2006) ont établit la convergence en norme L^p, la convergence presque complète et la normalité asymptotique sous l'hypothèse de concentration de la mesure de probabilité de la variable explicative. Plusieurs approches ont été développées pour l'estimation des quantiles conditionnels. L'approche paramétrique peut parfois être mal adaptée à la réalité des données en particulier biologiques. Une approche non paramétrique du problème a alors été développée afin de pallier les problèmes d'hypothèses et de modélisation paramétriques. De nombreux travaux récents ont été menés pour l'estimation non paramétrique des quantiles conditionnels aussi bien dans un cadre théorique que sur le plan des applications. Ces méthodes ne nécessitent pas d'hypothèse sur la nature de la distribution. Les quantiles conditionnels sont fréquemment utilisés en statistique. par exemple, la médiane est un indicateur robuste de la tendance centrale d'une population, l'intervalle interquartile est un bon indicateur de sa dispersion. Dans la pratique, les domaines d'utilisation

de quantiles sont assez variés. Les quantiles représentent également un moyen robuste de prévision. En pratique, ces quantiles sont calculés suivant un crictère d'ordre sur l'observations. Pour mieux apprendre les quantiles conditionnels, nous allons approfondir cette étude au cas ded donnés aléatoirement tronquées à gauche. Gürler, Stute et Wang ([14], 1993) ont donné une représentation pour la fonction quantile et sa normalité asymptotique. Son extension aux analyses de séries temporelles a été obtenue par Lemdani, Ould-Saïd et Poulin ([17], 2005).
Une note sur l'estimateur de la limite de produit sous la censure à droite et la troncature à gauche (1987)), Anderson, Borgan et Keiding ([1], 1993). Ces données se produisent en astronomie et en économie (voir Woodroofe ([24]), He et Yang ([15]).
Le présent mémoire est divisé en quatre chapitre et il considère les deux cas des observations : i.i.d et le cas mélane fort.
Le premier chapitre est consacré à l'introduction des définitions et outils techniques utilisées pour l'élaboration de nos résultat. En particulier, nous rappelons la définition de processus de mélange, les définitions de différents modes de convergence. On trouvera aussi dans ce chapitre un nombre très important outils intéressant pour l'élaboration des résultat obtenus.
Dans le deuxième chapitre, on considère le cas des observations indépendantes identiquement distribuées et on établit sous des conditions générales la convergence presque complète et la convergence en normalité asymptotique de l'estimateur à noyau du quantile conditionnel. On trouvera aussi le cas des obervations dépendantes, nous avons opté pour des conditions un peu plus restrictives, mais, qui donne la même vitesse de convergence du cas i.i.d.
Dans le troisième chapitre, on obordera le cas fonctinnel, nous avons opté pour des conditions dans un espace de dimension infini.
Le quatrième chapitre est consacré à des données incomplètes dans le cas tronquée à gauche par des données indépendants et dépendants.
Le dernier chapitre est consacré à l'étude de ce dernier modèle au cas fonctionnel.
On trouvera aussi quelques commentaires sur notre résultat.

Chapitre 1

Définitions et Notations

L'analyse de survie est un domaine des statistiques qui trouve sa place dans tous les champs d'application où l'on étudie la survenue d'un évènement. L'objectif de cette analyse réside dans l'analyse du délai de survenue d'un évènement dans un ou plusieurs groupes d'individus. Dans le domaine biomédical, par exemple, plusieurs évènements sont intéressants à étudier : le développement d'une maladie, la réponse à un traitement donné, la rechute d'une maladie ou le décés. Une des caractéristiques des données de survie est l'existence d'observations incomplètes. La censure et la troncature font partie de processus générant ce type de données. Dans ce chapitre nous allons rappeler quelques notions de base dans l'analyse de survie pour rendre la lecture plus facile.

1.1 La durée de survie et la date d'origine

La durée de survie, notée par T, définie comme le délai écoulé entre deux états (états 0 et1). Pour définir ce délai il est nécessaire de définir une date d'origine qui est la date de début du phénomène étudié. Par exemple, dans l'étude de l'évolution d'une maladie, la date d'origine T_0 est la date de début de la maladie et si on s'intéresse à l'âge du sujet à la survenue de l'évènement, la date d'origine sera la date de naissance du sujet $T_0 = 0$. Chaque individu peut avoir une date d'origine différente.

1.2 Troncature

Les troncatures différent des censures au sens où elles concernent l'échantillonnage lui même. Ainsi une variable X est tronquée par un sous ensemble éventuellement aléatoire A de $\mathbb{R}_+$ si au lieu de X, on observe X uniquement si $X \in A$. les points de l'échantillon tronqué appartient tous à A, et suivant donc la loi de T conditionnée par l'appartance à A. Il ne faut pas confondre censure et troncature. S'il y a troncature, une partie des individus (donc des X_i) ne sont pas observables et on n'étendie qu'un sous-échantillon (problème d'échantillonage). le biais de sélection est un particulier de troncature.

1.2.1 La troncature à gauche

Soit Z une variable aléatoire indépendante de X, on dit qu'il y a troncature à gauche lorsque X n'est observable que si $X > Z$. On observe le couple (X, Z) avec $X < Z$. Par exemple, si la durée de vie d'une population est étudiée à partir d'une cohorte tirée au sort dans cette population, seule la survie des sujets vivants à l'inclusion pourra être étudiée (il y a troncature à gauche car seuls les sujets vivants à l'inclusion pourra être étudiée (il y a troncature à gauche car seuls ayants survécu jusqu'a la date d'inclusion dans la cohorte sont observables.

1.2.2 La troncature à droite

De même il y a troncature à droiÌte lorsque X n'est observable que si $X < Z$.

1.2.3 La troncature par intervale

Quand une durée est tronquée à droite et à gauche, on dit qu'elle est tronquée par intervalle.

1.3 Définitions et Notations

Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $\{\Delta_i\}_{i\in\mathcal{Z}}$ une famille des variables aléatoires définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans un espace probabilisable $(\mathbb{E}, \xi)$. On note $(\sigma_i^j)_{i\neq j}$ dans $\mathbb{Z}\cup\{-\infty,+\infty\}$,

la tribu engendrée par $\{\Delta_k, i < k < j\}$ et par $L_2(\sigma_i^j)$ l'espace des variables aléatoires σ_i^j-mesurable et de carrée sommable.

Définition 1.3.1. Soit $\{\Delta_i\}_{i\in\mathcal{Z}}$ une famille des variables aléatoires définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans un espace probabilisable $(\mathbb{E}, \xi)$. On dit que la famille $\{\Delta_i, i \in \mathcal{Z}\}$ est α- mélangent si la suite

$$\alpha_n = \sup_{\left\{k\in\mathcal{Z}, A\in\sigma_{-\infty}^{k}, B\in\sigma_{n+k}^{+\infty}\right\}} |P(A\cap B) - P(A)P(B)|$$

tend vers 0 quand n tend vers l'infinie. La suite α_n est appelée coefficient de mélange forte.

Définition 1.3.2. On dit que une famille $\{\Delta_i, i \in Z\}$ de variables aléatoire á valeurs dans un même espace probabilisable ($\mathbb{E}$,ε) est algébriquement α-mélangeante, s'il existe deux constantes c$\in\mathbb{R}^{*+}$ et a$\in\mathbb{R}^{*+}$ telles que les coefficients de melange verifient

$$\alpha(n) \leq cn^{-a}.$$

Définition 1.3.3. On dit que une famille $\{\Delta_i, i \in Z\}$ de variables aléatoire á valeurs dans un même espace probabilisable ($\mathbb{E}$,ε) est géometriquement α-mélangeante, s'il existe deux constantes s$\in\mathbb{R}^{*+}$ et t$\in]0,1[$ telles que les coefficients de melange verifient

$$\alpha(n) \leq st^{n}.$$

Définition 1.3.4. Une fonction K de $\mathbb{R}^p$ dans $\mathbb{R}$ est dite noyau d'ordre k, $k \in \mathbb{N}^*$, si :

$$T_{i_1,\ldots,i_p}(K) = 0,\ \forall(i_1,\ldots,i_p) \in \mathbb{R}^p_* \quad \textit{vérifiant}\ \ i_j < k,\ 1 \leq j \leq p.$$

et

$$T_j(K) \neq 0,\ \forall j \leq k.$$

où

$$T_{i_1,\ldots,i_p}(K) = \int_{\mathbb{R}^p} u_1^{i_1},\ldots,u_p^{i_p} K(u_1,\ldots,u_p)du_1,\ldots,du_p.$$

et

$$T_j(K) = \int_{\mathbb{R}^p} u_j^k K(u_1,\ldots,u_p)du_1,\ldots,du_p.$$

1.4 Outils

Lemme 1.4.1. ([10]) Soit $\Delta_1,\ldots,\Delta_n$ des variables aléatoires réelles centrées, indépendantes et identiquement distribuées, telles qu'il existe deux réels positifs d et δ vérifiant :

$$|\Delta_1| \leq d \text{ et } \quad E\Delta_1^2 \leq \delta^2.$$

Alors, pour tout $\varepsilon \in]0, \frac{\delta^2}{d}[$ on a

$$p\left[n^{-1}\left|\sum_{i=1}^n \Delta_i\right| > \varepsilon\right] \leq 2e^{-\frac{n\varepsilon^2}{4\delta^2}}.$$

cette inégalité a été donnée par W.Hoeffding en (1963, ([10])). Les lemmes suivants donnent les deux version de l'inégalité de Fuk Nagave

Lemme 1.4.2. ([10]) Soit$\{\Delta_i, i\epsilon\mathbb{N}\}$ une suite des variables aléatoires réelles α-mélangeante, de coefficient de mélange α_n vérifiant :

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \qquad \alpha(n) \leq cn^{-a}$$

et si $\forall i, \quad \|\Delta_i\|_\infty < \infty$, alors, pour tout $\varepsilon > 0$ et $r > 0$, on a

$$P\left[|\sum_{k=1}^n \Delta_k| > 4\varepsilon\right] \leq \left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{2r}{\varepsilon}\right)^{a+1}. \tag{1.1}$$

Pour calculer l'expression de S_n^2, définie dans le lemme précédent, on utilise le lemme suivant :

Lemme 1.4.3. ([10]) Soit$\{\Delta_i, i\varepsilon\mathbb{N}\}$ une suite des variables aléatoires réelles α-mélangeante, de coefficient de mélange α_n vérifiant :

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \qquad \alpha(n) \leq cn^{-a}$$

et si $\|\Delta_i\|_\infty < \infty, \forall i$, alors, pour tout $\varepsilon > 0$ et $r > 0$

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n |cov(\Delta_i, \Delta_j)|.$$

pour calculer l'expression de S_n^2, définie dans le lemme précédent, on utilise le lemme suivant :

Lemme 1.4.4. ([10]) Soit$\{\Delta_i, i\varepsilon\mathbb{N}\}$ une suite des variables aléatoires réelle α-mélangeante, de coefficient des mélange α_n, telle que $\|\Delta_i\|_\infty < \infty, \forall i$. On a pour tout $i \neq j$:

$$|cov(\Delta_i, \Delta_j)| \leq 4\|\Delta_i\|_\infty \|\Delta_j\|_\infty \alpha_{|i-j|}.$$

Lemme 1.4.5. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur le même espace probabilité, indépendantes. Supposons que, pour $n \geq 1$, X_n ait une espérance finie μ_n et un écart-type fini σ_n, et posons $S_n^2 = \sum_{i=1}^{n} \sigma_i^2$ et $Z_n = \frac{1}{S_n}\sum_{i=1}^{n}(X_i - \mu_i)$. Si, pour tout $\epsilon > 0$, $\lim_{n\longrightarrow\infty} \frac{1}{S_n^2}\sum_{i=1}^{n} \mathbb{E}\left((X_i-\mu_i)^2 1_{|X_i-\mu_i|>\epsilon S_n}\right) = 0$ alors la loi de Z_n converge vers la loi normale centrée réduite $\mathcal{N}(0,1)$.

Lemme 1.4.6. ([7], 2001) Soit $\mathcal{F}$ une fonction mesurable borné et pour quelque soit σ^2 et u tel que

$$\sigma^2 \geq \sup_{f\in\mathcal{F}} Var_f, \qquad u \geq \sup_{f\in\mathcal{F}} ||f||_\infty, \quad \text{et} \quad 0 < \sigma \leq u$$

En suite, ils y a trois constantes c, k dépend seulement par la méthode de validation vroisée par A et v dans la classe $\mathcal{F}$, tel que

$$Pr\left\{||\sum_{i=1}^{n} f(\zeta_i) - \mathbb{E}f(\zeta_i)||_f > t\right) \quad \leq \quad k\exp\left(-\frac{1}{k}\frac{t}{u}\log\left(1 + \frac{tu}{k(\sqrt{n}\sigma + u\sqrt{\log\left(\frac{Au}{\sigma}\right)}}\right)\right)$$

Tel que

$$t \geq \left(u\log\frac{Au}{\sigma} + \sqrt{n}\sigma\sqrt{\log\frac{Au}{\sigma}}\right)$$

Chapitre 2

L'estimation non paramétrique du quantile conditionnel : Cas Réel

Dans ce chapitre, nous étudions l'estimateur à noyau du quantile conditionnel dans le cas des observations sont indépendant et identiquement distribuée. Ce chapitre est divisé en deux sections sections. Dans la première section nous établissons la convergence presque complète de notre estimateur. La convergence en normalité asymptotique se trouve dans la deuxième setcion.

2.1 Modèle

Soit (X, Y) un couple des variable aléatoire à valeurs dans $\mathbb{R} \times \mathbb{R}$. Pour tout $x \in \mathbb{R}$, on désigne par F^X la fonction de répartition conditionnelle de$Y$ sachant $X = x$ (resp.$F^{x(j)}$) la dérivée d'ordre j de F^x. Pour $\alpha \in]0, 1[$, on définit le quatile conditionnel d'ordre α, notée t_α par

$$F^x(t_\alpha) = \alpha$$

pour assurée l'unicité des quantiles conditionnels, on suppose que la fonction de répartition conditionnelle F^x est strictement croissente. Etant donnée $(X_1, Y_1), ..., (X_1, Y_1)$ une suite des observations de (X, Y), l'estimateur à noyau de la fonction de répartition F^x notée $\widehat{F}^x$, est défini par :

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^n K\Big(h_K^{-1}(x - X_i)\Big) H\Big(h_H^{-1}(y - Y_i)\Big)}{\sum_{i=1}^n K\Big(h_K^{-1}(x - X_i)\Big)}, \qquad \forall y \in \mathbb{R},$$

avec la convention $\frac{0}{0} = 0$, K est un noyau dans $\mathbb{R}$, H est un fonction de répartition et $h_K = h_{K,n}(resp.h_H = h_{H,n})$ est un suite des nombers réels positifs. L'estimateur naturel de t_α est :

$$\widehat{F}^x(\widehat{t_\alpha}) = \alpha$$

.

2.2 Cas i.i.d

On suppose que les observations sont indépendants et identiquement distribuées.

2.2.1 Convergence presque complete

Pour établir la convergence presque complète, on a besoin des hypothèses suivants :

2.2.2 Hypothèses

- (H1) la densité conditionnelle est k-fois continument différentiables relativement à (x, y).
- (H2) La densité f de la variable explicative est strictement positive au point$x$.
- (H3) Le noyau K est supposé d'ordre K ,borné ,positif et à support compact.
- (H4) Le paramètre de lissage est telque :

$$\lim_{n\to\infty} h_K = 0, \quad \exists\alpha > 0, \quad \lim_{n\to\infty} n^\alpha h_H = \infty \quad \text{et} \quad \lim_{n\to\infty} \frac{nh_K}{\log n} = \infty$$

.
- (H5) la fonction H est strictement croissante d'ordre k tel que

$$\int |t|^k H^{(1)}(t)dt < \infty$$

.

Remarque 2.2.2.1.
- Les hypothèses $(H1) - (H2)$ sont des conditions de régularité de l'espace fonctionnelle de F^x
- les hypothèses $(H3) - (H5)$ sont des hypothèses techniques nécessaires pour le calcule.

Théorème 2.2.1. Sous les hypothèses $(H1) - (H5)$, on a :

$$\widehat{t_\alpha} - t_\alpha = O(h_K^k) + O(h_H^k) + O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad \text{p.co.}$$

Démonstration des résultats téchniques
On considère le développement du Taylor d'ordre 1 pour l'estimateur $\widehat{F^x}$ au point $\hat{t}_\alpha$, ceci est possible car cette fonction est de classe C^1, alors :

$$\widehat{F}^x(\widehat{t_\alpha}) - \widehat{F}^x(t_\alpha) = (\widehat{t_\alpha} - t_\alpha)\widehat{F}^{x(1)}(t_\alpha^*)$$

où t_α^* est un point entre t_α et $\widehat{t_\alpha}$.

$$\left|\widehat{t_\alpha} - t_\alpha\right| = \frac{1}{\left|\widehat{F}^{x(1)}(t_\alpha^*)\right|}\left[\left|\widehat{F}^x(t_\alpha) - F^x(t_\alpha)\right|\right]$$

Il suffit de montrer que

$$\widehat{F}^x(t_\alpha) - F^x(t_\alpha) \quad = \quad O(h_K^k) + O(h_H^k) + O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad p.co. \tag{2.1}$$

et

$$\exists \delta > 0 \quad \text{tel que} \quad \sum_n \mathbb{P}\left[\left|\widehat{F}^{x(1)}(t^*_\alpha)\right| < \delta\right] < \infty \tag{2.2}$$

La preuve de l'équation (2.1) repose sur la décomposition suivante :

$$\begin{aligned}\widehat{F}^x(t_\alpha) - F^x(t_\alpha) &= \frac{1}{\widehat{f}_n(x)}\left[\Big(\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha)\Big) - \Big(F^x(t_\alpha)f(x) - \mathbb{E}\widehat{g}_n(x,t_\alpha)\Big)\right] \\ &+ \frac{F^x(t_\alpha)}{\widehat{f}_n(x)}\Big(f(x) - \widehat{f}_n(x)\Big)\end{aligned}$$

Où

$$\widehat{g}_n(x,t_\alpha) = \frac{1}{nh_K}\sum_{i=1}^n K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right)$$

et

$$\widehat{f}_n(x) = \frac{1}{nh_K}\sum_{i=1}^n K\left(\frac{x-X_i}{h_K}\right),$$

Ainsi, le théorème (3.2.1) est une conséquence directe des lemmes suivants :

Lemme 2.2.2.1. Sous les hypothèses $(H1)-(H5)$, on a :

$$\mathbb{E}\widehat{g}_n(x,t_\alpha) - F^x(t_\alpha)f(x) = O(h_K^k) + O(h_H^k), \qquad \text{p.co.}$$

Lemme 2.2.2.2. Sous les hypothèses $(H1)-(H5)$, on a :

$$\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha) = O\left(\sqrt{\frac{\log n}{nh_K}}\right) \quad \text{P.co.}$$

Lemme 2.2.2.3. Sous les conditions $(H1)-(H4)$, on a,

$$f(x) - \widehat{f}_n(x) \;=\; O(h_K^k) + O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad \text{p.co.} \tag{2.3}$$

et

$$\exists \delta > 0 \quad , \quad \text{tel que,} \qquad \sum_n \mathbb{P}\left(\left|\widehat{F}_n(x)\right| < \infty\right). \tag{2.4}$$

Corollaire 2.1. Sous les hypothèses $(H1)-(H5)$, on a :

$$\exists \tau > 0, \quad \text{tel que} \quad \sum_n \mathbb{P}\left[\left|\widehat{F}^{x(1)}(t^*_\alpha)\right| < \tau\right] < \infty.$$

Démonstration du lemme (2.2.2.1)
Les observations (X_i, Y_i) étant équidistribuées, alors,

$$\begin{aligned}
\mathbb{E}\widehat{g}_n(t_\alpha) &= \frac{1}{nh_K}\mathbb{E}\Bigg[\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_K}\right)H\left(\left(\frac{t_\alpha - Y_i}{h_H}\right)\right)\Bigg] \\
&= \frac{1}{h_K}\mathbb{E}\Bigg[K\left(\frac{x-X}{h_K}\right)\mathbb{E}H\left(\left(\frac{t_\alpha - Y}{h_H}\right)/X\right)\Bigg],
\end{aligned}$$

Une intégration par partie, ensuite un changement de variable $h_H^{-1}(x-u) = t$

$$\begin{aligned}
\mathbb{E}H\left(\left(\frac{t_\alpha - Y}{h_H}\right)/X\right) &= \int H^{(1)}\left(\frac{t_\alpha - u}{h_H}\right)F^x(z)dz \\
&= \int H^1(t)F^X(t_\alpha - h_K t)dt
\end{aligned}$$

Donc,

$$\mathbb{E}\widehat{g}_n(t_\alpha) = \frac{1}{h_K}\int\int K\left(\frac{x-z}{h_K}\right)H^1(t)F^X(t_\alpha - h_K t)f(u)dudt$$

Le changement de variable $t = h_K^{-1}(t_\alpha - u)$ donne

$$\mathbb{E}\widehat{g}_n(t_\alpha) = \frac{1}{h_K}\int\int K\left(z\right)H^1(t)F^{x-h_K t}(t_\alpha - h_K t)f(x-h_K z)dzdt$$

On pose $g(x, t_\alpha) = F^x(t_\alpha)f(x)$ tel que g est une fonction définie sur $\mathbb{R}^2$.

$$\mathbb{E}\widehat{g}_n(t_\alpha) = \frac{1}{h_K}\int\int K\left(z\right)H^1(t)(g(x-h_K z, t_\alpha - h_K t) - g(x, t_\alpha))dzdt \tag{2.5}$$

On utilise le développement de Taylor de g au voisinage (x, t_α) qui est de classe C^k, alors

$$\begin{aligned}
g(x-h_K z, t_\alpha - h_K t) - g(x, t_\alpha) &= -\sum_{i_1+i_2=j=1}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}((h_K z_1)^{i_1}(h_H t_\alpha)^{i_2}) \\
&+ \frac{1}{2}\sum_{i_1+i_2=j=2}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}((h_K z_1)^{i_1}(h_H t)^{i}) + \cdots \\
&+ \frac{(-1)^k}{k!}\sum_{i_1+i_2=j=k}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}((h_K z_1)^{i_1}(h_H t)^{i}) \\
&+ 0(h_K^k) + 0(h_H^k) \\
&= \sum_{j=1}^{n}\frac{(-1)^j}{j!}\sum_{i_1+i_2=j}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}((h_K z_1)^{i_1}(h_H t)^{i_2}) \\
&+ 0(h_K^k) + 0(h_H^k)
\end{aligned}$$

On remplace dans l'équation (4.2.1),

$$\begin{aligned}
\mathbb{E}\widehat{g}_n(x,t_\alpha) - F^x(t_\alpha)f(x) &= \int\int K(z)H^{(1)}(t)\sum_{j=1}^{k}\frac{(-1)^j}{j!}\sum_{i_1+i_2=j}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}((h_K z_1)^{i_1}(h_H t)^{i_2})dzdt \\
&+ O(h_K^k) + O(h_H^k) \\
&= \sum_{j=1}^{k}\frac{(-1)^j}{j!}\sum_{i_1+i_2=j}\frac{\partial^j g(x,t_\alpha)}{\partial x_1^{i_1}\partial t_\alpha^{i_2}}\int K(z)(h_K z)^{i_1}dz\int H^{(1)}(t)(h_H t)^{i_2}dt \\
&+ O(h_K^k) + O(h_H^k).
\end{aligned}$$

Finallement, il suffit d'utiliser le fait que le noyau K (resp H) est d'ordre k pour achever le lemme.

$$\mathbb{E}\widehat{g}_n(x,t_\alpha) - F^x(t_\alpha)f(x) = O(h_K^k) + O(h_H^k)$$

Démonstration du lemme (2.2.2.2)
On a,

$$\begin{aligned}
\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha) &= \frac{1}{nh_K}\sum_{i=1}^{n}K\Big(h_K^{-1}(x-X_i)\Big)H\Big(h_H^{-1}(t_\alpha - Y_i)\Big) \\
&- \mathbb{E}\left[\frac{1}{nh_K}\sum_{i=1}^{n}K\Big(h_K^{-1}(x-X_i)\Big)H\Big(h_H^{-1}(t_\alpha - Y_i)\Big)\right]
\end{aligned}$$

On pose :

$$\triangle_i = \frac{1}{h_K}\left[K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right) - \mathbb{E}K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right)\right].$$

Pour appliquer l'intégalité de Hoeffding (lemme 4.2.1, [10]) aux variables $\triangle_i$, il est nécessaire de majorer les deux termes : $|\Delta_i|$ et $\mathbb{E}\triangle_i^2$.
En effet, le noyau K est borné et $0 < H < 1$ donne

$$|\triangle_i| \leq \frac{C}{h_K}$$

Pour le moment d'ordre deux de $\triangle_i$, on pose

$$\begin{aligned}
\tau_i &= \frac{1}{h_K}\left[K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right)\right] \\
&< \frac{c}{h_K}
\end{aligned}$$

Donc

$$\mathbb{E}\tau_i^2 =\leq \frac{C}{h_K^2}\int K^2\left(\frac{x-u}{h_K}\right)f(u)du.$$

On considère le changement de variable $z = \dfrac{x-u}{h_K}$, alors,

$$du = -h_K^p dz,$$

et
$$\mathbb{E}\tau_i^2 \leq \frac{C}{h_K} \int f(x - zh_K)K^2(z)dz.$$
Puisque la densité f est continue et le nouau K est à support compact, alors il existe une constante C telle que :
$$\mathbb{E}\tau_i^2 \leq \frac{C}{h_K}.$$
De la définition de $\triangle_i$ et τ_i on constate que le moment d'ordre deux de $\triangle_i$ est égale à la variance de τ_i, ainsi, il est évident que :
$$\mathbb{E}\triangle_i^2 = var(\tau_i) \leq \mathbb{E}\tau_i^2 \leq \frac{C}{h_K}.$$
Le lemme (1.1) permet alors d'écrire que , pour tout ϵ, on a :
$$\mathbb{P}\left[\frac{1}{n}\left|\sum_{j=1}^{n} \triangle_i\right| > \epsilon\right] = \mathbb{P}\left[\left|\widehat{g}_n(x, t_\alpha) - \mathbb{E}\widehat{g}_n(x, t_\alpha)\right| > \epsilon\right] \leq 2\exp\left(\frac{-n\epsilon^2 C}{4h_K}\right).$$
En prenant :
$$\epsilon = \epsilon_0\sqrt{\frac{\log n}{nh_K}}.$$
On arrive pour tout $\epsilon_0 > 0$,
$$\mathbb{P}\left[\left|\widehat{g}_n(x, t_\alpha) - \mathbb{E}\widehat{g}_n(x, t_\alpha)\right| > \epsilon_0\sqrt{\frac{\log n}{nh_K}}\right] \leq 2n^{-C\epsilon_0^2}.$$
Donc,
$$\sum_n \mathbb{P}\left[\left|\widehat{g}_n(x, t_\alpha) - \mathbb{E}\widehat{g}_n(x, t_\alpha)\right| > \epsilon_0\sqrt{\frac{\log n}{nh_K}}\right] \leq \sum 2n^{-C\epsilon_0^2}.$$
Il suffit de choisir $\epsilon_0 > \frac{1}{\sqrt{C}}$ pour que la série converge.

Démonstration du lemme (2.2.2.3) La démonstration est similaire à celle qui procéde, il suffit de H est égale à 1.

Démonstration du corollaire (2.1)

Le lemme (2.2.2.3) entraine en particulier la convergence persque complète de $\widehat{f}_n(x)$ vers $f(x)$. Ainsi, pour tout $\epsilon > 0$,on a :
$$\sum_{n=1}^{\infty} \mathbb{P}\left[\left|\widehat{f}_n(x) - f(x)\right| > \epsilon\right] < \infty.$$
On remarque que,
$$\widehat{f}_n(x) \leq \frac{f(x)}{2} \Rightarrow \left|\widehat{f}_n(x) - f(x)\right| \geq \frac{f(x)}{2}.$$
On peut écrire alors,
$$\mathbb{P}\left[\left|\widehat{f}_n(x)\right| \leq \frac{f(x)}{2}\right] \leq \mathbb{P}\left[\left|\widehat{f}_n(x) - f(x)\right| > \frac{f(x)}{2}\right].$$

Il suffit maintenant de prendre $\delta = \dfrac{f(x)}{2}$.

Démonstration du corollaire
Puisque la fonction H est strictement croissante, l'estimateur $\widehat{F}^x$ est inversible et d'inverse continue, donc,

$$\forall \epsilon > 0, \quad \exists \delta(\epsilon) > 0, \quad \forall y, \quad \left|\widehat{F}^x(y) - \widehat{F}^x(t_\alpha)\right| \leq \delta(\epsilon) \Rightarrow |y - t_\alpha| \leq \epsilon,$$

alors,

$$\forall \epsilon > 0, \quad \exists \delta(\epsilon) > 0, \quad \mathbb{P}\left(\left|\widehat{t}_\alpha - t_\alpha\right| > \epsilon\right) \leq \mathbb{P}\left(\left|\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha)\right| \geq \delta(\epsilon)\right).$$

Ce qui nous permet de déduire la convergence persque complète de t^*_α vers t_α. Finallement, en suivant les même lignes de la décomposition du théorème précédente, en remplaçant H par $H^{(1)}(h_H^{-1})$, pour obtenir le résultat.

2.2.3 Normalité asymptotique

On s'intéresse dans ce paragraphe à la normalité asymptotique, pour laquelle, on garde les hypothèses de la partie précédente et on ajoute la condition suivante :
On fixe un point $x \in \mathbb{R}$,

- $(H6)$ il existe deux constontes c_1, c_n tq $\forall y \in \mathbb{R}, \forall j \geq k, 0 < c_1 \leq \dfrac{\partial^j \mathbb{F}^x(y)}{\partial y^j} \leq \varepsilon$

2.2.4 Résultats

Théorème 2.2.2. Sous les hypothèses $(H1) - (H5)$ et $(H6)$, on a :

$$\sqrt{nh_K}\left[\widehat{t_\alpha} - t_\alpha\right] \to \mathcal{N}\left(0, \sigma(x)\right)$$

où

$$\sigma^2(x) = \frac{\alpha(\alpha - 1)}{f(x)F^{x(1)}(t_\alpha)}$$

Démonstration
Le développement de taylor d'ordre 1 de l'estimateur $\widehat{F}^x$ au point $\widehat{t}_\alpha$, ceci est possible car la fonction est de classe C^1, alors :

$$\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha) = (\widehat{t}_\alpha - t_\alpha)\widehat{F}^{x(1)}(t^*_\alpha)$$

avec $t^*_\alpha \in (t_\alpha, \widehat{t}_\alpha)$, Alors

$$\left|\widehat{t}_\alpha - t_\alpha\right| = \frac{1}{|\widehat{F}^{x(1)}(t^*_\alpha)|}\left[|\widehat{F}^x(t_\alpha) - F^x(t_\alpha)|\right].$$

Ainsi, pour montrer le Théorème (2.2.2), il suffit de prouver que

Lemme 2.2.4.1. Sous les hypothèses $(H1)-(H5)$ et $(H6)$, on a :

$$\sqrt{nh_K}\left\{(\widehat{F}^x(t_\alpha)-F^x(t_\alpha)\right\}) \to \mathcal{N}(0,\sigma(x)), \qquad \text{p.co.}$$

Corollaire 2.2. Sous les hypothèses $(H1)-(H5)$,on a :

$$\exists \tau>0, \qquad \text{tel que} \quad \sum_n \mathbb{P}[|\widehat{F}^{x(1)}(t_\alpha^*)|<\tau\Big] < \infty.$$

Démonstration du lemme 2.2.4.1
La preuve est basé sur la décomposition suivante :

$$\sqrt{nh_K}(\widehat{F}^x(\widehat{t}_\alpha)-\widehat{F}^x(t_\alpha)=\sqrt{nh_k}((\widehat{t}_\alpha-t_\alpha)\widehat{F}^{x(1)}(t_\alpha^*), \quad \text{avec} \quad t_\alpha^*\in(t_\alpha,\widehat{t}_\alpha)$$

Dans ce qui suit, nous avons

$$\widehat{F}^x(t_\alpha)=\frac{\widehat{g}_n(x,t_\alpha)}{\widehat{f}_n(x)},$$

Alors

$$\sqrt{nh_K}\left(\widehat{g}_n(\widehat{t}_\alpha)-g_n(t_\alpha)\right)=\sqrt{nh_K}((\widehat{t}_\alpha-t_\alpha)\widehat{g}^{x(1)}(t_\alpha^*)$$

donc,

$$\sqrt{nh_K}\left(\alpha f_n(x)-g_n(t_\alpha)\right) \quad = \quad \sqrt{nh_K}((\widehat{t}_\alpha-t_\alpha)\widehat{g}^{x(1)}(t_\alpha^*) \quad \text{avec} \quad t_\alpha^*\in(t_\alpha,\widehat{t}_\alpha)$$

Alors

$$\sqrt{nh_K}\left(\widehat{t}_\alpha-t_\alpha\right)=\sqrt{nh_K}\frac{\alpha f_n(x)-g_n(t_\alpha)}{\widehat{g}^{x(1)}(t_\alpha^*)}$$

Ainsi, ce lemme est une conséquence directe des lemmes suivants :

Lemme 2.2.4.2. Sous les hypothèses $(H1)-(H5)$ et $(H6)$, on a :

$$\sqrt{nh_K}\left(\alpha f_n(n)-g_n(t_\alpha)-\mathbb{E}\Big(\alpha f_n(x)-g_n(t_\alpha)\Big)\right)\to\mathcal{N}(0,\sqrt{\alpha(\alpha-1)}), \qquad p.co.$$

Corollaire 2.3. Sous les hypothèses $(H1)-(H5)$, on a :

$$\exists \tau>0, \qquad \text{tel que} \qquad \sum_n \mathbb{P}[|\widehat{g}^{x(1)}(t_\alpha^*)|<\tau\Big] < \infty.$$

Démonstration du lemme 2.2.4.2 La preuve repose sur le Théorème de centrale limite (4.2.1). En effet, il suffit de montrer qu'il existe $\delta>0$ vérifiant :

$$\frac{\sum_{i=1}^n \mathbb{E}\left|\sqrt{nh}\frac{1}{n}|L_i(x)-\mathbb{E}L_i(x)|\right|^{2+\delta}}{\left(var\left(\sqrt{nh}\frac{1}{n}\sum_{i=1}^n L_i(x)\right)\right)^{1+\frac{\delta}{2}}}\longrightarrow 0$$

Où

$$L_i(x) = \frac{1}{h_K}\Big[K_iH_i - \alpha K_i\Big]$$

Pour le terme de variance, nous avons

$$\begin{aligned}
\frac{1}{n}var\sum_{i=1}^{n} L_i(n) &= var(g_n(t_\alpha) - \alpha f_n(x)) \\
&= \frac{1}{(nh)^2}\sum_{i=1}^{n} var(K_iH_i(t_\alpha) - \alpha K_i) \\
&= \frac{1}{nh^2}var(K_1H_1(t_\alpha) - \alpha K_1) \\
&= \frac{1}{nh^2}\mathbb{E}\left(K_1H_1(t_\alpha) - \alpha K_i\right)^2 - \frac{1}{nh^2}\mathbb{E}^2\left(K_1H_1(t_\alpha) - \alpha K_i\right)
\end{aligned}$$

Alors,

$$nh_K var\left(g_n(t_\alpha) - \alpha f_n(x)\right) = \frac{1}{h}\mathbb{E}K_1^2(H_1(t_\alpha) - \alpha)^2 - \left(\mathbb{E}K_1(H_1(t_\alpha) - \alpha)\right)^2 \tag{2.6}$$

Puisque

$$\begin{aligned}
\mathbb{E}\left(H_1(t_\alpha)/x\right) &= \int H_1\left(\frac{t_\alpha - z}{h}\right)dF^x(Z) \\
&= \frac{1}{h_K}\int H^{(1)}\left(\frac{t_\alpha - z}{h_H}\right)F^x(z) \\
&= \int H^{(1)}(y)\Big[F^x(t_\alpha - h_H(y)\Big]dy \\
&= F^x(t_\alpha) + O(h_H)
\end{aligned}$$

donc

$$\mathbb{E}\left(\frac{1}{h_K}K_1(H - \alpha)\right) = \frac{1}{h_K}\int K\left(h_K^{-1}(x - z)\right)[F^z(t_\alpha) - F^x(t_\alpha)]f(z)dt$$

Par le changement de variable $t = (h_K^{-1}(x - z))$, on peut écrire

$$\mathbb{E}\left(K_1H_1(t_\alpha)\right) = \int K(t)\left[F^{x - h_H t}(t_\alpha) - F^x(t_\alpha)\right]f(x - h_H t)dt \to 0$$

Passons maintenant au premier terme du côté droit de (2.6), nous avons

$$\begin{aligned}
\frac{1}{h_K}\mathbb{E}\left(K_1(H_1(t_\alpha) - \alpha)\right)^2 &= \frac{1}{h_K}\mathbb{E}K_1^2H_1^2(t_\alpha) + \alpha^2\frac{1}{h_K}\mathbb{E}K_1^2 - 2\alpha\frac{1}{h_K}\mathbb{E}K_1^2H_1(t_\alpha) \\
&= \frac{1}{h_K}\mathbb{E}(K_1^2(H_1^2(t_\alpha) - \alpha) - \frac{2}{h_K}\alpha\mathbb{E}(K_1^2(H_1(t_\alpha) - \alpha) \\
&+ \frac{1}{h_K}\mathbb{E}K_1^2(\alpha - \alpha^2)
\end{aligned}$$

puisque $2\int H(t)H^{'}(t)dt = 1$ et par intégration par partie, on obtient

$$\lim_{n\to\infty}\frac{1}{h_K}\mathbb{E}((H_i(t_\alpha)-\alpha)^2 = \alpha-\alpha^2$$

Donc

$$nh_K var((\alpha f_n(n) - g_n(t_\alpha) \to \alpha(\alpha-1)$$

On a évaluer à la limite de nomirateur, pour laquelle on utilise l'inégalité C_r.

$$\begin{aligned}(n^{-1}h_K)^{1+\frac{\delta}{2}}\sum_{i=1}^{n}(|\mathbb{E}L_i-\mathbb{E}L_i|^{2+\delta} &\leq 2^{1+\delta}(n^{-1}h_K)^{1+\frac{\delta}{2}}|\sum_{i=1}^{n}\mathbb{E}L_i^{2+\delta}| \\ &+ 2^{1+\delta}(n^{-1}h_K)^{1+\frac{\delta}{2}}|\sum_{i=1}^{n}(\mathbb{E}L_i)^{2+\delta}|\end{aligned}$$

Pour le premier terme, on a

$$\begin{aligned}2^{1+\delta}(n^{-1}h_K)^{1+\frac{\delta}{2}}\left|\sum_{i=1}^{n}\mathbb{E}L_i^{2+\delta}\right| &= 2^{1+\delta}n^{-\delta/2}h_K^{1+\delta/2}\mathbb{E}(K^{2+\delta}(H_i(t_\alpha)-\alpha)^{2+\delta}) \\ &\leq Cn^{-\delta/2}h_K^{2+\delta/2}\end{aligned}$$

Le deuxième terme,

$$\begin{aligned}2^{1+\delta}(n^{-1}h_K)^{1+\frac{\delta}{2}}|\sum_{i=1}^{n}(\mathbb{E}L_i)^{2+\delta}| &\leq 2^{1+\delta}n^{-\delta/2}h_K^{1+\delta/2}(\mathbb{E}|K(H_i(t_\alpha)-\alpha)|)^{2+\delta} \\ &\leq Cn^{-\delta/2}h_K^{3+3\delta/2}\longrightarrow 0\end{aligned}$$

Il reste de montrer que,

$$\mathbb{E}(\alpha f_n(n) - g_n(t_\alpha)) \quad\to\quad 0 \tag{2.7}$$

Alors :

$$\begin{aligned}\mathbb{E}(\alpha f_n(x) - g_n(t_\alpha)) &= \frac{1}{h_K}\mathbb{E}(\alpha K_1(x) - \mathbb{E}K_1(x)H_1(t_\alpha)) \\ &= \frac{1}{h_K}\mathbb{E}K_1(x)(\alpha - \mathbb{E}(H_i(t_\alpha)/x)) \longrightarrow 0\end{aligned}$$

Démonstration du corollaire (2.3) D'aprés les hypothèses (H1), (H3)-(H5) et (H6), on a

$$\begin{aligned}g_n^{(1)}(t_\alpha^*) - F^{x(1)}f(x) &= f_n(x)[\widehat{F}^{x(1)}(t_\alpha^*) - F^{x(1)}(t_\alpha^*)] \\ &+ F^{x(1)}(t_\alpha^*)[f_n(x)-f(x)] + f(x)[F^{x(1)}(t_\alpha^*) - F^{x(1)}(t_\alpha)]\end{aligned}$$

Puisque la fonction H est strictement croissante, l'estimateur $\widehat{F}^x$ est inversible et d'inverse continue, donc,

$$\forall\epsilon>0,\quad \exists\delta(\epsilon)>0,\quad \forall y, |\widehat{F}^x(y) - \widehat{F}^x(t_\alpha)| \leq \delta(\epsilon) \Rightarrow |y-t_\alpha|\leq\epsilon,$$

Alors,

$$\begin{aligned}\forall \epsilon > 0, \quad \exists \delta(\epsilon) > 0, \quad \mathbb{P}(|\widehat{t}_\alpha - t_\alpha| > \epsilon) &\leq \mathbb{P}(|\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha)| \geq \delta(\epsilon)) \\ &\leq \mathbb{P}(|F^x(t_\alpha) - \widehat{F}^x(t_\alpha)| \geq \delta(\epsilon)\end{aligned}$$

On déduit la convergence persque complète de t^*_α vers t_α. Comme t^*_α est entre ces deux derniers, on obtient le résultat. Finallement, on utilise la continuité de $F^{x(1)}(.)$ pour obtenir le résultat.

2.3 Cas de Mélange

Le but de ce chapitre est la généralisation du résultat donné dans le cas précédent à des observations mélangeantes. On établit, sous des conditions plus restrictives la convergence presque complète.

2.3.1 Convergence presque complète

Dans cette section, on étudie la convergence presque complète de l'estimateur à noyau du quantile conditionnel.

2.3.2 Hypothèses

On considère des observations $\alpha-$mélangeant et on garde les mêmes notations, ainsi, les mêmes hypothèses du cas précédent et on ajoute les conditions ci-dessous permettant de trouver la même vitesse de convergence que le cas i.i.d.

- (H6) La suite $(X_i, Y_i), i = 1, \cdots, n$ est $\alpha-$mélangeante
- (H7) Le couple des variables aléatoires (X_i, X_j) admet une densité continue notée f_{ij}, pour tout $i \neq j$
- (H8) Il existe deux constantes $C \in \mathbb{R}^{*+}$ et $a \in \mathbb{R}^{*+}$ telles que

$$\alpha(n) \leq Cn^{-a}$$

- (H9) $\lim_{n \longrightarrow \infty} h_H = 0$
- (H10) $\lim_{n \longrightarrow \infty} h_K = 0, \ \lim_{n \longrightarrow \infty} \dfrac{\log n}{nh_K} = 0$

$$\exists \beta_2 > 0, c_1 > 0, c_2 > 0, c_2 n^{\frac{3-a}{a+1}+\beta_2} \leq h_K \leq c_2 n^{\frac{1}{1-a}-\beta_1}$$

Remarque 2.3.2.1. Les hypothèses $(H6) - (H10)$ sont ajoutés pour éviter l'expression de covaraince dans la vitesse de convergence. Autrement dit, on peut démontrer la convergence presque complète sans ces hypothèses. Cependant, la vitesse de convegence sera donnée en fonction de covariance des observations et elle sera lente par rapport à la vitesse du cas indépendant. Ainsi nous établissons la convergence presque complète avec la même précision, mais, sous des conditions un peux plus fort que le cas i.i.d.

2.3.3 Résultats

Théorème 2.3.1. Sous les hypothèses $(H1)-(H9)$,on a :

$$\widehat{t}_\alpha - t_\alpha = O(h_K^k) + O(h_H^k) + O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad \text{p.co.}$$

Démonstration : Le caractère de l'indépendance des observations n'intervient pas dans la partie biais. Autrement dit, les vitesses de convergence de la partie biais du cas précédent seront les mêmes dans le cas de mélange. Cependant, la partie dispersion est basé sur les deux lemmes suivants :

Lemme 2.3.3.1. Sous les hypothèses $(H1)-(H9)$, on a :

$$\mathbb{E}\widehat{g}_n(x,t_\alpha) - \widehat{g}_n(x,t_\alpha) = O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad \text{p.co.}$$

Lemme 2.3.3.2. Sous les hypothèses $(H1)-(H9)$, on a :

$$\widehat{f}_n(x) - \mathbb{E}\widehat{f}_n(x) = O\left(\sqrt{\frac{\log n}{nh_K}}\right), \qquad \text{p.co.}$$

Démonstration du lemme 2.3.3.1
On pose

$$\Delta_i = K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right) - \mathbb{E}K\left(\frac{x-X_i}{h_K}\right)H\left(\frac{t_\alpha - Y_i}{h_H}\right).$$

Alors

$$\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha) = \frac{1}{nh_K}\sum_{i=1}^{n}\Delta_i$$

Donc, nous appliquons l'inégalté Fuk-Nagaev (Rio [10], 2005) pour obtenir pour tout $r>0$ et $\epsilon>0$,

$$\mathbb{P}\left(\left|\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha)\right| > 4\epsilon\right) = \mathbb{P}\left(\frac{1}{nh_K}\left|\sum_{i=1}^{n}\Delta_i\right| > 4\epsilon\right)$$

$$\leq 4\left(\underbrace{1+\frac{n^2h_K\epsilon^2}{16rS_n^2}}_{Q_1}\right)^{-r/2} + 2nCr^{-1}\left(\underbrace{\frac{8r}{nh_K\epsilon}}_{Q_1}\right)^{a+1}$$

Tel que

$$S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(\Delta_i,\Delta_j)$$

On vérifie aisément que

$$S_n^2 = S_n^{2*} + nvar(\Delta_1).$$

Par ailleurs,

$$\begin{aligned} cov(\Delta_i,\Delta_j) &= \mathbb{E}\Bigg[\Bigg(K\Bigg(\frac{x-X_i}{h_K}\Bigg)H\Bigg(\frac{t_\alpha-Y_i}{h_H}\Bigg)\Bigg)\Bigg(K\Bigg(\frac{x-X_j}{h_K}\Bigg)H\Bigg(\frac{t_\alpha-Y_j}{h_H}\Bigg)\Bigg)\Bigg] \\ &- \mathbb{E}\Bigg[\Bigg(K\Bigg(\frac{x-X_i}{h_K}\Bigg)H\Bigg(\frac{t_\alpha-Y_i}{h_H}\Bigg)\Bigg] \\ &\leq \mathbb{E}\Bigg[K\Bigg(\frac{x-X_i}{h_K}\Bigg)K\Bigg(\frac{x-X_j}{h_K}\Bigg)\Bigg]-\mathbb{E}\Bigg[K\Bigg(\frac{x-X_i}{h_K}\Bigg)\Bigg]\mathbb{E}\Bigg[K\Bigg(\frac{x-X_j}{h_K}\Bigg)\Bigg] \\ &\leq C\int\int K\Bigg(\frac{x-u}{h_K}\Bigg)K\Bigg(\frac{x-v}{h_H}\Bigg)\Big[f_{ij}(u,v)-f(u)f(v)\Big]dudv. \end{aligned}$$

On prend le changement des variable usuelles $\dfrac{x-u}{h_K}=z$ et $\dfrac{x-v}{h_K}=t$, d'où

$$|cov(\Delta_i,\Delta_j)| \leq Ch_K^2\int\int |K(z)K(t)\left[f_{ij}(x-h_Kt,x-h_Kt)-f(x-h_Kz)f(x-h_Kt)\right]|\,dzdt.$$

Pour n assez grand, on trouver une constante C telle que :

$$|cov(\Delta_i,\Delta_j)| \leq Ch_K^2\int\int \big|K(z)K(t)\left[f_{ij}(x,x)-f^2(x)\right]\big|\,dzdt.$$

Puisque les fonction K, f et f_{ij} sont bornées, alors,

$$|cov(\Delta_i,\Delta_j)| = O\left(h_K^2\right).$$

D'autre part, on peut majorer cette covariance en utilisant l'inégalité du lemme (4.2.1, [10])

$$|cov(\Delta_i,\Delta_j)| \leq 4||\Delta_i||_\infty.||\Delta_j||_\infty\alpha(|i-j|).$$

Par la suite, on utilise les téchniques de Masry ([10]), en considèrent une suite u_n des entiers naturels et on montre à l'aide des majorations précédentes que,

$$S_n^{2*} \leq c\left[\sum\sum_{0<|i-j|\leq u_n} h_K^2+\sum\sum_{|i-j|>u_n}\alpha(|i-j|)\right]=O(h_K^2nu_n+n^2\alpha(u_n)),$$

le choix de $u_n=\left[\dfrac{1}{h_K\log n}\right]$ implique que,

$$\begin{aligned} S_n^{2*} &= O\Bigg(h_K^2n.\frac{1}{h_K\log n}+n^2\alpha(h_K\log n)^{-1}\Bigg) \\ &= O\Bigg(\frac{nh_K}{\log n}\Bigg)+O(n^2\alpha(h_K\log n)^{-1}) \end{aligned}$$

Il claire que $\dfrac{nh_K}{\log n}=O(nh_K)$, ainsi,

$$\begin{aligned} S_n^{2*} &= O(nh_K)+O(n^2\alpha(h_K\log n)^{-1}) \\ &= O(nh_K) \end{aligned}$$

Pour le deuxième terme de quantité S_n^2, on a :

$$\begin{aligned} var(\Delta_1) &= \mathbb{E}\left[K^2\left(\frac{x-X_1}{h_K}\right)H^2\left(\frac{t_\alpha-Y_i}{h_H}\right)\right] - \mathbb{E}^2\left[K\left(\frac{x-X_1}{h_K}\right)H\left(\frac{t_\alpha-Y_1}{h_H}\right)\right) \\ &\leq C\int K^2\left(\frac{x-u}{h_K}\right)f(u)du - \left[C\int K\left(\frac{x-u}{h_K}\right)f(u)du\right]^2 \\ &\leq C\int K^2\left(\frac{x-u}{h_K}\right)f(u)du. \end{aligned}$$

On prend le changement de variable $\dfrac{x-u}{h_K} = t$, d'où,

$$var(\Delta_1) \leq Ch_K\int K^2 f(x-h_K t)dt,$$

Pour n assez grand , on obtient,

$$var(\Delta_1) \leq Ch_K\int K^2 f(x)dt,$$

Puisque les fonctions K, f sont bornées ,alors,

$$var(\Delta_1) = O(h_K)$$

Cela nous permet pour un choix de $r = C(\log n)^2$ et $(H9)$ que,

$$Q_1 \leq C\epsilon^{-(a+1)}(\log n)^{2(a+1)}n^{-1-[a+\frac{a+1}{1-a}+\beta_2]},$$

Alors, il existe $v > 0$,

$$Q_1 = O(n^{-1-v}) \tag{2.8}$$

Pour le deuxième terme, on a :

$$Q_2 \leq Cn^{\frac{-\epsilon^2}{32}} \tag{2.9}$$

Les deux équations (2.8) et (2.9) implique que,

$$\sum_n \mathbb{P}\left[\left|\widehat{g}_n(x,t_\alpha) - \mathbb{E}\widehat{g}_n(x,t_\alpha)\right| > \epsilon\sqrt{\frac{\log n}{nh_K}}\right] < \infty$$

Démonstration du lemme 2.3.3.2 La démonstration est smilaire à celle qui précède ,pour reprendre H est égale à 1.

Chapitre 3

L'estimation non paramétrique du quantile conditionnel : Cas fonctionnel

Nous rappelons dans ce chapitre quelques résultats concernant l'estimation non paramétrique de la fonction du quantile conditionnelle lorsque la variable explicative prend ses valeurs dans un espace de dimension infinie. Sous certaines conditions de régularité, nous établissons la convergence presque complète et la normalité asymptotique. Ces propriétés asymptotique sont obtenues sous certains conditions d'indépendance et de dépendance (Mélange).

3.1 Modèle

On considère le champs aléatoire $Z_i = (X_i, Y_i)$ à valeur dans $\mathcal{F} \times \mathbb{R}$ où $(\mathcal{F}, d)$ est un espace semi-métrique de dimension éventuellement infinie. Pour $x \in \mathcal{F}$ on définit la fonction de répartitin conditionnelle de Y sachant $X = x$, notée F^x par

$$F^x(y) = P(Y \leq y | X = x) \qquad \forall y \in \mathbb{R}$$

On suppose que cette distribution conditionnelle de Y sachant X est absolument continue par rapport à la measure de Lebesgue dont, on désigne par f^x la densité conditionnelle et par $f^{x(j)}$ la dérivée d'ordre j de cette densité conditionnelle. On estime la fonction de répartition conditionnelle par

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^{n} K(h_K^{-1}(x - X_i))H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_K^{-1}(x - X_i)}, \qquad \forall y \in \mathbb{R},$$

avec la convention $\frac{0}{0} = 0$, K est un noyau dans H est un fonction de répartition et $h_K = h_{K,n}(resp.h_H = h_{H,n})$ est un suite des nombers réels positifs. En ce qui concerne l'estimation du quantile conditionnel d'ordre $\alpha \in]0, 1[$, on note $q_\alpha(x)$ ce quantile, solution de l'équation

$$F^x(q_\alpha(x)) = \alpha$$

Pour assurer l'existence et l'unicité de $q_\alpha(x)$ on suppose que la fonctin de répartition conditionnelle F^x est strictement croissante. On estime le quantile conditionnel $q_\alpha(x)$ de $\hat{q}_\alpha(x)$ tel que

$$\hat{F}^x(\hat{q}_\alpha(x)) = \alpha$$

3.2 Cas i.i.d.

Dans cette Section, nous considérons le cas où les observations sont indépendantes identiquements distribuées et nous établissons la convergence presque complète de l'estimateur à noyau de $\widehat{t_\alpha}$.

3.2.1 Convergence presque complète

On fixe un point $x \in \mathcal{F}$ et on introduit les conditions suivants :

- $(H1)$ $\mathbb{P}(X \in B(x, h_K) = \phi_x(h_K) > 0$
- $(H2)$ la fonction f^x est j-fois continument différentiable relativement à y sur $(t_\alpha - \epsilon, t_\alpha + \epsilon)$ verifiant $F^{x(l)}(t_\alpha) = 0$ si $1 \leq l < j$ et $0 < |F^{x(j)}(t_\alpha)| < \infty$
- $(H3)$ $\forall (y_1, y_2) \in \mathbb{R} \times \mathbb{R}, \forall (\chi_1, \chi_2) \in \mathcal{N}_{\chi_1} \times \mathcal{N}_{\chi_2}$,

$$\mid F^{x_1}(y_1) - F^{x_2}(y_2) \mid \leq C_x(d(x_1, x_2)^{b_1} + \mid y_1 - y_2 \mid^{b_2} .$$

- $(H4)$

$$\mid F^{x_1(j)}(y_1) - F^{x_2(j)}(y_2) \mid \leq C_x(d(x_1, x_2)^{b_1} + \mid y_1 - y_2 \mid^{b_2} .$$

- $(H5)$ Le paramètre de lissage est tel que :

$$\lim_{n\to\infty} h_K = 0, \quad \exists \alpha > 0, \lim_{n\to\infty} n^\alpha h_H = \infty \quad \text{et} \quad \lim_{n\to\infty} \frac{n\phi_\chi(h_K)}{\log n} = \infty$$

.

- $(H6)$ K est une fonction à support compact dans $[0, 1]$ verifiant :

$$0 < C_1 < K(t) < C_2 < \infty$$

- $(H7)$ $\begin{cases} \int |t|^{b_2} H^{(1)}(t)dt < \infty \\ \exists \sigma > 0, \lim\limits_{y\to 0} \mid y \mid^{1+\sigma} \mid H(y) \mid = 0 \\ \text{H est une fonction strictement croissante,} \end{cases}$

Théorème 3.2.1. Sous les hypothèses $(H1) - (H7)$, on a :

$$\widehat{t_\alpha} - t_\alpha = O(h_K^{b_1/j} + h_H^{b_2/j}) + O\left(\frac{\log n}{n\phi_x(h_K)}\right)^{1/2j}, \qquad p.co.$$

Démonstration des résultats techniques Ce dernier théorème va découler des lemmes suivants :

Lemme 3.2.1.1. Sous les hypothèses (H1), (H3), (H5)-(H7), on a

$$\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha) = O(h_K^{b_1}) + 0(h_H^{b_2}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)$$

Lemme 3.2.1.2. Si les hypothèses (H1), (H4)-(H7) sont satisfaits, alors

$$\widehat{F}^{x(j)}(\widehat{t}_\alpha) - \widehat{F}^{x(j)}(t_\alpha) = O(h_K^{b_1}) + O(h_H^{b_2}) + O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)$$

Lemme 3.2.1.3. Sous les hypothèses du théorème, on a

$$\widehat{t}_\alpha \longrightarrow t_\alpha, \qquad p.co$$

Lemme 3.2.1.4. Sous les conditions (H1) (H7), on a

$$\exists \tau > 0, \qquad \text{tel que,} \qquad \sum_n \mathbb{P}[|\widehat{F}^{x(j)}(t_\alpha^*)| > \tau\Big] < \infty.$$

Démonstration du lemme 3.2.1.1 La démonstration suit le même démarche de la décomposition du cas précédent :

$$\begin{aligned}\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha) &= \frac{1}{\widehat{F}_D^x}\left(\widehat{F}_N^x(\widehat{t}_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right) - \left(F^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right) \\ &- \frac{F^x(t_\alpha)}{\widehat{F}_D^x}\left(\mathbb{E}\widehat{F}_D^x - \widehat{F}_D^x\right)\end{aligned}$$

Tel que

$$\hat{F}_N^x(t_\alpha) = \frac{1}{EK_1}\sum_{i=1}^n K(h_K^{-1}d(\chi, X_i))H(h_H^{-1}(t_\alpha - Y_i))$$

Et

$$\hat{F}_D^x = \frac{1}{EK_1}\sum_{i=1}^n K(h_K^{-1)}d(\chi, X_i))$$

On pose

$$H_i(t_\alpha) = H(h_H^{-1}(t_\alpha - Y_i)), \quad \text{et} \quad K_i = K(h_K^{-1}(t_\alpha - Y_i))$$

La preuve est trés voisine de celle du Théorème précedent, avec une modification dans les fonctions $\hat{F}_N$ resp $\hat{F}_D$ qui devient dans $\mathcal{F}$. Le lemme (3.2.1.1) est un conséquance directe des lemme suivants :

Lemme 3.2.1.5. Sous les hytpothèses (H1),(H6)-(H5), on a :

$$\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x = O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.co.$$

Lemme 3.2.1.6. Sous les hytpothèses (H1),(H3)-(H7), on a :

$$F^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha) = O(h_K^{b_1}) + (h_H^{b_2}), \qquad p.co.$$

Lemme 3.2.1.7. Sous les hytpothèses (H1),(H3)-(H5), on a :

$$\widehat{F}_N^x(\widehat{t}_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha) = O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.co.$$

Corollaire 3.1. Sous les hytpothèses (H1),(H6)-(H5), on a :

$$\sum_{n=1} \mathbb{P}(\widehat{F}_D^x < 1/2) < \infty. \tag{3.1}$$

Et

$$\exists \tau > 0, \quad \text{tel que} \quad \sum_{i=1}^{n} \mathbb{P}\left(|\widehat{F}^{x(1)}(t_\alpha^*)| < \tau\right) < \infty. \tag{3.2}$$

Démonstration du lemme 3.2.1.5

On a

$$\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{K_i}{\mathbb{E}K_i} - 1\right) = \frac{1}{n}\sum_{i=1}^{n}\Delta_i.$$

Avec

$$\Delta_i = \frac{1}{\mathbb{E}K_i}\left(K_i - \mathbb{E}K_i\right)$$

D'aprés l'hypothèse $(H1)$ et $(H6)$, on a,

$$C\phi_x(h_K) < \mathbb{E}K_i < C^{'}\phi_x(h_K)$$

et puisque la fonction K est bornée et $H < 1$, alors on peut écrire,

$$|\Delta_i| < C/\phi_x(h_K) = d$$

et

$$\mathbb{E}\left|\Delta_i\right|^2 < C^{'}/\phi_x(h_K) = \delta$$

On applique maintenant l'inégalité de Hoeffging (lemme (4.2.1, [10])) aux variables Δ_i :

$$\mathbb{P}\left(\left|\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x\right| > \epsilon\right) = \mathbb{P}\left(\sum_{i=1}^{n}|\Delta_i| > \epsilon\right) \leq 2\exp\left(\frac{-n\epsilon^2 C}{4\phi_x(h_K)}\right).$$

En prenant

$$\epsilon = \epsilon_0\sqrt{\frac{\log n}{n\phi_x(h_K)}}$$

On arrive, pour tout $\epsilon_0 > 0$ à

$$\mathbb{P}\left(\left|\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x\right| > \epsilon_0\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right) \leq C^{'}n^{-C\epsilon^2}.$$

Donc,

$$\sum_n \mathbb{P}(\left(\left|\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x\right| > \epsilon_0 \sqrt{\frac{\log n}{n\phi_x(h_K)}}\right) \leq \sum_n C' n^{-C\epsilon^2}.$$

Il suffit de choisir $\epsilon > \frac{1}{\sqrt{C}}$ pour que série converge,
Démonstration du lemme 3.2.1.6 On a, par équidistribution des observations,

$$\mathbb{E}\widehat{F}_N^x(t_\alpha) - F^x(t_\alpha) = \frac{1}{\mathbb{E}K_1}\mathbb{E}(K_1(\mathbb{E}(H_1(t_\alpha)/X)) - F^x(t_\alpha),$$

avec

$$\mathbb{E}(H_1(t_\alpha)/X) = \int_{\mathbb{R}} H(h_H^{-1}(t_\alpha - z))f^x(z)dz.$$

On applique l'intégrale par partie et par le changement de variable $t = (h_H^{-1}(t_\alpha - z))$ pour obtenir,

$$\mathbb{E}(H_1(t_\alpha)/X) = \int_{\mathbb{R}} H^{(1)}(t)F^x(t_\alpha - h_H t)dt,$$

donc,

$$|\mathbb{E}(H_1(t_\alpha)/X) - F^x(t_\alpha)| \leq \int_{\mathbb{R}} H^{(1)}(t)|F^x(t_\alpha) - h_H t) - F^x(t_\alpha)|dt,$$

L'hypothèse (H3) permet d'obtenir,

$$|\mathbb{E}(H_1(t_\alpha)/X) - F^x(t_\alpha)| \leq C_x \int_{\mathbb{R}} H^{(1)}(t)(h_K^{b_1} + |t|^{b_2} h_K^{b_2})dt,$$

et puisque, $\int_{\mathbb{R}} H^{(1)}(t)dt = 1$ et $\int_{\mathbb{R}} H^{(1)}(t)|t|^{b_2}dt < \infty$, alors,

$$\mathbb{E}\widehat{F}_N^x(t_\alpha) - F^x(t_\alpha) = O(h_K^{b_1}) + 0(h_H^{b_2}).$$

Démonstration du lemme 3.2.1.7 On a :

$$\widehat{F}_N^\chi(\widehat{t_\alpha}) - \mathbb{E}\widehat{F}_N^\chi(t_\alpha) = \frac{1}{\mathbb{E}K_1}\sum_{i=1}^n \left(K\left(\frac{d(x,X)}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right) - \mathbb{E}K\left(\frac{d(x,X)}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)\right).$$

on pose

$$\Delta_i = \frac{1}{\mathbb{E}K_1}\left(K\left(\frac{d(x,X)}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right) - \mathbb{E}K\left(\frac{d(x,X)}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)\right)$$

D'aprés l'hypothèse (H6), on a :

$$C\mathbb{E}\mathbb{I}_{B(x,h_K)}(X_i) < \mathbb{E}K_1 < C'\mathbb{E}\mathbb{I}_{B(x,h_K)}(X_i),$$

Alors,

$$C\phi_x(h_K) < \mathbb{E}K_1 < C'\phi_x(h_K),$$

Et puisque la fonction K est bornée et $H < 1$, on peut majorer directement

$$|\Delta_i| \leq \frac{C}{\phi_x(h_K)}, \qquad \mathbb{E}\Delta_i^2 \leq \frac{C'}{\phi_x(h_K)}.$$

On applique l'inégalité de Hoeffging (4.2.1,[10]) aux variable$|\Delta_i|$, on obtient ,

$$\mathbb{P}\left(\left|\widehat{F}_N^x(\widehat{t}_\alpha)-\mathbb{E}\widehat{F}_N^x(t_\alpha)\right|>\epsilon\right)=\mathbb{P}(\left(\frac{1}{n}\left|\sum_{i=1}^n\Delta_i\right|>\epsilon\right)\leq 2\exp\left(\frac{-n\epsilon^2C}{4\phi_x(h_K)}\right)$$

En prenant

$$\epsilon=\epsilon_0\sqrt{\frac{\log n}{n\phi_x(h_K)}}$$

On arrive, pour tout $\epsilon_0>0$,

$$\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha)-\mathbb{E}\widehat{F}_N^x(t_\alpha\right|>\epsilon_0\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)\leq 2n^{-C\epsilon_0^2}.$$

donc,

$$\sum_{i=1}^n\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha)-\mathbb{E}\widehat{F}_N^x(t_\alpha)\right|>\epsilon_0\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right)\leq\sum_{i=1}^n 2n^{-C\epsilon_0^2}.$$

Il suffit de choisir $\epsilon>\dfrac{1}{\sqrt{C}}$ pour que la série converge. Ceci achève la preuve de lemme.
Démonstration du corollaire (3.1) On remarque que :

$$|\widehat{F}_D^x|\leq 1/2\Rightarrow|\widehat{F}_D^x-1|>1/2\Rightarrow|\widehat{F}_D^x-\mathbb{E}\widehat{F}_D^x|>1/2.$$

Alors, on peut écrire,

$$\mathbb{P}[|\widehat{F}_D^x|\leq 1/2\Big]\leq\mathbb{P}[|\widehat{F}_D^x-\mathbb{E}\widehat{F}_D^x|>1/2\Big].$$

Donc,

$$\sum_{n=1}^\infty\mathbb{P}[|\widehat{F}_D^x|\leq 1/2\Big]<\infty.$$

Démonstration du lemme (3.2.1.3) Puisque la fonction H est strictement croissante, l'estimateur $\widehat{F}^x$ est inversible et d'inverse continue, donc,

$$\forall\epsilon>0,\exists\delta(\epsilon)>0,\forall y,|\widehat{F}^x(y)-\widehat{F}^x(t_\alpha)|\leq\delta(\epsilon)\Rightarrow|y-t_\alpha|\leq\epsilon,$$

alors,

$$\begin{aligned}\forall\epsilon>0,\exists\delta(\epsilon)>0,\mathbb{P}(|\widehat{t}_\alpha-t_\alpha|>\epsilon &\leq \mathbb{P}(|\widehat{F}^x(\widehat{t}_\alpha)-\widehat{F}^x(t_\alpha)|\geq\delta(\epsilon)\\ &\leq \mathbb{P}(|\widehat{F}^x(t_\alpha)-F^x(t_\alpha)|\geq\delta(\epsilon).\end{aligned}$$

Ce qui nous permet de déduire la convergence persque complète de t_α^* vers t_α Finallement, en suivant les même lignes de la décomposition du théorème précedente, en remplacant H par $H^{(1)}(h_H^{-1})$,pour obtenir le résultat.

Dmonstration du lemme 3.2.1.2 La démonstration est essentiellement basée sur la décomposition suivante :

$$\begin{aligned}\widehat{F}^{\chi(j)}(\widehat{t}_\alpha) - \widehat{F}^{\chi(j)}(t_\alpha) &= \frac{1}{\widehat{F}_D^\chi}\left(\left(\widehat{F}_N^{\chi(j)}(\widehat{t}_\alpha) - \mathbb{E}\widehat{F}_N^{\chi(j)}(t_\alpha)\right) - \left(F^{\chi(j)}(t_\alpha) - \mathbb{E}\widehat{F}_N^{\chi(j)}(t_\alpha)\right)\right) \\ &- \frac{F^{\chi(j)}(t_\alpha)}{\widehat{F}_D^\chi}\left(\mathbb{E}\widehat{F}_D^\chi - \widehat{F}_D^\chi\right).\end{aligned}$$

On pose

$$\widehat{F}_N^{\chi(j)}(\widehat{t}_\alpha) = \frac{1}{nh_H^j}\sum_{i=1}^n K(h_K^{-1}d(x,X_i))H^{(j)}(h_H^{-1}(t_\alpha - Y_i)),$$

Le lemme (3.2.1.1) est un conséquence directe des lemme suivants :

Lemme 3.2.1.8. Sous les hytpothèses (H1),(H6)-(H5),on a :

$$\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x = o\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.co.$$

Lemme 3.2.1.9. Sous les hytpothèses (H1),(H3)-(H7),on a :

$$F^{x(j)}(t_\alpha) - \mathbb{E}\widehat{F}_N^{x(j)}(t_\alpha) = o(h_K^{b_1}) + o(h_H^{b_2}), \qquad p.co.$$

Lemme 3.2.1.10. Sous les hytpothèses (H1),(H3)-(H5),on a :

$$\widehat{F}_N^{x(j)}(\widehat{t}_\alpha) - \mathbb{E}\widehat{F}_N^{x(j)}(t_\alpha) = o\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.co.$$

Corollaire 3.2. Sous les hytpothèses (H1),(H6)-(H5),on a :

$$\sum_{n=1}\mathbb{P}(\widehat{F}_D^x < 1/2) < \infty.$$

Corollaire 3.3. Sous les hytpothèses (H1)-(H5),on a :

$$\exists \tau > 0, \quad \text{tel que} \quad \sum_{i=1}^n \mathbb{P}\left(|\widehat{F}^{x(1)}(t_\alpha^*)| < \tau\right) < \infty.$$

Démonstration du lemme 3.2.1.9 Nous avons :

$$\mathbb{E}\widehat{F}_N^{\chi(j)}(t_\alpha) - F^{\chi(j)}(t_\alpha) = \frac{1}{h_H^j}\mathbb{E}(K_1(\mathbb{E}(H_1^j(t_\alpha)/X)) - F^{x(j)}(t_\alpha),$$

et

$$\mathbb{E}(H_1^{(j)}(t_\alpha)/X) = \int_{\mathbb{R}} H^{(j)}(h_H^{-1}(t_\alpha - z))f^x(z)dz.$$

Par le changement de variable $t = (h_H^{-1}(t_\alpha - z))$ on obtient,

$$\begin{aligned}\mathbb{E}(H_1^j(t_\alpha)/X) &= \int_{\mathbb{R}} H^{(j)}(h_H^{-1}(t_\alpha - z))f^x(z)dz \\ &= -h_H^{-1}\int_{\mathbb{R}} f^x(z)dH^{j-1}(h_H^{-1}(t_\alpha - z))\end{aligned}$$

On applique l'intégrale par partie

$$\begin{aligned}\mathbb{E}(H_1^j(t_\alpha)/X) &= h_H(f^x(z)H^{j-1}(h_H^{-1}(t_\alpha - z)) + h_H\int_{\mathbb{R}} f^{1(\chi)}(z)H^{j-1}(h_H^{-1}(t_\alpha - z))dz \\ &= -h_H(f^x(z)dH^{j-1}(h_H^{-1}(t_\alpha - z)) - h_H^2(f^x(z)H^{j-2}(h_H^{-1}(t_\alpha - z)) \\ &+ h_H^2\int_{\mathbb{R}} f^{2(\chi)}(z)H^{j-2}(h_H^{-1}(t_\alpha - z))dz \\ &= -\sum_{l=1}^{j} h_H^l(f^{(l-1)x}(z)dH^{j-l}(h_H^{-1}(t_\alpha - z)) \\ &+ h_H^j\int_{\mathbb{R}} f^{j(\chi)}(z)H(h_H^{-1}(t_\alpha - z))dz\end{aligned}$$

D'aprés l'hypothèse (H7) et par une deuxième intégration, on a

$$\begin{aligned}\mathbb{E}(H_1^j(t_\alpha)/X) &= h_H^j\int_{\mathbb{R}} f^{j(\chi)}(z)H(h_H^{-1}(t_\alpha - z))dz \\ &= h_H^{j-1}\int_{\mathbb{R}} F^{j(\chi)}(z)H^{(1)}(h_H^{-1}(t_\alpha - z))dz\end{aligned}$$

Pour $t = \dfrac{t_\alpha - z}{h_H}$, donc

$$\mathbb{E}(H_1^j(t_\alpha)/X) = h_H^j\int_{\mathbb{R}} F^{j(\chi)}(t_\alpha - h_H t)H^{(1)}(t)dt$$

Alors,

$$|\mathbb{E}(H_1^j(t_\alpha)/X) - h_H^j F^{x(j)}(t_\alpha)| \leq h_H^j\int_{\mathbb{R}} H^{(1)}(t)|F^{j(\chi)}(t_\alpha - h_H t) - F^{j(\chi)}(t_\alpha)|dt,$$

L'hypothèse $(H3)$ permet d'obtenir,

$$|\mathbb{E}(H_1^j(t_\alpha)/X) - h_H^j F^{x(j)}(t_\alpha)| \leq C_x h_H^j\left(\int_{\mathbb{R}} H^{(1)}(t)h_K^{b_1}dt + \int_{\mathbb{R}} H^{(1)}(t)|t|^{b_2}h_K^{b_2}dt\right),$$

Comme $\int_{\mathbb{R}} H^{(1)}(t)dt = 1$ et $\int_{\mathbb{R}} H^{(1)}(t)|t|^{b_2}dt < \infty$, alors

$$\mathbb{E}\widehat{F}_N^{\chi(t)}(t_\alpha) - F^{\chi(t)}(t_\alpha) = o(h_K^{b_1}) + o(h_H^{b_2}).$$

Démonstration du lemme 3.2.1.10 On a :

$$\begin{aligned}\widehat{F}_N^{\chi(j)}(t_\alpha) - \mathbb{E}\widehat{F}_N^{\chi(j)}(t_\alpha) &= \frac{1}{nh_H^j \mathbb{E}K_1} \sum_{i=1}^{n} \left(K\left(\frac{d(x,X)}{h_K}\right) H^{(j)}\left(\frac{y-Y_i}{h_H}\right) \right. \\ &- \left. \mathbb{E}K\left(\frac{d(x,X)}{h_K}\right) H^{(j)}\left(\frac{y-Y_i}{h_H}\right)\right).\end{aligned}$$

on pose :

$$\Delta_i = \frac{1}{h_H^{(j)}\mathbb{E}K_1}\left(K\left(\frac{d(x,X)}{h_K}\right) H^{(j)}\left(\frac{y-Y_i}{h_H}\right) - \mathbb{E}K\left(\frac{d(x,X)}{h_K}\right) H^{(j)}\left(\frac{y-Y_i}{h_H}\right)\right)$$

D'aprés l'hypothèse $(H5)$, on a,

$$C\mathbb{E}\mathbb{I}_{B(x,h_K)}(X_i) < \mathbb{E}K_1 < C^{'}\mathbb{E}\mathbb{I}_{B(x,h_K)}(X_i),$$

alors,

$$C\phi_x(h_K) < \mathbb{E}K_1 < C^{'}\phi_x(h_K),$$

et puisque la fonction K est bornée et $H < 1$, on peut majorer directement

$$|\Delta_i| \leq \frac{C}{h_H^j\phi_x(h_K)}, \qquad \mathbb{E}\Delta_i^2 \leq \frac{C^{'}}{h_H^{2j-1}\phi_x(h_K)}.$$

On applique maintenant l'inégalité de Hoeffging (4.2.1,[10]) aux variable$|\Delta_i|$,

$$\mathbb{P}\left(\left|\widehat{F}_N^{\chi(j)}(\widehat{t_\alpha}) - \mathbb{E}\widehat{F}_N^{\chi(j)}(t_\alpha)\right| > \epsilon\right) = \mathbb{P}(\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Delta_i\right| > \epsilon\right) \leq 2\exp\left(\frac{-n\epsilon^2 C}{4h_H^{2j-1}\phi_\chi(h_K)}\right)$$

Avec

$$\epsilon = \epsilon_0\sqrt{\frac{\log n}{nh_H^{2j-1}\phi_x(h_K)}}$$

On arrive, pour tout $\epsilon_0 > 0$,

$$\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha\right| > \epsilon_0\sqrt{\frac{\log n}{nh_H^{2j-1}\phi_x(h_K)}}\right) \leq 2n^{-C\epsilon_0^2}.$$

donc,

$$\sum_n \mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right| > \epsilon_0\sqrt{\frac{\log n}{nh_H^{2j-1}\phi_x(h_K)}}\right) \leq \sum_{i=1}^{n} 2n^{-C\epsilon_0^2}.$$

Il suffit de choisir $\epsilon > \dfrac{1}{\sqrt{C}}$ pour que la série converge. Ceci achève la preuve de lemme.

3.2.2 Normalité asymptotique

Dans cette section, nous établissons la normalité asymptotique de l'estimateur du quantile conditionnel lorsque la variable explicative est fonctionnelle.

3.2.3 Hypothèses

On garde les hypothéses $(H1)$ -$(H5)$ du cas paragraphe précédent et on ajoute les hypothèses suivants :

- $(H6) H$ est un nouyon vers fonction

i) il $\exists$ une fonction intégrable g tel que,

$$\left|H(t) - H(s)\right| \leq Cg\left|t - s\right|$$

ii) $\displaystyle\int |t|^{b_2} H^{(1)}(t) dt < \infty$

Théorème 3.2.2. Sous les hypothèses $(H1)$- $(H5)$ et $(H6)$, on a

$$\left(\frac{n\psi_K(h_K)}{f^{\chi}(t_\alpha)\psi_K^2(h_K)}\right)^{1/2} (\hat{t}_\alpha - t_\alpha) \longrightarrow \mathcal{N}(0, \sqrt{\alpha(\alpha-1)}). \tag{3.3}$$

Où $\psi_K(h_K) = -\int g^{(1)}(t)\phi_\chi(h_K t)dt$

Démonstration des résultats techniques
On utilise le développement de taylor d'ordre 1 de l'estimateur $\widehat{F}^x$ au point $\widehat{t}_\alpha$ car la fonction H est de classe C^1, alors :

$$\widehat{F}^x(\widehat{t}_\alpha) - \widehat{F}^x(t_\alpha) = (\widehat{t}_\alpha - t_\alpha)\widehat{F}^{x(1)}(t^*_\alpha)$$

avec $t^*_\alpha \in (t_\alpha, \widehat{t}_\alpha)$

$$\left|\widehat{t}_\alpha - t_\alpha\right| = \frac{1}{\left|\widehat{F}^{x(1)}(t^*_\alpha)\right|}\left[\left|\widehat{F}^x(t_\alpha) - F^x(t_\alpha)\right|\right].$$

Alors,

$$\sqrt{\frac{n(\psi_g(h_K))^2}{f^n(t_\alpha)\psi_{k^2}(h_K)}}\left|\widehat{t}_\alpha - t_\alpha\right| = \sqrt{\frac{n(\psi_K(h_K))^2}{f^{\chi}(t_\alpha)\psi_{K^2}(h_K)}}\frac{1}{\left|\widehat{F}^{x(1)}(t^*_\alpha)\right|}\left[\left|\widehat{F}^x(t_\alpha) - F^x(t_\alpha)\right|\right].$$

La démonstration est basé sur les lemmes suivants :

Lemme 3.2.3.1. Sous les hypothèses $(H1) - (H5)$ et $(H6)$, on a,

$$\sqrt{\frac{n(\psi_K(h_K))^2}{f^n(t_\alpha)\psi_{K^2}(h_K)}}\Big(\alpha\widehat{F}^x_D - \widehat{F}^x_N(t_\alpha) - \mathbb{E}(\alpha\widehat{F}^x_D - \widehat{F}^x_N(t_\alpha))\Big) \to \mathcal{N}\Big(0, \sqrt{\alpha(\alpha-1)}\Big), \quad \text{P.co.}$$

Corollaire 3.4. Sous les hypothèses $(H1) - (H5)$, on a,

$$\exists \tau > 0, \quad \text{Tel que} \quad \sum_n \mathbb{P}\left[\left|\widehat{F}^{x(1)}_N(t^*_\alpha)\right| < \tau\right] < \infty.$$

Démonstration du lemme 3.2.3.1
D'aprés le Théorème de centrale limite, il existe $\delta > 0$, tel que

$$\frac{\sum_{i=1}^{n}\mathbb{E}\left|\sqrt{\frac{n(\psi_g(h_K))^2}{f^X(t_\alpha)\psi_{K^2}(h_K)}}\frac{1}{n}\mid L_i(x)-\mathbb{E}L_i(x)\mid\right|^{2+\delta}}{\left(var\sqrt{\frac{n(\psi_g(h_K))^2}{f^x(t_\alpha)\psi_{K^2}(h_K)}}\frac{1}{n}\sum_{n}L_i(x)\right)^{1+\frac{\delta}{2}}}\longrightarrow 0$$

Avec

$$L_i(x)=\frac{1}{\mathbb{E}K_i}\Big[K_iH_i-\alpha K_i\Big]$$

Donc,

$$\begin{aligned}
\frac{1}{n}var\sum_{i=1}^{n}L_i(x) &= var(\hat{F}_N^x(t_\alpha)-\alpha\hat{F}_D^x)\\
&= \frac{1}{(n\mathbb{E}K_1)^2}\sum_{i=1}^{n}var(K_iH_i(t_\alpha)-\alpha K_i)\\
&= \frac{1}{n(\mathbb{E}K_1)^2}var(K_1H_1(t_\alpha)-\alpha K_1)\\
&= \frac{1}{n(\mathbb{E}K_1)^2}\mathbb{E}(K_1H_1(t_\alpha)-\alpha K_1)^2-\mathbb{E}^2\Big(K_1H_1(t_\alpha)-\alpha K_i\Big)\\
&= \frac{1}{n(\mathbb{E}K_1)^2}\mathbb{E}(K_1^2\mathbb{E}\left(\frac{K_1^2(H_1(t_\alpha)-\alpha)^2}{\mathbb{E}(K_1^2}\right)-\frac{1}{n}\left(\mathbb{E}\left(\frac{K_1(H_1(t_\alpha)-\alpha}{\mathbb{E}K_1}\right)\right)^2
\end{aligned}$$

D'autre part,

$$\mathbb{E}\bigg(K\frac{d(x,X_i)}{h_K}\bigg)=\int_0^1 K(u)d\mathbb{P}^{\frac{d(x,X_i)}{h_K}}(u)$$

On a

$$\int_0^u K^{'}(t)dt=K(u)-K(0)$$

Donc

$$\begin{aligned}
\mathbb{E}\bigg(K\frac{d(x,X_i)}{h_K}\bigg) &= \int_0^1 K(0)dP^{\frac{d(\chi,x_1)}{h_K}}(u)+\int_0^1\int_0^u K'(t)dtdP^{\frac{d(\chi,x_1)}{h_K}}(u)\\
&= K(0)\phi_x(h_K)+\int_0^1\int_0^1 K^{'}(t)\mathbb{1}_{[0,u]}(t)dtd\mathbb{P}^{\frac{d(x,X_i)}{h_K}}(u)\\
&= K(0)\phi_x(h_K)+\int_0^1\int_0^1 K^{'}(t)\mathbb{1}_{[t,1]}(t)dtd\mathbb{P}^{\frac{d(x,X_i)}{h_K}}(u)\\
&= K(0)\phi_x(h_K)+\int_0^1 K^{'}(u)\mathbb{P}\Big(t<\frac{d(\chi,x_1)}{h_K}<1\Big)dt\\
&= -\int K^{(1)}(t)\phi_x(h_Kt)dt, \quad \text{car} \quad K(1)=0\\
&= \psi_K(h_K)
\end{aligned}$$

De la même manière, on trouve d'où

$$\frac{(\mathbb{E}K_i)^2}{\mathbb{E}K_i^2} = O(1)$$

Et à l'aide de l'hypothèses (), on trouve

$$|\mathbb{E}(H_i(t_\alpha)/X) - \alpha)| \leq C_\chi \left(h_K^{b_1} + h_H^{b_2} \int_{\mathbb{R}} |t|^{b_2} H'(t)dt \right) \tag{3.4}$$

Et

$$\left|\mathbb{E}(H_1^2(t_\alpha)/X) - \alpha)\right| \leq C_\chi \left(h_K^{b_1} + h_H^{b_2} \int_{\mathbb{R}} 2H(t)H'(t)|t|^{b_2} dt \right) \tag{3.5}$$

Il suffit d'écrire

$$\begin{aligned} \frac{1}{\mathbb{E}K_1}\mathbb{E}\Big(K_i(H_i-\alpha)\Big)^2 &= \frac{1}{\mathbb{E}K_1}\left(\mathbb{E}K_i^2H_i^2 + \alpha^2\mathbb{E}K_i^2 - 2\alpha\mathbb{E}K_i^2H_i\right) \\ &= \frac{1}{\mathbb{E}K_1}\left(\mathbb{E}(K_i^2(H_i^2-\alpha) - 2\alpha\mathbb{E}(K_i^2(H_i-\alpha) + (\alpha-\alpha^2)\right) \end{aligned}$$

Donc

$$\frac{n(\psi_K(h_K))^2}{\psi_{K^2}(h_K)} var\Big(\alpha\widehat{F}_D^x - \widehat{F}_n^x(t_\alpha)\Big) \to \alpha(\alpha-1)$$

Il reste à évaluer la limite, pour laquelle

$$n^{-1}\Big(\frac{(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2}\Big)^{1+\frac{\delta}{2}}\sum_{i=1}^n\Big|\mathbb{E}L_i - \mathbb{E}L_i\Big|^{2+\delta} \leq \underbrace{2^{1+\delta}n^{-1}\Big(\frac{(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2}\Big)^{1+\frac{\delta}{2}}\Big|\sum_{i=1}^n\mathbb{E}L_i^{2+\delta}\Big|}_{D_1}$$

$$\underbrace{+\,2^{1+\delta}n^{-1}\Big(\frac{(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2}\Big)^{1+\frac{\delta}{2}}\Big|\sum_{i=1}^n\mathbb{E}L_i^{2+\delta}\Big|}_{D_2}$$

Comme on a :

$$\frac{(\mathbb{E}K_1)^l}{\mathbb{E}K_1^l} = O(1)$$

Pour le terme D_1, on a

$$\begin{aligned} D_1 &\leq 2^{1+\delta}n^{-\frac{\delta}{2}}(\mathbb{E}K_1)^{-1-\frac{\delta}{2}}\mathbb{E}K_1^{2+\delta}\Big[H_1(t_\alpha)-\alpha\Big]^{2+\delta} \\ &\leq C\Big(n\mathbb{E}K_i^2\Big)^{\frac{-\delta}{2}}\mathbb{E}K_1^2/\mathbb{E}K_i^{2+\delta} \\ &\leq C\Big(n\phi_x(h_K)\Big)^{\frac{-\delta}{2}} \to 0 \end{aligned}$$

Puis D_2,

$$D_2 \leq 2^{1+\delta} n^{-\frac{\delta}{2}} \Big(\frac{(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2}\Big)^{1+\frac{\delta}{2}} \Big(\mathbb{E}(K_1(H_i(t_\alpha) - \alpha)\Big)^{2+\delta}$$

$$\leq Cn^{-\frac{\delta}{2}} (\mathbb{E}K_i)^{3+\frac{3\delta}{2}} \Big(\frac{\mathbb{E}K_i}{\mathbb{E}K_1^2}\Big)^{1+\frac{\delta}{2}} \to 0$$

Démonstration On a,

$$\begin{aligned} \mathbb{E}\Big(\alpha \widehat{F}_D^x - \widehat{F}_n^x(t_\alpha)\Big) &= \mathbb{E}\Big(\alpha\Big(\frac{1}{nh_K}\sum_{i=1}^{n} K_i(x)\Big) - \frac{1}{nEK_1}\sum_{i=1}^{n} K_i(x)H_i(t_\alpha)\Big) \\ &= \frac{1}{nEK_1} EK_1\left(\alpha - \mathbb{E}H_1(t_\alpha/X)\right) \longrightarrow 0 \end{aligned}$$

$$\frac{n(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2} var\Big(\alpha \widehat{F}_D^x - \widehat{F}_n^x(t_\alpha)\Big) = \mathbb{E}\left(\frac{K_1^2(H_1(t_\alpha) - \alpha)^2}{\mathbb{E}K_1^2}\right) - \frac{(\mathbb{E}K_1)^2}{\mathbb{E}K_1^2}\mathbb{E}\Big(\frac{K_1(H_1(t_\alpha) - \alpha)}{\mathbb{E}K_1}\Big)^2$$

3.3 Cas de mélange

Dans cette Section , nous considérons le cas où les observations sont α-mélangeantes et nous établissons la convergence presque complète de l'estimateur à noyau de $\widehat{t_\alpha}$. Nous sommes amenès à fait sur les hypothèses sont du même ordre que celles discutées dans le i.i.d et aussi s'ajoute les conditions suivants :

- $(H8)$ $(X_i, Y_i)_{i\in N}$ une suite des variable aléatoires α-mélangeantes,de coefficient de mélange α_n verifiant :

$$\exists a > 0, \exists c > 0 : \forall n \in \mathbb{N}, \alpha(n) \leq cn^{-a}$$

- $(H9) O < \sup_{i\neq j} \mathbb{P}\Big[(X_i, Y_i) \in B(x, h_K) \times B(x, h_K)\Big] = o\Big(\frac{(\phi_x(h_K))^{(a+1)/a}}{n^{1/a}}\Big).$
- $(H10)$

$$\lim_{n\to\infty} h_H = 0$$

$$\exists \eta > 0, \quad Cn^{\frac{3-a}{a+1}+\eta} \leq \phi_x(h_K) \leq C' n^{\frac{1}{1-a}}, \quad \text{avec} \quad a > (5 + \sqrt{17})/2.$$

3.3.1 Propriétés asymptotiques

Théorème 3.3.1. Sous les hypothèses $(H1) - (H9)$, on a :

$$\widehat{t}_\alpha - t_\alpha = o(h_K^{b_1} + h_k^{b_2}) + o\left(\frac{\log n}{n\phi_x(h_K)}\right)^{1/2}, p.co.$$

En reprend les mêmes arguments employées dans la démonstration du lemme (3.2.1.6), donc on va démontrer seulement la partie de dispersion

Lemme 3.3.1.1. Sous les hypothèses $(H1)-(H4)$ et $(H6)-(H9)$, on a :

$$\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x = o\Big(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\Big), \qquad \text{P.co.}$$

Lemme 3.3.1.2. Sous les hypothèses $(H1)-(H4)$ et $(H6)-(H9)$, on a :

$$\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha) = o\Big(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\Big), \qquad \text{P.co.}$$

Démonstration du lemme 3.3.1.1 Il suffit de montrer que

$$\mathbb{P}\left(\left|\hat{F}_N^x(t_\alpha) - \mathbb{E}\hat{F}_N^x(t_\alpha)\right| > \eta\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right) \leq Cn^{-1-\nu}$$

Pour cela, on applique l'inégalité de Fuk-Nagaevs([10]), on obtient,

$$\mathbb{P}\Big(\Big|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\Big| > 4\lambda\Big) = \mathbb{P}\Big(\Big|\frac{1}{\mathbb{E}K_1}\sum_{i=1}^n \Delta_i\Big| > 4\lambda\Big) \leq 2nCr^{-1}\Big(\frac{2r}{\lambda}\Big)^{a+1} + \Big(1+\frac{\lambda^2}{rS_n^2}\Big)^{-r/2},$$

Tel que

$$\Delta_i = K_iH_i - \mathbb{E}K_iH_i$$

En ce qui concerne le deuxième terme, on a

$$\begin{aligned} S_n^2 &= \sum_{i=1}^n\sum_{j=1}^n |cov(\Delta_i\Delta_j)| \\ &= \underbrace{\sum_{i\neq j} cov(\Delta_i,\Delta_j)}_{S_n^{cov}} + \underbrace{\sum_{i\neq j} var(\Delta_i)}_{S_n^{var}} \end{aligned}$$

Pour la quantité S_n^{cov}, on a

$$\begin{aligned} \Big|Cov(\Delta_i,\Delta_j)\Big| &= \Big|\mathbb{E}(\Delta_i,\Delta_j) - \mathbb{E}(\Delta_i)\mathbb{E}(\Delta_j)\Big| \\ &= \mathbb{E}(K_iH_iK_jH_j) + \mathbb{E}(K_iH_i)\mathbb{E}(K_jH_j) \\ &\leq C\mathbb{E}(K_iK_j) + \mathbb{E}(K_i)\mathbb{E}(K_j) \end{aligned}$$

Comme le noyau K est bornée à support compact $[0,1]$ alors, il existe une constante vérifiant

$$\begin{aligned} \Big|Cov(\Delta_i,\Delta_j)\Big| &\leq C\mathbb{E}(\mathbb{I}_{B(\chi,h_K)}(X_i)\mathbb{I}_{B(\chi,h_K)}(X_j)) + C\mathbb{E}(\mathbb{I}_{B(\chi,h_K)}(X_j)\mathbb{E}(\mathbb{I}_{B(\chi,h_K)}(X_i) \\ &= \leq C\mathbb{P}\Big((X_i,X_j) \in B(x,h_K)\times B(x,h_K)\Big) \\ &+ C'\mathbb{P}\Big(X_i \in B(x,h_K)\Big)\mathbb{P}\Big(X_j \in B(x,h_K)\Big) \\ &\leq C\left(\chi_\chi(h_K)\phi_x(h_K) + C'\phi_x^2(h_K)\right) \end{aligned}$$

La dernière égalité à cause de l'hypothèse $(H10)$, donc

$$\Big|Cov(\Delta_i,\Delta_j)\Big| = O\left(\chi_\chi(h_K)\phi_x(h_K)\right) \tag{3.6}$$

D'autre part, nous avons d'aprés l'inégalité de covariance (4.2.1,[10])

$$\forall i \neq j, \qquad \Big|Cov(\Delta_i,\Delta_j)\Big| \leq C\alpha(|i-j|) \tag{3.7}$$

En effet, on utilise les techniques de Masry ([10], 1986) et on partage la somme sur les deux ensembles

$$S_n^{cov} = \sum \sum_{0<|i-j|\leq u_n} \Big|Cov(\Delta_i,\Delta_j)\Big| + \sum \sum_{|i-j|>u_n} \Big|Cov(\Delta_i,\Delta_j)\Big|$$

Ou u_n est une suite des entiers tend vers l'infinie quand n tend vers l'infinie. Alors,

$$\begin{aligned} S_n^{cov} &\leq C\left[\sum \sum_{0<|i-j|\leq u_n} \left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \sum \sum_{|i-j|>u_n} \alpha(|i-j|)\right] \\ &= O\Bigg(nu_n\Bigg(\chi_\chi(h_K)\phi_x(h_K) + n^2\alpha(u_n)\Bigg)\Bigg), \end{aligned}$$

Si $u_n = \left(\dfrac{1}{\chi_x(h_K)}\right)$, alors

$$\begin{aligned} S_n^{cov} &= O\Bigg(n\phi_x(h_K) + n^2\alpha\left(\frac{1}{\chi_\chi(h_K)}\right)\Bigg) \\ &\leq C(n\phi_x(h_K) + n^2(\chi_x(h_K))^a) \end{aligned}$$

Pour la quantité S_n^{var}, on a

$$var(\Delta_i) = \mathbb{E}(K_1H_1)^2 - \mathbb{E}^2(K_1H_1)$$

On utilise l'hypothèse $(H1)$ pour on obtenir

$$S_n^{var} \leq C\Big(\phi_x(h_K)) + (\phi_x(h_K))^2\Big),$$

D'où,

$$S_n^2 = O(n\phi_x(h_K)).$$

Nous appliquons, maintenant l'inégalité de Fuk-Nagaev (4.2.1, [10]) aux variables Δ, on obtient,

$$\begin{aligned} \mathbb{P}\Big(\Big|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\Big| > \lambda\Big) &= \mathbb{P}\Bigg(\Bigg|\sum_{i=1}^n \Delta_i\Bigg| > n\mathbb{E}K_1\lambda\Bigg) \\ &\leq \mathbb{P}\Bigg(\Bigg|\sum_{i=1}^n \Delta_i\Bigg| > \frac{4n\chi_\chi(h_K)\lambda}{4}\Bigg) \\ &= \leq 2nCr^{-1}\left(\frac{8r}{\lambda n\chi_\chi(h_K)}\right)^{a+1} \\ &= \left(1 + \frac{\lambda^2 n^2(\chi_\chi(h_K))^2}{16rS_n^2}\right)^{-r/2} \end{aligned}$$

Si on prend $\lambda = \lambda_0\sqrt{\dfrac{\log n}{n\chi_\chi(h_K)}}$, on obtient alors

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right| > \lambda\sqrt{\frac{\log n}{n\chi_\chi(h_K)}}\right) &\leq 4\left(1+\frac{\eta^2\log n}{16r}\right)^{-r|2} \\
&+ 2ncr^{-1}\left(\frac{8r}{n\chi_\chi(h_K)\eta\sqrt{\dfrac{\log n}{n\chi_\chi(h_K)}}}\right)^{a+1} \\
&\leq 4\left(1-\frac{\eta^2\log n\ 32}{-r|2}\right)^{-r|2} \\
&+ 2nCr^{-1}\left(\frac{8r}{\eta}\right)^{a+1}(n\log n\chi_\chi(h_K))^{-(a+1)|2}
\end{aligned}$$

Donc,

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right| > \lambda\sqrt{\frac{\log n}{n\chi_\chi(h_K)}}\right) &\leq 4n^{-\eta^2|32} \\
&+ C\eta^{-(a+1)}n^{(1-(a+1)|2)}r^a n^{((a-3)|2-\beta(a+1)|2)}(\log n)^{-(a+1)|2}
\end{aligned}$$

On choisi $r = C(\log n)^2$, on trouve

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right| > \lambda\sqrt{\frac{\log n}{n\chi_\chi(h_K)}}\right) &\leq Cn^{-\eta^2|32} \\
&+ Cn^{-1-\beta(a+1)|2}(\log n)^{(3a-1)|2}
\end{aligned}$$

Pour η suffisament grand, on a

$$\exists\nu > 0, \qquad \mathbb{P}\left(\left|\widehat{F}_N^x(t_\alpha) - \mathbb{E}\widehat{F}_N^x(t_\alpha)\right| > \lambda\sqrt{\frac{\log n}{n\chi_\chi(h_K)}}\right) \leq Cn^{-1-\nu}$$

Démonstration du lemme 3.3.1.2 En reprend les même arguments du lemme précédente.

Chapitre 4

L'estimation du modèle de troncature à gauche : Cas réel

Ce travail effectué dans le cadre de ce mémoire considère la tronquée a gauche où n'est pas observé que si $(Y_i \geq T_i)$. la question posé est d'expliquer la durée de survie y sujette à une variable T par une variable aléatoire réelle X. pour ce faire, nous disposons de n copies indépendants (X, Y_i, T_i) du triplet (X, Y, T).

4.1 Cas i.i.d

4.1.1 Modèle

Considérons une suite de variables aléatoires indépendantes $Y_1, \cdots, Y_N$ de même fonction de répartition F inconnue. Ces variables aléatoires sont regardées comme les durées de vie des sujets étudiés. La troncature aléatoire à gauche peut notamment avoir lieu si le temps d'origine de la durée de vie étudiée précéde le temps d'origine de l'étude. Ce modèle peut survenir dans différents champs d'applications comme l'astronomie et les études médicales. Soit $T_1, \cdots, T_N$ une suite de variables aléatoires indépendantes de même fonction de répartition G inconnue. On suppose aussi que ces variables sont indépendantes des Y_i. La taille de l'échantillon N est déterministe mais inconnue Dans le modèle de troncature à gauche, (X_i, T_i) est observé lorsque $Y_i \geq T_i$, si non rien n'est observé. Pour éviter la confusion, on note $(X_i, T_i, i = 1, \cdots, n)(n \leq N)$ l'échantillon observé $(i.eY_i \geq T_i)$. Une conséquence de la troncature, la taille de l'échantillon vraiment observé n est une variable aléatoire distribuée selon la loi Binomiale de paramètre N Il est clair que si $\mu = 0$, aucune donnée peut observer. Pour cela, nous supposons, dorénavant, que $\mu \neq 0$ par la loi forte des grands nombres on a, lorsque N tend vers ∞

$$\hat{\mu} = \frac{n}{N} \to \mu, \quad \text{p.s}$$

Lemdani et Ould-Said (2007) ont prouvé que la propriété i.i.d de l'échantillon observé de taille n est déduite de celle de l'échantillon de taille N. Sous le modèle de troncature à gauche la distribution conjointe conditionnelle (Stute (1993), et Zhou (1996) d'un (Y,T)observé devient

$$\begin{aligned}
J^*(y,t) &= \mathbb{P}(Y \le y, T \le t) \\
&= \mathbb{P}(Y \le y, T \le t / Y \ge T) \\
&= \frac{\mathbb{P}(Y \le y, T \le t, Y \ge T)}{\mathbb{P}(Y \ge T)} \\
&= \frac{\mathbb{P}(Y \le y, T \le t, T \le Y)}{\mathbb{P}(Y \ge T)} \\
&= \mu^{-1} \int_{-\infty}^{y} G(t \wedge u) dF(u)
\end{aligned}$$

Où $t \wedge u = \min(t,u)$. Les distributions marginales sont donc définies par

$$\begin{aligned}
F^*(y) &= J^*(y, \infty) \\
&= \mu^{-1} \int_{-\infty}^{y} G(u) dF(u)
\end{aligned}$$

et

$$\begin{aligned}
G^*(t) &= J^*(\infty, t) \\
&= \mu^{-1} \int_{-\infty}^{\infty} G(u \wedge u) dF(u) \\
&= \mu^{-1} \int_{-\infty}^{t} \int_{-\infty}^{t \wedge u} dG(v) dF(u) \\
&= \mu^{-1} \int_{-\infty}^{t} dG(v) \int_{v}^{\infty} dF(u) \\
&= \mu^{-1} \int_{-\infty}^{t} (1 - F(v)) dG(v),
\end{aligned}$$

Qui peuvent être estimés respectivenment par $F_n^* = n^{-1} \sum_{i=1}^{n} I_{\{Y_i \le y\}}$ et $G_n^* = n^{-1} \sum_{i=1}^{n} I_{\{T_i \le t\}}$

Soit la fonction $C(.)$ définie par

$$\begin{aligned}
C(y) &= \mathbb{P}(T \le y \le y | y \ge t) \\
&= G^*(y) - F^*(y) \\
&= \frac{\mathbb{P}(T \le y \le Y, Y \ge T)}{\mathbb{P}(Y \ge T)} \\
&= \frac{\mathbb{P}(T \le y \le Y)}{\mathbb{P}(Y \ge T)} \\
&= \mu^{-1} \mathbb{P}(T \le y) \mathbb{P}(Y \ge y) \\
&= \mu^{-1} G(y)(1 - F(u))
\end{aligned}$$

On peut être estimé par

$$\begin{aligned}
C_n(y) &= G_n^*(y) - F_n^*(y) \\
&= n^{-1} \sum_{i=1}^{n} 1_{\{T_i \le y \le Y_i\}}
\end{aligned}$$

Lynden-Bell 1971 introduit les estimateurs de maximum de vraisemblance non paramétriques de F et G donnés par les estimateurs produit-limite suivants

$$F_n(y) = 1 - \prod_{i/Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)}\right], \quad \text{et} \quad G_n(y) = 1 - \prod_{i/T_i > \leq y} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)}\right]$$

Woodroofe (1985) établit la convergence presque sûre des estimateurs de Lynden- Bell et ainsi que les conditions dŠidentifabilité du modèle et a remarqué que F et G peuvent être estimés complétement seulement si $a_G \leq a_F$, $b_G \leq b_F$ et $\int_{a_F}^{\infty} \frac{dF}{G} < \infty$, (où a_G,b_G et a_F,b_F désignent les points finaux de G et F respectivemeny).

Maintenant, en plus de deux variables considérées auparavant Y et T, nous considérons un vecteur aléatoire $X \in R^d$ de covariables, supposé absolument continu de fonction de répartition $V(.)$ et une densité continue $v(.)$. On note $(X_i, Y_i, T_i), 1 \leq i \leq n$ lééchantillon observé $(i.e, Y_i, T_i)$. Dorénavant, on suppose que T est indépendante de (X, Y). Dans que se suit, nous allons construire les estimateurs de $V(.)$ et $v(.)$. Premièrement, l'estimateur à noyau naturel de la densité de la covariable $v(.)$ est donné par

$$v_n^* \;\; = \;\; \frac{1}{nh_n^d} \sum_{i=1}^{N} k_d\left(\frac{x - X_i}{h_N}\right)$$

Où $K_d : \mathbb{R}^d \longrightarrow \mathbb{R}$ est un noyau fixé avec $\int_{\mathbb{R}^d} K_d = 1$ et $(h_N)_{N \geq 1}$ une suite non negative tend vers zéro lorsque N tend vers l'infini. Comme N est inconné, on ne peut pas alors utiliser le dernier estimateur. On a,

$$v_n^* \;\; = \;\; \frac{1}{nh_n^d} \sum_{i=1}^{n} k_d\left(\frac{x - X_i}{h_n}\right)$$

est l'estimateur de la densité conditionnel $v^*(x)$ donnée $y_i \geq \mathfrak{T}_1$. Notons que dans la somme est sur i telle que $G_n(Y_i) \neq 0$. Finalement, pour estimer la fonction v sous le modèle tronquée on a l'estimateur de la densité de X de la forme

$$v_n^* \;\; = \;\; \frac{\alpha_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} k_d\left(\frac{x - X_i}{h_n}\right)$$

Par une méthode analogue, on peut obtenir un estimateur de $F(y|x)$ comme suit

$$\begin{aligned} F_n(y/x) \;\; &= \;\; \alpha_n \sum_{i=1}^{n} W_{i,n} \frac{1}{G_n(Y_i)} K_0\left(\frac{y - Y_i}{h_n}\right) \\ &= \;\; \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) k_d\left(\frac{x - X_i}{h_n}\right) K_0\left(\frac{y - Y_i}{h_n}\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) k_d\left(\frac{x - X_i}{h_n}\right)} \end{aligned}$$

4.1.2 Convergence uniforme presque s$\hat{u}$re

On suppose que $0 = a_G < a_F$ et $b_G < b_F.\mathcal{T}_i$ et (χ_i, y_i) avec $1 \leq i \leq N$ sont indépendants. On considère deux réels a et b tel que $a_F < a < b < b_F$.
Soit $\Omega_0 = \{x \in \mathbb{R}^d / \upsilon(x) > 0\}$ et Ω est un sous compact de Ω_0.

$$\gamma := \inf_{x \in \Omega} \upsilon(x) > 0.$$

- $(H1)$ Sur le paramètre de lissage

$$h_n \searrow 0 \quad \text{et} \quad nh_n^d / \log n \to_{n \to \infty} \infty$$

- $(H2)$ le noyau k_d est une densité de probabilité avec support compact de C^1
- $(H3)$ K_0 est differenteable de C^1 et assez compact
- $(H4)$ les noyaux K_d et K_0 sont vérifiant

$$\int tk_0(t)dt = 0 \quad \text{et} \quad \sum_{i=1}^{d} \int r_i k_d(r)dr = 0,$$

avec $r = (r_1, ..., r_d)^t$
- $(H5)$ la densite conjonte $f(.,.)$ est bornée, differentiable et continue.

Théorème 4.1.1. Sous les hypothèses (H1)-(H5), on a pour $p \in (0,1)$ la densité conditionnel $\inf_{x \in \Omega} f(\varepsilon_p(x)|x) > 0$

$$\sup_{x \in \Omega} \Big| \varepsilon_{p,n}(x) - \varepsilon_p(x) \Big| = O\Bigg(\max \left\{ \sqrt{\frac{\log n}{nh_n^d}}, h_n^2 \right\} \Bigg), \qquad \text{P.s}$$

Démonstration

Soit $x \in \Omega$. Comme $F(./x)$ et $F_n(./x)$ sont continues, on a $F\Big(\varepsilon_p(x)/x\Big) = F\Big(\varepsilon_{p,n}(x)/x\Big) = p$. En suite

$$\begin{aligned} \Big| F\Big(\varepsilon_{p,n}(x)/x\Big) - F\Big(\varepsilon_p(x)/x\Big) \Big| &= \Big| F\Big(\varepsilon_{p,n}(x)/x\Big) - F_n\Big(\varepsilon_p(n)/n\Big) \Big| \\ &\leq \sup_{a \leq y \leq b} \Big| F_n(y/n) - F(y/n) \Big| \end{aligned}$$

Et

$$F\Big(\varepsilon_{p,n}(x)/x\Big) - F\Big(\varepsilon_p(x)/x\Big) = \Big(\varepsilon_{p,n}(x) - \varepsilon_p(n) f\Big(\varepsilon_p^*(x)/x\Big)\Big)$$

Où $\varepsilon^*(x)$ est entre $\varepsilon_p(x)$ et $\varepsilon_{p,n}(x)$. Donc,

$$\sup_{x \in \Omega} \Big| \varepsilon_{p,n}(x) - \varepsilon_p(x) \Big| \Big| f\Big(\varepsilon_p^*(x)/x\Big) \Big| \leq \sup_{x \in \Omega} \sup_{a \leq y \leq b} \Big| F_n(y/x) - F(y/x) \Big|$$

Il suffit de montrer que

$$\begin{aligned} \sup_{x \in \Omega} \sup_{a \leq y \leq b} |F_n(y/x) - F(y/x)| &= \frac{1}{\gamma - \sup_{x \in \Omega} |v_n(x) - v(x)|} \sup_{x \in \Omega} \sup_{a \leq y \leq b} |F_{1,n}(y/x) - F_1(y/x)| \\ &+ \sup_{x \in \Omega} \sup_{a \leq y \leq b} |F(y/x)| \, |v_n(x) - v(x)| \end{aligned}$$

Pour montrer les résultats précédents, nous avons besoin des lemmes suivants.

Lemme 4.1.2.1. sous les hypothèses H1, H2, H4 et H5, on a

$$\sup_{x\in\Omega}\Big|v_n^*(x)-v^*(x)\Big|=O\bigg(\max\bigg\{\sqrt{\frac{\log n}{nh_n^d}},h_n^2\bigg\}\bigg),\quad p.s\quad n\to\infty$$

la preuve est basée sur la décomposition suivantes

$$\begin{aligned}\sup_{x\in\Omega}|v_n^*(x)-v^*(x)| &\leq \sup_{x\in\Omega}|v_n^*(x)-\mathrm{E}(v_n^*(x))|+\sup_{x\in\Omega}|\mathrm{E}(v_n^*(x))-v^*(x)|\\ &= C_1+C_2.\end{aligned}$$

Sous l'hypothése $(H2)$, On pose

$$\Delta_x=\frac{1}{nh_n^d}k_d\bigg(\frac{x-X_i}{h_n}\bigg),\qquad x\in\mathbb{R}^d$$

Ainsi

$$\mathbb{E}\Big(\Delta_x(X)\Big)\leq\frac{||v^*||_\infty}{n},\quad\text{et}\quad\mathbb{E}(\Delta_i^2)\leq\frac{M_d||v^*||_\infty}{n^2h_n^d}$$

On obtient alors d'aprés l'inégalité de Talagrand ([7], 2001) avec $t=D\sqrt{\frac{\log n}{nh_n^d}}$, où $D\geq 0$. Pour tout c_1 et c_2

$$\mathbb{P}\bigg\{\sup_{\theta_n\in F_1}\sum_{i=1}^n[\Delta_x(X_i)-\mathbb{E}\Delta_x(X_1)]\geq\varepsilon_0\sqrt{\frac{\log n}{nh^d}}\bigg\}$$

$$\leq C_1\exp\Bigg\{-\frac{1}{C_1}\varepsilon_0\frac{1}{||\vartheta^*||_\infty}\sqrt{\frac{\log n}{nh_n^d}}n\log\Bigg[1+\frac{\frac{\varepsilon_0\sqrt{\frac{\log n}{nh^d}}}{n}}{C_1||\vartheta^*||_\infty\Big[\frac{\sqrt{n}}{nh_n^{d/2}}+\frac{\sqrt{\log C_2h_n^{d/2}}}{n}\Big]^2}\Bigg]\Bigg\}$$

$$\leq C_1\exp\Bigg\{-\frac{1}{C_1}\varepsilon_0\frac{1}{||\vartheta^*||}\sqrt{\frac{n\log n}{h_n^d}}\frac{1}{n}\varepsilon_0\sqrt{\frac{\log n}{nh_n^d}}\frac{nh_n^d}{C_1||\vartheta||_\infty}\Bigg\}$$

$$\leq C_1 n^{-\frac{\varepsilon_0^2}{C_1^2||v^*||_\infty^2}}$$

Puisque $\log(1+t)\leq t$, pour n large et le choix de ϵ_0, on trouve $0(n^{-3/2})$. On a sous le lemme de Borel Cantelli et (H_1)

$$C_1 \;=\; 0\left(\sqrt{\frac{\log n}{nh_n^d}}\right) \tag{4.1}$$

Pour le deuxième terme, on obtient sous le développement de Taylor et les hypothèses (A4) et (A5)

$$C_2 \;=\; 0(h_n^2) \tag{4.2}$$

En utilisant (4.1) et (4.2), pour achever le résultat.

Lemme 4.1.2.2. sous les hypothèses H1, H2, H4 et H5, on a

$$\sup_{x\in\Omega}\sup_{a\leq\leq b}\Big|F_{1,n}(x,y)-F_1(x,y)\Big| = O\bigg(\max\bigg\{\sqrt{\frac{\log n}{nh_n^d}}, h_n^2\bigg\}\bigg), \qquad \text{p.s}, \qquad n\to\infty$$

On a,

$$\begin{aligned}\sup_{x\in\Omega}\sup_{a\leq y\leq b}\Big|F_{1,n}(x,y)-F_1(x,y)\Big| &= \sup_{x\in\Omega}\sup_{y\in[a,b]}\Big|F_{1,n}(x,y)-\widetilde{F}_{1,n}(x,y)\Big| \\ &+ \sup_{x\in\Omega}\sup_{y\in[a,b]}\Big|\widetilde{F}_{1,n}(x,y)-\mathbb{E}\Big[\widetilde{F}_{1,n}(x,y)\Big]\Big| \\ &+ \sup_{x\in\Omega}\Big|\mathbb{E}\Big[\widetilde{F}_{1,n}(x,y)\Big]-F_1(x,y)\Big| \\ &= S_1+S_2+S_3\end{aligned}$$

Pour le premier terme S_1
D'aprés l'hypothèse (H3), K_0 est borné par 1

$$\Big|F_{1,n}(x,y)-\widetilde{F}_{1,n}(x,y)\Big| \leq \bigg\{\frac{|\alpha_n-\alpha|}{G_n(a_F)} + \frac{\alpha}{G_n(a_F)G(a_F)}\sup_{a\leq y\leq b}\Big|G_n(y)-G(y)\Big|\bigg\}\Big|v_n^*(n)\Big|$$

On a $|\alpha_n-\alpha| = 0\Big(n^{-1/2}\Big)$, d'aprés le théorme (3.2) dans ([15]). En plus, $G_n(a_F)\longrightarrow G(a_F)>0$. On autre, $\sup_{a\leq y\leq b}|G_n(y)-G(y)| = 0(n^{-1/2})$ voir le remarque (6) dans ([24]) on obtient directement

$$S_1 \;=\; 0(n^{-1/2}) \tag{4.3}$$

Le deuxième terme S_2, On pose

$$\eta_y(.) = \frac{1}{G(y)}K_0\left(\frac{y-.}{h_n}\right), \qquad y\in\mathbb{R}$$

Et

$$\Delta_{x,y}(r,w) = \frac{1}{nh_nG(y)}k_d\bigg(\frac{x-r}{h_n}\bigg)K_0\bigg(\frac{y-w}{h_n}\bigg), \quad \text{tel que} \quad x\in\mathbb{R}^d, \quad y\in\mathbb{R}$$

En reprenant les calcules éffectués pour pouver

$$S_2 = 0\bigg(\sqrt{\frac{1}{n}}\bigg)$$

Le troisiéme terme S_3,

$$\mathbb{E}(\widetilde{F}_{1,n}(x,y)) = \mathbb{E}\bigg[\frac{1}{h^d}k_d\bigg(\frac{x-X_1}{h_n}\bigg)\mathbb{E}\bigg(\frac{\alpha}{G(y_1)}K_0\bigg(\frac{y-y_1}{h_n}\bigg)/X_1\bigg)$$

On remarque que

$$\begin{aligned}\mathbb{E}\biggl(\frac{\alpha}{G(y_1)}K_0\biggl(\frac{y-y_i}{h_n}\biggr)/X_1\biggr)\biggr] &= \int \frac{\alpha}{G(y_1)}K_0\biggl(\frac{y-z}{h_n}\biggr)f^*(z/X_1)dz \\ &= \int K_0\biggl(\frac{y-z}{h_n}\biggr)f(z/X_1)dz \\ &= \int k_0(s)F(y-sh_n/X_1)ds\end{aligned}$$

Il suffit alors de développer la fonction $\widetilde{F}_{1,n}$ au voisinage de x et par intégration par partie, on a sous les hypothéses $(H4)$ et $(H5)$

$$\begin{aligned}\mathbb{E}\widetilde{F}_{1,n}(x,y)-F_1(x,y) &= \int\int K_d(u)K_0(v)\Bigl[F_1(x-uh_n,y-vh_n)-F_1(x,y)\Bigr]dudv \\ &= 0(h_n^2)\end{aligned}$$

Lemme 4.1.2.3. sous les hypothèses H1, H2, H4 et H5, on a

$$\sup_{x\in\Omega}\Bigl|v_n(x)-v(x)\Bigr| = O\biggl(\max\biggl\{\sqrt{\frac{\log n}{nh_n^d}},h_n^2\biggr\}\biggr), \quad p.s \quad n\to\infty$$

La démonstration est similaire à celle que le cas dans le lemme précédent.

4.1.3 Normalité asymptotique

On garde les mémes notations du section précédente et on ajoute l'hypothèse ci-dessous permettant de trouver la méme vitesse de convergence que le cas i.i.d

– $H6$ Le paramétre de lissage h_n

$$h_n \searrow 0 \quad \text{et} \quad nh_n^{d+1}/\log n \to_{n\to\infty} \infty$$

– $H7$ Le paramétre h_n satisfait

$$h_n \searrow 0 \quad \text{et} \quad nh_n^{d+4} \to 0$$

Théorème 5.1 Sous ces hypothéses (H2)-(H7), on a pour tout $p \in (0.1)$ et pour tout $x \in \Omega_0$ tel que $f_n(\zeta_p(x)/x) \neq 0$

$$\sqrt{nh_n^d}(\zeta_{p,n}(x)-\zeta_p(x)) \to^D \mathcal{N}\left(0,\frac{\sigma^2(x,\zeta_p(x))}{f(\zeta_p(x)/x)}\right),$$

où

$$\sigma^2(x,y) = k\frac{\sum_1(x,y)v^2(x)+\sum_2(x)F_1^2(x,y)-2\sum_1(x,y)F_1(x,y)v(x)}{\alpha v^4(x)}$$

avec

$$k=\int k_d^2(r)dr, \quad \sum_1(x,y)=\int_{-\infty}^{y}\frac{f(x,y)}{G(y)}ds \quad \text{et} \quad \sum_2(x)=\int_{\mathbb{R}}\frac{f(x,s)}{G(s)}ds$$

Démonstration :
On pose $F(\zeta(x)|x) = p = F_n(\zeta_{p,n}(x)|x)$, d'aprés le développement de Taylor, on a

$$F(\zeta_{p,n}(x)|x) - F(\zeta_p(x)|x) = (\zeta_{p,n}(x) - \zeta_p(x))f(\zeta_p^*(x)|x)$$

Où $\zeta_p^*(x)$ est entre $\zeta_{p,n}(x)$ et $\zeta_p(x)$. Donc, d'aprés la continuité de $f(.|y)$ et le Théorème précédente implique que la convergence en probabilité du doménateur $f(\zeta_p(x)|x)$.
Soit

$$\begin{aligned}
\frac{F_{1,n}(x,y)}{\alpha_n} - \frac{F_1(x,y)}{\alpha} &= \frac{F_{1,n}(x,y)}{\alpha_n} - \frac{\widetilde{F}_{1,n}(x,y)}{\alpha} + \frac{\widetilde{F}_{1,n}(x,y)}{\alpha} - \frac{\mathbb{E}(\widetilde{F}_{1,n}(x,y))}{\alpha} \\
&+ \frac{\mathbb{E}(\widetilde{F}_{1,n}(x,y))}{\alpha} - \frac{F(x,y)}{\alpha} \\
&= \Lambda_{n1}(x,y) + \Lambda_{n2}(x,y) + \Lambda_{n3}(x,y)
\end{aligned}$$

D'aprés (H1), (H2), (H4) et (H5), pour tout x et y

$$\begin{aligned}
\sqrt{nh_n^d}\Lambda_{n1}(x,y) &= \sqrt{nh_n^d}\Big[\frac{1}{nh_n}\sum_{i=1}^n \frac{1}{G_n(Y_i)} k_d\Big(\frac{x-X_i}{h_n}\Big)K_0\Big(\frac{y-Y_i}{h_n}\Big) \\
&- \frac{1}{nh_n}\sum_{i=1}^n \frac{1}{G(Y_i)} k_d\Big(\frac{x-X_i}{h_n}\Big)K_0\Big(\frac{y-Y_i}{h_n}\Big)\Big] \\
&\leq \; \leq \sqrt{nh_n^d}\frac{\sup_y |G_n(Y) - G(Y)|}{G(a_F)G_n(a_F)}\upsilon_n^*(x) \\
&\longrightarrow 0_p(1) \qquad \text{quand n} \to \infty
\end{aligned}$$

à l'aide (4.1.2.3), on obtient

$$\begin{aligned}
\sqrt{nh_n^d}\Lambda_{n3}(x,y) &= \alpha^{-1}\sqrt{nh_n^d}\Big\{\mathbb{E}\Big[\widetilde{F}_{1,n}(x,y)\Big] - F_1(x,y)\Big\} \\
&= O\Big(\sqrt{nh_n^{d+4}}\Big)
\end{aligned}$$

Lemme 4.1.3.1. Sous les hypothéses $(K), (D1)$ et $(H1 : a)$, Nous avons pour tout x, $\upsilon(x) > 0$ et $y < b_i$

$$\sqrt{nh_n^d}(\Gamma_{n2}(x,y), \Lambda_{n2}(x,y))^T \to^D \mathcal{N}(0, \mu^{-1}k\sum(x,y))$$

Tel que

$$\sum(x,y) = \begin{pmatrix} \sum_1(x,y) & \sum_1(x,y) \\ \sum_1(x,y) & \sum_2(x,y) \end{pmatrix}$$

Démonstration du lemme 4.1.3.1 En appliquant le développement de Taylor et le changement de variable pour

$$\begin{aligned}
Var\left[\sqrt{nh_n^d}\Lambda_{n2}(x,y)\right] &= \frac{1}{nh_n^d}\times nVar\left[\frac{1}{G_n(Y_i)}k_d\Big(\frac{x-X_i}{h_n}\Big)K_0\Big(\frac{y-Y_i}{h_n}\Big)\right] \\
&= \frac{1}{\alpha h_n^d}\int\int\frac{1}{G(s)}k_d^2\left(\frac{x-r}{h_n}\right)K_0^2\left(\frac{y-s}{h_n}\right)f(r,s)drds \\
&- \frac{1}{\alpha^2 h_n^d}\left\{\int\int k_d\left(\frac{x-r}{h_n}\right)K_0\left(\frac{y-s}{h_n}\right)f(r,s)drds\right\}^2 \\
&= \frac{1}{\alpha}\int\int\frac{1}{G(s)}k_d^2(u)K_0^2\left(\frac{y-s}{h_n}\right)f(x,s)duds+O(h_n) \\
&= \frac{k}{\alpha}\sum_1(x,y)+o(1)
\end{aligned}$$

De la méme manière,

$$\begin{aligned}
Cov\left[\sqrt{nh_n^d}\Lambda_{n2}(x,y),\sqrt{nh_n^d}\Gamma_{n2}(x,y)\right] &= \frac{1}{\alpha h_n}\int\int\frac{1}{G(s)}k_d^2\left(\frac{x-r}{h_n}\right)K_0\left(\frac{y-s}{h_n}\right)f(r,s)drds+o(1) \\
&= \frac{1}{\alpha}\int\int\frac{1}{G(s)}k_d^2(u)K_0\left(\frac{y-s}{h_n}\right)f(x,s)duds+o(1) \\
&= \frac{k}{\alpha}\sum_1(x,y)+o(1)
\end{aligned}$$

On pose $c=(c_1,c_2)^t$

$$\Delta_n(x,y)=\sqrt{nh_n^d}[c_1\Lambda_{n2}(x,y)+c_2\Gamma_{n2}(x,y)]=\sum_{i=1}^n\Delta_{ni}(x,y)$$

où $\Delta_{ni}(x,y)$ est réellement indépendant. Soit $\rho_{ni}^3(x,y)=\mathbb{E}\left[\Big|\Delta_{ni}(x,y)\Big|^3\right]$
Puis,

$$\rho_{ni}^3(x,y)\leq 4\left(\frac{h_n^d}{n}\right)^{3/2}\left\{c_1^3\mathbb{E}\left[\left|\frac{1}{h_n^dG(Y_1)}k_d\left(\frac{x-X_1}{h_n}\right)K_0\left(\frac{y-Y_1}{h_n}\right)\right|^3\right]+c_2^3\mathbb{E}\left[\left|\frac{1}{h_nG(Y_1)}k_d\left(\frac{x-X_1}{h_n}\right)\right|^3\right]\right\}$$

Ce que implique

$$\rho_n^3(x,y)=\sum_{i=1}^n\rho_{ni}^3(x,y)=O(n^{-1/2}h_n^{3d/2})=o(1)$$

En outre,

$$s_n^2(x,y)=Var\left\{\sqrt{nh_n^d}[c_1\Lambda_{n2}(x,y)+c_2\Gamma_{n2}(x,y)]\right\}\to^{n\to\infty}\alpha^{-1}kc^t\sum(x,y)c>0$$

Pour tout $c\neq 0$, $\upsilon(x)>0$, $\lim\limits_{n\to\infty}\rho_n(x,y)/s_n(x,y)\to 0$. le résultat est donc une conséquence directe par le Théoème Berry-Eséen (voir ([25], p322). On suit les mêmes étapes pour démonter

$$\Gamma_{n2}(x)=\frac{1}{nh_n^d}\sum_{i=1}^n\frac{1}{G_n(Y_i)}K_d\Big(\frac{x-X_i}{h_n}\Big)-\mathbb{E}\left(\frac{1}{h_n^d}\frac{1}{G_n(Y_1)}K_d\Big(\frac{x-X_1}{h_n}\Big)\right)\longrightarrow\mathcal{N}(0,1)$$

Et

$$Var(\sqrt{nh_n^d}\Gamma_{n2}(x)) = \alpha^{-1}k\sum_2(x) + 0(1)$$

4.2 Cas de mélange

Pour pouvoir étendre les résultats obtenus ci-dessus, il est nécessaire d'introduire une structure probabiliste qui permette de contrôler la dépendance entre les variables constituant l'échantillon statistique. Une manière naturelle de faire consiste à introduire une condition d'indépendance asymptotique. Nous ferons içi l'hypothèse de mélange.

4.2.1 Modèle

Pour éviter tout confusion, nous allons noter $(X_i, Y_i, T_i); 1 \leq i \leq n, n < N$, la sous suite observée (i.e $Y_i \geq T_i$). La vraie taille n de l'échantillon observé est un variable aléatoire distribuée selon la loi Binomiale de paramètre N et μ où $\mu = \mathbb{P}(Y \geq T)$. Il est clair que si $\mu = 0$ aucune donnée n'est observée. Pour cela, nous supposons, dorénavant, que $\mu \neq 0$, d'aprés la loi fort de grand nombres on a, lorsque N tend vers ∞,

$$\widetilde{\mu}_n = \frac{n}{N} \to \mu, \quad p.s.$$

En utilisant les poids de Ould Said-Lemadani ([17]), nous obtenons l'estimateur de la fonction de répartition de Y sachant $X = x$ donné par

$$F_n(y/x) = \mu_n \sum_{i=1}^{n} \widetilde{W}_{i,n}(x) \frac{1}{G_n(Y_i)} H\Big(\frac{y - Y_i}{h_n}\Big)$$

$$= \frac{\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\Big(\frac{x - X_i}{h_n}\Big) H\Big(\frac{y - Y_i}{h_n}\Big)}{\sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\Big(\frac{x - X_i}{h_n}\Big)}$$

$$= \frac{F_{1,n}(x, y)}{\upsilon_n(x)},$$

où H est un fonction de répartition définie sur $\mathbb{R}$ et

$$F_{1,n}(x, y) = \frac{\mu_n}{nh_n^d} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\Big(\frac{x - X_i}{h_n}\Big) H\Big(\frac{y - Y_i}{h_n}\Big) \tag{4.4}$$

A partir de cet estimateur, on peut trouver l'estimateur de la densité conditionnelle de Y sachant $X = x$ définie par

$$f(y/.) = \frac{\partial F(y/.)}{\partial y}.$$

Cet estimateur est donné par

$$f_n(y/x) = \frac{f_{1,n}(x, y)}{\upsilon_n(x)},$$

où

$$f_{1,n}(x,y) = \frac{\mu_n}{nh_n^{d+1}} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K_d\Big(\frac{x-X_i}{h_n}\Big) H^{(1)}\Big(\frac{y-Y_i}{h_n}\Big)$$

Est l'estimateur de $f(x,y) = \dfrac{\partial F_1(x,y)}{\partial y}$ tel que $H^{(1)}$ est la dérivée de H.

4.2.2 Convergence uniforme presque s$\hat{u}$re

Dans que se suit, on s'intéresse au cas de covariable notée X univariée (i.e d=1). On note X de x et K de K_1. Supposons que $0 = a_G < a_F$ et $b_G < b_F$. Nous considérons deux nombres réels a et b tel que $a_F < a < b < b_F$. Soit Ω est un sous compact de $\Omega_0 = \{x \in \mathbb{R}/\upsilon(n) > 0\}$ et $\gamma := \inf_{x\in\Omega} \upsilon(x) > 0$. Considérons les hypothèses suivantes :

- $(K1)$ K est une densité de probabilité bornée, à valeur positive et continue au sens de Holder d'exposant $\beta > 0$ et satisfait
$$|u|K(u) \to 0 \qquad lorsque \qquad ||u|| \to +\infty$$
- $(K2)$ H est une fonction de répartition de densité de probabilité $H^{(1)}$ de classe C^1 qui est positive, bornée et a un support compact. Elle est continue au sens de Hölder d'exposant β.
- $(K3)$ $i)H^1$ et K sont deux noyaux de second ordre,
  $ii)$ $\int K^2(r)dr < \infty$.
- $(M1)$ $\{(X_i,Y_i), i \geq 1\}$ est une suite de variables aléatoires stationnaires et α-mélangeantes de coefficient $\alpha(n)$.
- $(M2)$ $\{T_i, i \geq 1\}$ est une suite de variables de troncature i.i.d et indépendante de $\{(X_i,Y_i), i \geq 1\}$ de fonction de répartition commune et continue G.
- $(M3)$ Il existe $\nu > 5 + 1/\beta$ pour $\beta > 1/7$ tel que $\alpha(n) = O(n^{-\nu})$.
- $(D1)$ La densité conditionnelle $\upsilon^*(.)$ est deux fois continûment différentiable.
- $(D2)$ La densité conditionnelle jointe $\upsilon^*(.,.)$ de (X_i, Y_j) existe et satisfait
$$\sup_{r,s} |\upsilon^*(r,s) - \upsilon^*(r)\upsilon^*(s)| \leq C < \infty,$$
  tell que C est une constante ne dépendant pas de (i,j).
- $(D3)$ La densité conditionnelle jointe de (X_i,Y_i,X_j,Y_j) et la densité conditionnelle jointe de (X_i,Y_i) notées respectivement par $f^*(.,.,.,.)$ et $f^*_{i,i}(.,.)$ existent et satisfont pour tout constante C
$$\sup_{r,s,t,u} |f^*(r,s,t,u) - f^*_{i,i}(r,s)f^*_{i,i}(t,u)| \leq C < \infty.$$
- $(D4)$ La densité jointe $f(.,.)$ est bornée et deux fois continument différentiable.
- $(D5)$ La densité marginale $\upsilon(.)$ est localement lipschitzienne sur Ω_0.
  La fenêtre $h_n := h$ satisfait :
- $(H1)$
$$h \downarrow 0, \quad \frac{\log n}{nh} \to 0 \qquad et \quad h = 0(1/\log n), \quad lorsque \quad n \to \infty.$$

– $(H2)$

$$Cn^{\frac{(3-\nu)\beta}{\beta(\nu+1)+4\beta+1}+\eta} < h < C^{'}n^{\frac{1}{1-\nu}},$$

où η satisfait

$$\frac{(2)\beta}{\beta(\nu+1)+4\beta+1} < \eta < \frac{(\nu-3)\beta}{\beta(\nu+1)+4\beta+1} + \frac{1}{1-\nu}$$

et β et ν vérifient la condition $(M3)$.

Théorème 4.2.1. Sous ces derniers hypothèses et pour tout $p \in (0,1)$ si la fonction q_p satisfait pour $\epsilon > 0$ tel que

$$\forall \eta_p : \Omega \to \mathbb{R}, \quad \sup_{x\in\Omega} |q_p(x) - \eta_p(x)| \geq \epsilon \Rightarrow \sup_{x\in\Omega} |F(q_p(x) - F(\eta_p(x))| \geq \beta$$

On a

$$\lim |q_{p,n}(x) - q_p(x)| = 0, \qquad \text{p.s}$$

De plus, on a

$$\sup_{x\in\Omega} \Big| q_{p,n}(n) - q_p(n) \Big| = O\Bigg(\max\left\{ \sqrt{\frac{\log n}{nh}}, h^2 \right\} \Bigg), \quad p.s \quad n \to \infty$$

Preuve

Pour montrer les résultats precédents ,nous avons besoin des lemmes suivants :

Lemme 4.2.2.1. Sous les hypothèses (K1), (K3), (M), (D1), (D2) et (H), on a

$$\sup_{x\in\Omega} \Big| \upsilon_n^*(x) - \upsilon^*(x) \Big| = O\Bigg(\max\left\{ \sqrt{\frac{\log n}{nh}}, h^2 \right\} \Bigg), \quad p.s, \qquad n \to \infty$$

Lemme 4.2.2.2. Sous les hypothèses (M), ona

$$\sup_{x\in\Omega} \Big| \mu_n - \mu \Big| = O\Bigg(\sqrt{\frac{\log\log n}{n}} \Bigg), \quad p.s, \qquad n \to \infty$$

Lemme 4.2.2.3. Sous les hypothèses du lemme (4.2.2.1) et (K2), o na

$$\sup_{x\in\Omega} \sup_{y\in[a,b]} \Big| F_{1,n}(x,y) - \widetilde{F}_{1,n}(x,y) \Big| = O\Bigg(\sqrt{\frac{\log\log n}{n}} \Bigg), \quad p.s, \qquad n \to \infty$$

Lemme 4.2.2.4. Sous les hypothèses (K), (M), (D3), (D4) et (H) on a

$$\sup_{x\in\Omega} \sup_{y\in[a,b]} \Big| \widetilde{F}_{1,n}(x,y) - \mathbb{E}\Big[\widetilde{F}_{1,n}(x,y) \Big] \Big| = O\Bigg(\sqrt{\frac{\log n}{nh}} \Bigg), \quad p.s, \qquad n \to \infty$$

Lemme 4.2.2.5. Sous les hypothèses (K3) et (D), on a

$$\sup_{x\in\Omega} \Big| \mathbb{E}\Big[\widetilde{F}_{1,n}(x,y) \Big] - F_1(x,y) \Big| = O(h^2), \quad p.s, \qquad n \to \infty$$

Lemme 4.2.2.6. Sous les hypothèses du lemme (4.2.2.1) et (D5), on a

$$\sup_{x\in\Omega}\Big|\upsilon_n(x)-\upsilon(x)\Big| = O\bigg(\max\bigg\{\sqrt{\frac{\log n}{nh}}, h^2\bigg\}\bigg), \quad p.s, \qquad n\to\infty$$

Démonstration du lemme 4.2.2.1

on a :

$$\begin{aligned}\sup_{x\in\Omega}|v_n^*(x)-v^*(x)| &\leq \sup_{x\in\Omega}|v_n^*(x)-\mathrm{E}(v_n^*(x))| + \sup_{x\in\Omega}|\mathrm{E}(v_n^*(x))-v^*(x)| \\ &= \mathrm{T}_{1n}+\mathrm{T}_{2n}.\end{aligned}$$

Nous commençons par étudier le terme variance T_{1n}. L$'$idée consiste à utiliser 1$'$ine|galité exponentielle prenant en compte la structure a-mélangeante. L$'$ensemble compact Ω peut être recouvert par un nombre fini dintervalle l_n de longueur $\omega_n = (n^{-1}h^{1+2\beta})^{\frac{1}{2\beta}}$, où β est 1$'$exposant de Hölder. On note $\mathrm{I}_k := \mathrm{I}(x_k,\ \omega_n); k = 1,\ \cdots,\ l_n$, l'intervalle centré en x_k. Comme Ω est borné, alors il existe une constante C tel que $\omega_n l_n \leq \mathrm{C}$. Pour tout x dans Ω, il existe I_k qui contient x tel que $|x-x_k|\leq\omega_n$. Nous commenȩons par écrire

$$\triangle_i(x) := \frac{1}{nh}\bigg\{\mathrm{K}\bigg(\frac{x-\mathrm{X}_i}{h}\bigg)-\mathrm{E}\bigg[\mathrm{K}\bigg(\frac{x-\mathrm{X}_i}{h}\bigg)\bigg]\bigg\}.$$

On a

$$\begin{aligned}\sum_{i=1}^n\triangle_i(x) &= \sum_{i=1}^n\frac{1}{nh}\bigg\{\mathrm{K}\bigg(\frac{x-\mathrm{X}_i}{h}\bigg)-\mathbb{E}\bigg[\mathbb{K}\bigg(\frac{x-\mathbb{X}_i}{h}\bigg)\bigg]\bigg\} \\ &= v_n^*(x)-\mathbb{E}(v_n^*(x)) \\ &= \{(v_n^*(x)-v_n^*(x_k))-(\mathbb{E}[v_n^*(x)]-\mathbb{E}[v_n^*(x_k)])\}+(v_n^*(x_k)-\mathbb{E}[v_n^*(x_k)]) \\ &= \sum_{i=1}^n\widetilde{\triangle}_i(x)+\sum_{i=1}^n\triangle_i(x_k).\end{aligned}$$

D'où

$$\begin{aligned}\sup_{x\in\Omega}\bigg|\sum_{i=1}^n\triangle_\iota(x)\bigg| &\leq \max_{1\leq k\leq l_n}\sup_{x\in\mathrm{I}_k}\bigg|\sum_{i=1}^n\widetilde{\triangle}_\iota(x)\bigg|+\max_{1\leq k\leq l_n}\bigg|\sum_{i=1}^n\triangle_\iota(x_k)\bigg| \\ &= \mathrm{S}_{1n}+\mathrm{S}_{2n}\end{aligned}$$

On a sous l'hypothèses (Kl),

$$
\begin{aligned}
\sup_{x\in \mathrm{I}_k}\Big|\sum_{i=1}^{n}\widetilde{\triangle}_\iota(x)\Big| &= \frac{1}{nh}\sum_{i=1}^{n}\Big\{\mathrm{K}\Big(\frac{x-\mathrm{X}_i}{h}\Big)-\mathrm{K}\Big(\frac{x_k-\mathrm{X}_i}{h}\Big)\Big\}\\
&- \frac{1}{h}\sum_{i=1}^{n}\Big\{\mathrm{E}\Big[\mathrm{K}\Big(\frac{x-\mathrm{X}_i}{h}\Big)-\mathrm{K}\Big(\frac{x_k-\mathrm{X}_i}{h}\Big)\Big]\Big\}\\
&\leq \frac{1}{nh}\sum_{i=1}^{n}\Big|\mathrm{K}\Big(\frac{x-\mathrm{X}_i}{h}\Big)-\mathrm{K}\Big(\frac{x_k-\mathrm{X}_i}{h}\Big)\Big|\\
&+ \frac{1}{h}\mathrm{E}\Big[\Big|\mathrm{K}\Big(\frac{x-\mathrm{X}_i}{h}\Big)-\mathrm{K}\Big(\frac{x_k-\mathrm{X}_i}{h}\Big)\Big|\Big]\\
&\leq \frac{2\sup_{x\in \mathrm{I}_k}|x-x_k|^\beta}{h^{1+\beta}}\\
&\leq \mathrm{C}\omega_n^\beta h^{-1-\beta}\\
&= \mathrm{O}((nh)^{-1/2})
\end{aligned}
$$

D'où, d'aprés $(H1)$ et pour n suffisamment grand, on obtient $\mathrm{S}_{1n}=o_{\mathrm{P}}(1)$. Etudons maintenant le terme S_{2n} dans (4.3). Sous (Kl), les variables aléatoires $\mathrm{U}_i = nh\triangle_i(x_k)$ sont centrées et bornées. L$'$utilisation de l'inégalité de Fuk-Nagaev ([10]), nous permet d$'$obtenir, pour tout $\epsilon>0$ et $r>1$

$$
\begin{aligned}
\mathrm{P}\Big\{\max_{1\leq k\leq l_n}\Big|\sum_{i=1}^{n}\triangle_i(x_k)\Big|>\epsilon\Big\} &\leq \sum_{i=1}^{l_n}\mathrm{P}\Big\{\Big|\sum_{i=1}^{n}\triangle_i(x_k)\Big|>\epsilon\Big\}\\
&\leq \mathrm{C}\omega_n^{-1}\Big\{\frac{n}{r}\Big(\frac{2r}{\epsilon nh}\Big)^{\nu+1}+\Big(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\Big)^{\frac{-r}{2}}\Big\}\\
&= \mathrm{T}_{11n}+\mathrm{T}_{12n}
\end{aligned}
$$

où

$$
s_n^2=\sum_{1\leq i\leq n}\sum_{1\leq j\leq n}|\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)|.
$$

Posons

$$
r=(\log n)^{1+\delta}, o\delta>0, \quad et \quad \epsilon=\epsilon_0\sqrt{\frac{\log n}{nh}}, pour \epsilon_0>0, \tag{4.5}
$$

on a

$$
\begin{aligned}
\mathrm{T}_{11n} &= \mathrm{C}\omega_n^{-1}\frac{n}{r}\big(\frac{r}{\epsilon nh}\big)^{\nu+1}\\
&= \mathrm{C}(n^{-1}h^{1+2\beta})^{\frac{-1}{2\beta}}\Bigg\{\frac{n}{(\log n)^{1+\delta}}\Bigg(\frac{(\log n)^{1+\delta}}{nh\epsilon_0\sqrt{\frac{\log n}{nh}}}\Bigg)^{\nu+1}\Bigg\}\\
&= \mathrm{C}n^{1-\frac{\nu+1}{2}+\frac{1}{2\beta}}h^{-(\frac{1}{2\beta}+1+\frac{\nu+1}{2})}(\log n)^{\nu(1+\delta)-\frac{\nu+1}{2}}\epsilon_0^{-(\nu+1)}
\end{aligned}
$$

Notons que sous (M3), il est facile de vérifier que $1'$hypothé se modifiée suivante (H$'$2) de (H2) est satisfaite,

$$Cn^{\frac{(3-\nu)\beta}{\beta(\nu+1)+4\beta+1}+\eta} < h < C'n^{\frac{1}{1-\nu}}, \tag{4.6}$$

où η satisfait

$$\frac{2}{\beta(\nu+1)+2\beta+1} < \eta < \frac{(\nu-3)\beta}{\beta(\nu+1)+2\beta+1} + \frac{1}{1-\nu} \tag{4.7}$$

et β et ν satisfont la condition (M3). Donc, d'aprés (4.2.1), on obtient

$$T_{11n} \leq C'(\log n)^{\nu((1+\delta)-\frac{\nu+1}{2}} n^{-1_{2\beta}(\beta(\nu+1)+2\beta+1-\frac{1}{\nu}} - L.$$

D'où, pour tout η vérifiant (4.7), T_{1nn} est bornée par le terme général d'une série finie. Etudions maintenant le terme T_{12n}, mais avant ça, nous allons étudier le comportement asymptotique de s_n^2. On a

$$\begin{aligned} s_n^2 &= \sum_{i=1}^{n} V\alpha r(U_i) + \sum_{i\neq j} |Cov(U_i,\ U_j)| \\ &=: s_n^{var} + s_n^{cov} \end{aligned}$$

Premièrement, d'après, (KI), (K3), et (Dl) et un changement de variable, nous obtenons

$$\begin{aligned} s_n^{var} &= \sum_{i=1}^{n} V\alpha r(U_i) \\ &= nV\alpha r(U_1) \\ &= n\{E(U_1^2) - E^2(U_1)\} \\ &= nE(U_i^2) \\ &= nE\left\{\left(K\left(\frac{x_k - X_1}{h}\right) - E\left[K\left(\frac{x_k - X_1}{h}\right)\right]\right)^2\right\} \\ &= n\left\{E\left[K^2\left(\frac{x_k - X_1}{h}\right)\right] - E^2\left[K\left(\frac{x_k - X_1}{h}\right)\right]\right\} \\ &= n\left\{\int K^2\left(\frac{x_k - t}{h}\right)v^*(t)dt - \left(\int K\left(\frac{x_k - t}{h}\right)v^*(t)dt\right)^2\right\} \\ &= n\left\{h\int K^2(z)v^*(x_k - zh)dz - \left(h\int K(z)v^*(x_k - zh)dz\right)^2\right\} \\ &= n\left\{h\int K^2(z)v^*(x_k)dz - \left(h\int K(z)v^*(x_k dz)\right)^2\right\} \\ &\leq Cnh - Cnh^2 \\ &= O(nh) - O(nh^2) \\ &= O(nh) \end{aligned} \tag{4.8}$$

Quant au second terme, en faisant un changement de variable, (Kl), (Ml) et (D2) entrainent,

$$\begin{aligned}
|\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| &= |\mathrm{E}(\mathrm{U}_i\mathrm{U}_j) - \mathrm{E}(\mathrm{U}_i)\mathrm{E}(\mathrm{U}_j)| = |\mathrm{E}(\mathrm{U}_i\mathrm{U}_j)| \\
&= \left|\mathrm{E}\left[\mathrm{K}\left(\frac{x_k - \mathrm{X}_i}{h}\right)\mathrm{K}\left(\frac{x_k - \mathrm{X}_j}{h}\right)\right] - \mathrm{E}\left[\mathrm{K}\left(\frac{x_k - \mathrm{X}_i}{h}\right)\right]\mathrm{E}\left[\mathrm{K}\left(\frac{x_k - \mathrm{X}_j}{h}\right)\right]\right| \\
&= \left|\int\int \mathrm{K}\left(\frac{x_k - r}{h}\right)\mathrm{K}\left(\frac{x_k - s}{h}\right)v^*(r,\ s)drds\right. \\
&- \left.\int\int \mathrm{K}\left(\frac{x_k - r}{h}\right)v^*(r)dr\int \mathrm{K}\left(\frac{x_k - s}{h}\right)v^*(s)ds\right| \\
&\leq \int\int \mathrm{K}\left(\frac{x_k - r}{h}\right)\mathrm{K}\left(\frac{x_k - s}{h}\right)|v^*(r,\ s) - v^*(r)v^*(s)|drds \\
&\leq \mathrm{C}\int\int \mathrm{K}\left(\frac{x_k - r}{h}\right)\mathrm{K}\left(\frac{x_k - s}{h}\right)drds \\
&= \mathrm{C}h^2\int\int \mathrm{K}(z)\mathrm{K}(z')dzdz' \\
&= \mathrm{O}(h^2).
\end{aligned}$$

Notons aussi que ces covariances peuvent être controlées en utilisant 1'inégalité de covariance de Davydov. On a

$$\begin{aligned}
\forall i \neq j, \quad |\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| &\leq 2\alpha\|\mathrm{U}_i\|_\infty\|\mathrm{U}_j\|_\infty \\
&\leq \mathrm{C}\alpha(|i-j|)
\end{aligned}$$

Pour évaluer s_n^{cov}, nous utilisons une théchnique développée dans Masry ([7],1986). Posons $\varphi_n = \lceil (n^{-1}h)^{\frac{-1}{\nu}} \rceil$ (où $\lceil . \rceil$ note le plus petit entier plus grand que $1'$argument), on peut écrire

$$s_n^{cov} = \sum_{0<|i-j|\leq\varphi_n} |\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| + \sum_{|i-j|>\varphi_n} |\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| \tag{4.9}$$

En utilisant la borne supérieure sur le premier terme de covariance pour obtenir

$$\sum_{0<|i-j|\leq\varphi_n} |\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| \leq \mathrm{C}nh^2\varphi_n \tag{4.10}$$

Pour le second terme,

$$\begin{aligned}
\sum_{|i-j|>\varphi_n} |\mathrm{C}ov(\mathrm{U}_i,\ \mathrm{U}_j)| &\leq \mathrm{C}\sum_{|i-j|>\varphi_n} \alpha(|i-j|) \\
&\leq \mathrm{C}n^2\alpha(\varphi_n) \qquad (4.11)
\end{aligned}$$

D'aprés ($H2$), en utilisant (M3),(4.10) et (4.11), nous obtenons

$$s_n^{cov} = \mathrm{O}(nh). \tag{4.12}$$

Finalement en utilisant (4.8) et (4.12) nous concluons directement que $s_n^2 = \mathrm{O}(nh)$. Ceci est suffisant pour étudier la quantité T_{12n}. En effet pour tout ϵ et r le développe- ment de Taylor de $\log(1+x)$ donne

$$\begin{aligned}
\mathrm{T}_{12n} &= \mathrm{C}\omega_n^{-1}\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}} \\
&= \mathrm{C}\omega_n^{-1}\exp\left\{\frac{-r}{2}\log\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)\right\} \\
&= \mathrm{C}\omega_n^{-1}\exp\left\{\frac{-r}{2}\log\left(1+\frac{\epsilon_0^2nh\log n}{rs_n^2}\right)\right\} \\
&= \mathrm{C}(n^{-1}h^{1+2\beta})^{\frac{-1}{2\beta}}\exp\left\{\frac{-r\epsilon_0^2nh\log n}{rs_n^2 2}\right\} \\
&\le \mathrm{c}^{\frac{1}{2\beta}-C'\varepsilon_0^2}h^{-1(1+\frac{1}{2\beta})} \\
&= \mathrm{c}^{\frac{1}{2\beta}-C'\varepsilon_0^2}h^{-\frac{1}{2\beta}(\beta(\nu+1)+2\beta+1)}h^{\frac{\nu+1}{2}}
\end{aligned}$$

En utilisant (H$'$2) et (M3), la dernière formule peut être considérée comme le terme général d$'$une série convergente. Comme $\sum_{n\ge1}(\mathrm{T}_{11n}+\mathrm{T}_{12n}) < \infty$ et en utilisant le lemme de Borel-Cantelli, nous avons

$$\mathrm{T}_{1n} = \mathrm{O}\left(\sqrt{\frac{\log n}{nh}}\right), \qquad p.s \quad lorsque \quad n\to\infty$$

.

Etudions maintenant le terme T_{2n}. D'aprés (K3),(Dl) et un développement de Taylor

$$\begin{aligned}
\mathrm{E}(v_n^*(x)) - v^*(x) &= \mathrm{E}\left[\frac{1}{nh}\sum_{i=1}^n \mathrm{K}\left(\frac{x-\mathrm{X}_i}{h}\right)\right] - v^*(x) \\
&= \frac{1}{h}\int \mathrm{K}(\frac{x-u}{h})v^*(u)du - v^*(x) \\
&= \frac{1}{h}\left\{h\int \mathrm{K}(z)v^*(x-zh)dz\right\} - v^*(x) \\
&= \int \mathrm{K}(z)\left(v^*(x)+zhv'^*(\tilde{x})+\frac{(zh)^2}{2}v''^*(\tilde{x})\right)dz - v^*(x) \\
&= \frac{h^2}{2}\int \mathrm{K}(z)z^2v''^*(\tilde{x})dz \\
&= \mathrm{O}(h^2),
\end{aligned}$$

où $\tilde{x} \in$ [x- zh, x], et on déduit que

$$\mathrm{T}_{2n} = \mathrm{O}(h^2)\ \mathrm{P}-p.s.\ n\to\infty$$

Remplaçant T_{1n} et T_{2n} dans l'égalité précédente, nous obtenons

$$\begin{aligned}
\sup_{x\in\Omega}|v_n^*(x)-v^*(x)| &= \mathrm{o}\Big(\sqrt{\frac{\log n}{nh}}\Big)+\mathrm{O}(h^2)\\
&= \mathrm{o}\Big(\max\Big\{\sqrt{\frac{\log n}{nh}},\ h^2\Big\}\Big), \quad p.s \quad n\to\infty
\end{aligned}$$

Démonstration du lemme 4.2.2.2 Voir Ould-Said et Tatachak $([20], 2009)$.
on pose

$$F_{1,n}(x,y)=\frac{\mu_n}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\Big(\frac{x-X_i}{h_n}\Big)H\Big(\frac{y-Y_i}{h_n}\Big)$$

Démonstration du lemme 4.2.2.3 Sous $(K2)$ la H est bornée par 1. Ainsi,on a

$$\begin{aligned}
\Big|F_{1,n}(x,y)-\widetilde{F}_{1,n}(x,y)\Big| &= \Big|\frac{\mu_n}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K_d\Big(\frac{x-X_i}{h_n}\Big)H\Big(\frac{y-Y_i}{h_n}\Big)\\
&- \frac{\mu}{nh_n^d}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K_d\Big(\frac{x-X_i}{h_n}\Big)H\Big(\frac{y-Y_i}{h_n}\Big)\Big|\\
&\leq \frac{1}{nh}\sum_{i=1}^{n}K\Big(\frac{x-X_i}{h}\Big)\Big|\frac{\mu_n}{G_n(Y_i)}-\frac{\mu}{G(Y_i)}\Big|\\
&\leq v_n^*(x)\Big|\frac{\mu_n}{G_n(a_F)}-\frac{\mu}{G(a_F)}\Big|\\
&= v_n^*(x)\Big|\frac{\mu_n}{G_n(a_F)}-\frac{\mu}{G(a_F)}-\frac{\mu}{G_n(a_F)}+\frac{\mu}{G_n(a_F)}\Big|\\
&\leq v_n^*(x)\Big\{\frac{|\mu_n-\mu|}{G_n(a_F)}+\frac{\mu\sup_{y\geq a_F}|G_n(y)-G(y)|}{G_n(a_F)G(a_F)}\Big\}
\end{aligned}$$

En utilisant les lemme $(4.2.2.1), (4.2.2.2)$ le lemme (3.4) de Liang,Li ei Oi $([18], 2009)$ et le fait que $G_n(a_F)\to G(a_F)\quad p.s\quad n\to\infty$,
nous obtenons le résultat.

Démonstration du lemme 4.2.2.4 Comme Ω et $[\alpha,\ b]$ sont des ensembles compacts, alors, il existe un recouvrement fini l_n et d_nd' intervalles $\mathrm{I}_1,\ \cdots,\ \mathrm{I}_{l_n}$ et $\mathrm{J}_1,\cdots,\ \mathrm{J}_{d_n}$ de longeurs ω_n définis dans le lemme (4.1) et $\lambda_n=(n^{-1}h^{2\beta})^{\frac{1}{2\beta}}$ et de centres $x_1,\ \cdots,\ x_{l_n}$ et $y_1,\ \cdots,\ y_{d_n}$ respectivement. Comme Ω et $[\alpha,\ b]$ sont bornés, ils existent deux constants C_1 et C_2 tels que $l_n\omega_n\leq\mathrm{C}_1$ et $d_n\lambda_n\leq\mathrm{C}_2$. Pour tout $(x,y)\in\Omega\times[\alpha,\ b]$ ils existent x_k et y_j tels que

$\|x - x_k\| \leq \omega_n$ et $\|y - y_j\| \leq \lambda_n$. Ainsi on a la décomposition suivante

$$\begin{aligned}
\sup_{x\in\Omega}\sup_{a\leq y\leq b}|\tilde{\mathrm{F}}_{1,n}(x,\ y)-\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x,\ y)]| &\leq \max_{1\leq k\leq l_n}\sup_{x\in \mathrm{I}_k}\sup_{y}|\tilde{\mathrm{F}}_{1,n}(x,\ y)-\tilde{\mathrm{F}}_{1,n}(x_k,\ y)| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}\sup_{y\in \mathrm{J}_j}|\tilde{\mathrm{F}}_{1,n}(x_k,\ y)-\tilde{\mathrm{F}}_{1,n}(x_k,\ y_j)| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}|\tilde{\mathrm{F}}_{1,n}(x_k,\ y_j)-\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x_k,\ y_j)]| \\
&+ \max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}\sup_{y\in \mathrm{J}_j}|\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x_k,\ y_j)]-\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x_k,\ y)]| \\
&+ \max_{1\leq k\leq l_n}\sup_{x\in \mathrm{I}_k}\sup_{y}|\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x_k,\ y)]-\mathrm{E}[\tilde{\mathrm{F}}_{1,n}(x,\ y)]| \\
&= \mathrm{J}_{1n}+\mathrm{J}_{2n}+\mathrm{J}_{3n}+\mathrm{J}_{4n}+\mathrm{J}_{5n}
\end{aligned}$$

Commenąons par étudier J_{1n}. Les hypothèses (Kl) et (K2) donnent

$$\begin{aligned}
\sup_{x\in \mathrm{I}_k}\sup_{y}|\tilde{\mathrm{F}}_{1,n}(x,\ y)-\tilde{\mathrm{F}}_{1,n}(x_k,\ y)| &= \left|\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{\mathrm{G}(\mathrm{Y}_i)}\mathrm{K}\left(\frac{x-\mathrm{X}_i}{h}\right)\mathrm{H}\left(\frac{y-\mathrm{Y}_i}{h}\right)\right. \\
&- \left.\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{\mathrm{G}(\mathrm{Y}_i)}\mathrm{K}\left(\frac{x_k-\mathrm{X}_i}{h}\right)\mathrm{H}\left(\frac{y-\mathrm{Y}_i}{h}\right)\right| \\
&= \left|\mathrm{H}\left(\frac{y-\mathrm{Y}_i}{h}\right)\right|\left|\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{\mathrm{G}(\mathrm{Y}_i)}\left(\mathrm{K}\left(\frac{x-\mathrm{X}_i}{h}\right)\right.\right. \\
&- \left.\left.\mathrm{K}\left(\frac{x_k-\mathrm{X}_i}{h}\right)\right)\right| \\
&\leq \sup_{y}\left|\mathrm{H}\left(\frac{y-\mathrm{Y}_i}{h}\right)\right|\frac{\mu||x-x_k\|}{\mathrm{G}(\alpha_{\mathrm{F}})h^{1+\beta}} \\
&\leq \sup_{y}\left|\mathrm{H}\left(\frac{y-\mathrm{Y}_i}{h}\right)\right|\frac{\mu \mathrm{C}\omega_n^{\beta}}{\mathrm{G}(\alpha_{\mathrm{F}})h^{1+\beta}} \\
&\leq \mathrm{C}'h^{-1-\beta}(n^{-1}h^{1+2\beta})^{\frac{\beta}{2\beta}} \\
&= \mathrm{O}((nh)^{-1/2}).
\end{aligned}$$

A partir de (Hl), nous obtenons

$$\sqrt{\frac{nh}{\log n}}\sup_{x\in\Omega}\sup_{a\leq y\leq b}|\tilde{\mathrm{F}}_{1,n}(x,\ y)-\tilde{\mathrm{F}}_{1,n}(x_k,\ y)| = o(1)$$

.

En utiliant les mêmes étapes, on peut montrer la même chose pour J_{5n}. Etudions maintenant le terme J_{2n}

$$\begin{aligned}
\sup_{y\in J_j}|\tilde{F}_{1,n}(x_k,\ y)-\tilde{F}_{1,n}(x_k,\ y_j)| &= \left|\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y-Y_i}{h}\right)\right.\\
&- \left.\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_j-Y_i}{h}\right)\right|\\
&= K\left(\frac{x_k-X_i}{h}\right)\left|\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}\left(H\left(\frac{y-Y_i}{h}\right)\right.\right.\\
&- \left.\left.H\left(\frac{y_j-Y_i}{h}\right)\right)\right|\\
&\leq K\left(\frac{x_k-X_i}{h}\right)\frac{\mu\|y-y_j\|}{G(\alpha_F)h^{1+\beta}}\\
&\leq K\left(\frac{x_k-X_i}{h}\right)\frac{\mu C\lambda_n^{\beta}}{G(\alpha_F)h^{1+\beta}}\\
&\leq C'h^{-1-\beta}(n^{-1}h^{2\beta})^{\frac{\beta}{2\beta}}\\
&= O((nh^2)^{-1/2}).
\end{aligned}$$

D'aprés (Hl), on obtient

$$\sqrt{\frac{nh}{\log n}}\sup_{x\in\Omega}\sup_{a\leq y\leq b}|\tilde{F}_{1,n}(x_k,\ y)-\tilde{F}_{1,n}(x_k,\ y_j)|=o(1)$$

En utilisant les mêmes étapes, on peut montrer la même chose pour J_{4n}. Passons maintenant au dernier terme J_{3n}, pour tout $\epsilon>0$, on a

$$P\{\max_{1\leq k\leq l_n}\max_{1\leq j\leq d_n}|\tilde{F}_{1,n}(x_k,\ y_j)-E[\tilde{F}_{1,n}(x_k,\ y_j)]|>\epsilon\}\leq l_n d_n P\{|\tilde{F}_{1,n}(x_k,\ y_j)-E[\tilde{F}_{1,n}(x_k,\ y_j)]|>\epsilon\}$$

. Posons, pour tout $i\geq 1$

$$\begin{aligned}
\Psi_n(x_k,\ y_j) &= \frac{\mu}{nh}\left\{\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_j-Y_i}{h}\right)\right.\\
&- \left.E\left[\frac{1}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_j-Y_i}{h}\right)\right]\right\}
\end{aligned}$$

Sous (Kl) et (K2), les variables aléatoires $V_i=nh\Psi(x_k,\ y_j)$ sont centrées et bornées par $\dfrac{2\mu M_0M_1}{G(a_F}=:C<\infty$. Alors, en appliquant l'inégalité de Fuck-Nagaev ([10]), on obtient pour tout $\epsilon > 0$ et $r > 1$

$$\begin{aligned}
\mathrm{P}\left\{\max_{1\le k\le l_n}\max_{1\le j\le d_n}\left|\sum_{i=1}^{n}\Psi_n(x_k,\ y_j)\right|>\epsilon\right\} &= \mathrm{P}\left\{\max_{1\le k\le l_n}\max_{1\le j\le d_n}\left|\sum_{i=1}^{n}\mathrm{V}_i\right|>nh\epsilon\right\}\\
&\le l_nd_n\left\{\frac{n}{r}\Big(\frac{r}{\epsilon nh}\Big)^{\nu+1}+\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}}\right\}\\
&\le \mathrm{C}_1\mathrm{C}_2(\omega_n\lambda_n)^{-1}\left\{\frac{n}{r}\Big(\frac{r}{\epsilon nh}\Big)^{\nu+1}+\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}}\right\}\\
&= \mathrm{C}(n^{-1}h^{1+2\beta})^{\frac{-1}{2\beta}}(n^{-1}h^{2\beta})^{\frac{-1}{2\beta}}\left\{\frac{n}{r}\Big(\frac{r}{\epsilon nh}\Big)^{\nu+1}\right.\\
&+ \left.\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}}\right\}\\
&\le \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\frac{n}{r}\left(\frac{r}{\epsilon nh}\right)^{\nu+1}\\
&+ \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}}\\
&= \mathrm{J}_{31n}+\mathrm{J}_{32n} \qquad (4.13)
\end{aligned}$$

Où

$$s_n^2=\sum_{1\le i\le n}\sum_{1\le j\le n}|\mathrm{C}ov(\mathrm{V}_i,\ \mathrm{V}_j)|.$$

En posant

$r=(\log n)^{1+\delta}$, où $\delta>0,\ et\ \epsilon=\epsilon_0\sqrt{\frac{\log n}{nh}}$, pour $\epsilon_0>0$

$$\begin{aligned}
\mathrm{J}_{31n} &= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\frac{n}{r}\left(\frac{r}{\epsilon nh}\right)^{\nu+1}\\
&= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\frac{n}{(\log n)^{1+\delta}}\left(\frac{(\log n)^{1+\delta\sqrt{nh}}}{\epsilon_0\sqrt{\log n}nh}\right)^{\nu}+1\\
&= \mathrm{C}n^{\frac{1}{\beta}+1-\frac{\nu+2}{2}}h^{-(\frac{1}{2\beta}+2)}h^{-\frac{\nu+2}{2}}(\log n)^{(1+\delta)\nu}(\log n)^{-\frac{\nu+2}{2}}\epsilon_0^{-(\nu+1)}\\
&= \mathrm{c}\epsilon_0^{-(\nu+1)}n^{\frac{1}{\beta}+1-\frac{\nu+2}{2}}(\log n)^{(1+\delta)\nu-\frac{\nu+2}{2}}h^{-\frac{1}{2\beta}(1+4\beta+\beta(1+\nu))}.
\end{aligned}$$

Donc, en utilisant (H2) et (Hl), nous obtenons

$$\mathrm{J}_{31n}\le \mathrm{C}\epsilon_0^{-(\nu+1)}(\log n)^{(1+\delta)\nu-\frac{\nu+1}{2}}n^{1+\frac{1}{\beta}-\frac{1+\nu}{2}}h^{-\frac{1}{2\beta}(1+4\beta+\beta(1+\nu))}.$$

Sous les conditions sur β et η, J_{31n} est le terme générale d'une série finie. Examinons maintenant le terme J_{32n}. premièrement, on commence par calculer

$$s_n^2=nV\alpha r(\mathrm{V}_1)+\sum_{i\ne j}|\mathrm{C}ov(\mathrm{V}_i,\ \mathrm{V}_j)|.$$

On a

$$
\begin{aligned}
V\alpha r(V_1) &= E[V_1^2] - E^2[V_1] \\
&= E\left[\frac{\mu^2}{G^2(Y_1)}K^2\left(\frac{x_k - X_1}{h}\right)H^2\left(\frac{y_j - Y_1}{h}\right)\right] \\
&- E^2\left[\frac{\mu}{G(Y_1)}K\left(\frac{x_k - X_1}{h}\right)H\left(\frac{y_j - Y_1}{h}\right)\right] \\
&= W_1 - W_2.
\end{aligned}
$$

On a

$$
\begin{aligned}
E\left[\frac{\mu^2}{G^2(Y_1)}H^2\left(\frac{y_j - Y_1}{h}\right)/X_1\right] &= \int \frac{\mu^2}{G^2(y_1)}H^2\left(\frac{y_j - y_1}{h}\right)f^*(y_1/X_1)dy_1 \\
&= \int \frac{\mu}{G(y_1)}H^2\left(\frac{y_j - y_1}{h}\right)f(y_1/X_1)dy_1 \\
&= E\left[\frac{\mu}{G(Y_1)}H^2\left(\frac{y_j - Y_1}{h}\right)\right]
\end{aligned}
$$

Alors

$$
\begin{aligned}
W_1 &= E\left[\frac{\mu^2}{G^2(Y_1)}K^2\left(\frac{x_k - X_1}{h}\right)H^2\left(\frac{y_j - Y_1}{h}\right)\right] \\
&= E\left[\frac{\mu}{G(Y_1)}K^2\left(\frac{x_k - X_1}{h}\right)H^2\left(\frac{y_j - Y_1}{h}\right)\right] \\
&\leq \frac{\mu}{G(\alpha_F)}E\left[K^2\left(\frac{x_k - X_1}{h}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\mu}{G(\alpha_F)}\int K^2\left(\frac{x_k - t}{h}\right)v^*(t)dt \\
&= \frac{h\mu}{G(\alpha_F)}\int K^2(z)v^*(x_k - zh)dz \\
&= \frac{h\mu}{G(\alpha_F)}\int K^2(z)v^*(x_k)dz \\
&\leq Ch \\
&= O(h)
\end{aligned}
$$

Un développement analogue donne que $W_2 = O(h^2)$, qui implique que $nV\alpha r(V_1) =$

$O(nh)$. Pour $|Cov(V_i,\ V_j)|,$ (MI), (KI), (K2), (D3) et un changement de variable donnent

$$\begin{aligned}
|Cov(V_i,\ V_j)| &= |E[V_iV_j]-E[V_i]E[V_j]| \\
&= \left|E\left[\frac{\mu^2}{G(Y_i)G(Y_j)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_i-Y_i}{h}\right)K\left(\frac{x_k-X_j}{h}\right)H\left(\frac{y_j-Y_j}{h}\right)\right]\right. \\
&- \left.E\left[\frac{\mu}{G(Y_i)}K\left(\frac{x_k-X_i}{h}\right)H\left(\frac{y_i-Y_i}{h}\right)\right]E\left[\frac{\mu}{G(Y_j)}K\left(\frac{x_k-X_j}{h}\right)H\left(\frac{y_i-Y_j}{h}\right)\right]\right| \\
&= \left|\int\int\int\int\frac{\mu^2}{G(r)G(t)}K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_i-r}{h}\right)K\left(\frac{x_k-s}{h}\right)H\left(\frac{y_i-t}{h}\right)\right. \\
&.\ \ f^*(u,r,s,t)dudrdsdt-\int\int\frac{\mu}{G(r)}K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_i-r}{h}\right)f^*_{i,i}(u,r)dudr \\
&.\ \ \left.\int\int K\left(\frac{x_k-s}{h}\right)H\left(\frac{y_i-t}{h}\right)f^*_{i,i}(s,t)dsdt\right| \\
&\le \frac{\mu^2}{G^2(\alpha_F)}\int\int\int\int K\left(\frac{x_k-u}{h}\right)H\left(\frac{y_i-r}{h}\right)K\left(\frac{x_k-s}{h}\right) \\
&.\ \ H\left(\frac{y_i-t}{h}\right)|f^*(u,r,s,t)-f^*_{i,i}(u,\ r)f^*_{i,i}(s,t)|dudrdsdt \\
&\le C\int\int\int\int K\left(\frac{x_k-r}{h}\right)H\left(\frac{y_i-r}{h}\right)K\left(\frac{x_k-s}{h}\right)H\left(\frac{y_i-t}{h}\right)dudrdsdt \\
&= Ch^4\int\int\int\int K(z)H(v)H(v')K(z')dzdvdz'dv' \\
&= O(h^4) \qquad (4.14)
\end{aligned}$$

A partir d'un résultat dans Bosq (1998, [4]), on a

$$|Cov(V_i,\ V_j)| = O(\alpha(|i-j|)). \qquad (4.15)$$

Donc, d'après(4.14) et (4.15) nous obtenons pour $\varphi_n=\lceil(n^{-1}h)^{\frac{-1}{\nu}}\rceil$ (où $\lceil.\rceil$ représente le plus petit entier plus grand que 1)

$$\begin{aligned}
\sum_{i\neq j}|Cov(V_i,\ V_j)| &= \sum_{0<|i-j|\le\varphi_n}|Cov(V_i,\ V_j)|+\sum_{|i-j|>\varphi_n}|Cov(V_i,\ V_j)|. \\
&\le \sum_{0<|i-j|\le\varphi_n}Ch^4+\sum_{|i-j|>\varphi_n}C\alpha(|i-j|) \\
&\le Cn\varphi_nh^4+Cn^2\alpha(\varphi_n).
\end{aligned}$$

D'après (H2) et (M3), on a

$$\sum_{i\neq j}|Cov(V_i,\ V_j)| = O(nh)$$

Donc $s_n^2 = \mathrm{O}(nh)$

Passant maintenant au terme J_{32n}. En prenant r et ϵ habituelles, on voit que

$$\begin{aligned}
\mathrm{J}_{32n} &= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)^{\frac{-r}{2}} \\
&= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\exp\left\{\frac{-r}{2}\log\left(1+\frac{\epsilon^2n^2h^2}{rs_n^2}\right)\right\} \\
&= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\exp\left\{\frac{-r}{2}\frac{\epsilon^2n^2h^2}{rs_n^2}\right\} \\
&\leq \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\exp\left\{\frac{-1}{2}\epsilon_0^2\log n\right\} \\
&= \mathrm{C}n^{\frac{1}{\beta}}h^{-(\frac{1}{2\beta}+2)}\exp\{\log n^{\frac{-1}{2}\in^2}0\} \\
&= \mathrm{C}n^{\frac{1}{\beta}-\frac{1}{2}\in^2}0h^{-(\frac{1}{2\beta}+2)}.
\end{aligned}$$

D'aprés (H2) et (M3), cette dernière peut être considérée comme un teme général d)une série finie. Par conséquent $\sum_{n\geq 1}(\mathrm{J}_{31n}+\mathrm{J}_{32n})<\infty$. Donc par le lemme de Borel- Cantelli le premier terme de (4.13) tend vers 0 p.s et pour n suffisament grand, on a $\mathrm{J}_{3n} = \mathrm{O}\left(\sqrt{\frac{\log n}{nh}}\right)$.

En regroupant les résultats de $\mathrm{J}_{1n},\ \mathrm{J}_{2n},\ \mathrm{J}_{3n},\ \mathrm{J}_{4n}$ et J_{5n}, on complète la démonstration du lemme. Pour le terme L_{2n}, on réprend les mêmes étapes de la preuve du terme T_{1n} dans la démonstration du lemme (4.1) et on montre que pour n suffisamment grand que

$$\mathrm{L}_{2n} = \mathrm{O}\left(\sqrt{\frac{\log n}{nh}}\right), \quad p.s, \qquad n\to\infty.$$

Finalement, par un changement de variable, un développement de Taylor, les conditions (K3) et (D5), on a

$$\begin{aligned}
\mathrm{E}[\tilde{v}_n(x)]-v(x)] &= \mathrm{E}\left(\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{\mathrm{G}(\mathrm{Y}_i)}\mathrm{K}(\frac{x-\mathrm{X}_i}{h})\right)-v(x) \\
&= \frac{1}{h}\int \mathrm{K}(\frac{x-u}{h})v(u)du - v(x) \\
&= \frac{1}{h}\{h\int \mathrm{K}(z)v(x-zh)dz\} - v(x) \\
&= \int \mathrm{K}(z)(v(x)+zhv'(\tilde{x})+\frac{(zh)^2}{2}v''(\tilde{x}))dz - v(x) \\
&= \frac{h^2}{2}\int \mathrm{K}(z)z^2v''(\tilde{x})dz,
\end{aligned}$$

où $\tilde{x}\in[x\text{–}zh,\ x]$. Alors, on obtient

$$\mathrm{L}_{3n} = \mathrm{O}(h^2), \quad \text{p.s} \quad n\to\infty.$$

En regroupant les résultats précédentes, on conclut le résultat final.
Démonstration du lemme 4.2.2.5 Voir lemme $(6.2, [17])$ dans Lemdani, Ould-Said et Poulin (2009).
Démonstration du lemme 4.2.2.6
on définit

$$\widetilde{\upsilon}_n(x) = \frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\Big(\frac{x-X_i}{h}\Big)$$

On a

$$\begin{aligned}\sup_{x\in\Omega}\Big|\upsilon_n(x)-\upsilon(x)\Big| &\leq \sup_{x\in\Omega}\Big|\upsilon_n(x)-\widetilde{\upsilon}_n(x)\Big| + \sup_{x\in\Omega}\Big|\widetilde{\upsilon}_n(x)-\mathbb{E}\Big[\widetilde{\upsilon}_n(x)\Big]\Big| \\ &+ \sup_{x\in\Omega}\Big|\mathbb{E}\Big[\widetilde{\upsilon}_n(x)\Big]-\upsilon(x)\Big| \\ &= L_{1n}+L_{2n}+L_{3n}\end{aligned}$$

Commençons par le terme L_{1n}. on a

$$\begin{aligned}\sup_{x\in\Omega}\Big|\upsilon_n(x)-\upsilon(x)\Big| &= \Big|\frac{\mu_n}{nh}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K\Big(\frac{x-X_i}{h}\Big) \\ &- \frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\Big(\frac{x-X_i}{h}\Big)\Big| \\ &\leq \frac{1}{nh}\sum_{i=1}^{n}K\Big(\frac{x-X_i}{h}\Big)\Big|\frac{\mu_n}{G_n(a_F)}-\frac{\mu}{G(a_F)}\Big| \\ &= \upsilon_n^*(x)\Big|\frac{\mu_n}{G_n(a_F)}-\frac{\mu}{G(a_F)}-\frac{\mu}{G_n(a_F)}+\frac{\mu}{G_n(a_F)}\Big| \\ &\leq \upsilon_n^*(x)\Big\{\frac{|\mu_n-\mu|}{G_n(a_F)}+\frac{\mu\sup_{y\geq a_F}|G_n(y)-G(y)|}{G_n(a_F)G(a_F)}\Big\}\end{aligned}$$

En utilisant les lemme $(4.2.2.1), (4.2.2.2)$ le lemme (3.4) de Liang,Li ei Oi $([18], 2009)$ et le fait que
$G_n(a_F)\rightarrow G(a_F) \quad p.s \quad n\rightarrow\infty$, nous obtenons

$$L_{1n} = O\Bigg(\sqrt{\frac{\log\log n}{n}}\Bigg), \quad p.s \quad n\rightarrow\infty \tag{4.16}$$

Et

$$L_{2n} = O\Bigg(\sqrt{\frac{\log n}{n}}\Bigg), \quad p.s \quad n\rightarrow\infty \tag{4.17}$$

Finalement,par un changement de variable ,un développement de Taylor,les conditions $(K3)$

et $(D5)$, on a

$$\begin{aligned}
\mathbb{E}\Big[\widetilde{\upsilon}_n(x)\Big] - \upsilon(x) &= \mathbb{E}\Bigg[\frac{\mu}{nh}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\Big(\frac{x-X_i}{h}\Big)\Bigg] - \upsilon(x) \\
&= \frac{1}{h}\int K\Big(\frac{x-u}{h}\Big)\upsilon(u)du - \upsilon(x) \\
&= \frac{1}{h}\Big\{h\int K(z)\upsilon(x-zh)dz\Big\} - \upsilon(x) \\
&= \int K(z)\Big(\upsilon(x) + zh\upsilon^{'}(\widetilde{x}) + \frac{(zh)^2}{2}\upsilon^{''}(\widetilde{x})\Big)dz - \upsilon(x) \\
&= \frac{h^2}{2}\int K(z)z^2\upsilon^{''}(\widetilde{x})dz,
\end{aligned}$$

Où $\widetilde{x} \in [x - zh, x]$. Alors,on obtient

$$L_{3n} = o(h^2) \qquad \text{p.s} \qquad n \to \infty \tag{4.18}$$

En regroupant les résultats (4.16), (4.17) et (4.18) pour conclure le résultat final.

Chapitre 5

L'estimation du modèle de troncature à gauche : Cas fonctionnel

Dans ce chapitre, on se propose d'étudier la convergence uniforme presque s$\hat{u}$re de l'estimateur à noyau du quantile conditionnel pour des données tronquées à gauche défini dans le chapitre précédent sous des conditions d'un certain type de l'indépendance dans un espace de dimension infiné.

5.1 Modèle

Soit Y et T deux variables aléatoires réelles de fonctions de répartition inconnues F et G respectivement. Soit X un vecteur aléatoire de covariable de dimension d de fonction de répartition $V(.)$ et une densité $v(.)$. Soit (X_i, Y_i, T_i), N copies iid de (X, Y, T), sous le modèle aléatoire de troncature à gauche Y et T sont observées seulement si $\{Y \geq T\}$. Pour éviter toute confusion, nous allons noter $(X_i, Y_i, T_i), \quad 1 \leq i \leq n, \quad n < N$, la sous suite observée (i.e $Y_i \geq T_i$)). La vrai taille n de l'échantillon observé est une variable aléatoire distribuée selon la loi Binomiale de Paramètre N et μ où $\mu = \mathbb{P}(Y \geq T)$. Il est clair que si $\alpha = 0$, aucun donnée n'est observée. Pour cela, nous supposons, dorénavant, que $\mu \neq 0$ d'aprés la loi forte de grand nombres on a, lorsque N tend vers ∞.

$$\widehat{\alpha}_n = \frac{n}{N} \longrightarrow \alpha, \quad p.s$$

L'estimateur de la fonction de répartition de Y sachant $X = x$ donnée par

$$\begin{aligned} F_n(y/x) &= \alpha_n \sum_{i=1}^{n} \widehat{W}_{in} \frac{1}{G_n(Y_i)} H\left(\frac{y - Y_i}{h_{n,H}}\right) \\ &= \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(\frac{||x - X_i||}{h_{n,H}}\right) H\left(\frac{y - Y_i}{h_{n,H}}\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(\frac{||x - X_i||}{h_{n,H}}\right)} \end{aligned}$$

Où K est un noyau, H est la fonction de répartition et $h_{n,K} = h_K$ (resp. $h_{n,H} = h_H$) est une suite des valeurs réels tend vers zero quand n tend vers l'infini.

Pour tout $p \in (0,1)$, l'estimateur naturel du quantile conditionnel d'ordre p est donné par

$$\zeta_{p,n} = \inf\{y : F_n(y/x) \geq p\}$$

5.2 Cas i.i.d

5.2.1 Convergence uniforme presque s$\hat{u}$re

On suppose que $0 = a_G < a_F$ et $b_G < b_F, T_i$ et (X_i, Y_i) sont indépendant et $a_F < a < b < b_F$. On note $B(x,h)$ une boule de centre x et de Rayon h, $W_i = ||x - X_i||$ et

$$\mathbb{P}\Big(X_i \in B(x,h)\Big) = \mathbb{P}(W_i \leq h) = F_x(h)$$

et ε est un sous compact de S et $\eta(u)$ est un voisinage de u

- $(H1)$ Il existe trois fonctions $g(.), \phi(.)$ (supposée croissante) et $\varepsilon_0(.)$ tel que

  i $F_x(h_K) = g(x)\phi(h_K) + 0(\phi(h_K))$

  ii pour tout $u \in [0,1]$, $\quad \lim\limits_{h_K \to \infty} \dfrac{\phi(uh_K)}{\phi(h_K)} = \lim\limits_{h_K \to \infty} \varepsilon_{h_K}(u) = \varepsilon_0(u)$

  iii $\lim\limits_{h_K \to 0} \sup\limits_{x \in \varepsilon} \Big| \dfrac{F_x(h_K)}{\phi(h_K)} - a_1 g(x) \Big| = 0, \quad$ où $\quad a_1 = K(1) - \displaystyle\int_0^1 K^{'}(t)\varepsilon_h(u)du.$

- $(H2)$ le noyau K est positive et à assez compact dans $[0,1]$ de classe C^1 sur $[0,1]$, $K(0) > 0$ et $K(1) > 0$ et sa dérivée $K^{'}$ est tel que,

$$-\infty < C_1 < K^{'}(t) < C_2 < 0, \quad \text{sur} \quad [0,1]$$

- $(H3)$la probabilité conditionnel vérifiée la condition de Hölder pout tout variable et il existe deux constentes strictement positive β et ν tel que

  i $\forall (y_1, y_2) \in \mathbb{R}^2, (x_1, x_2) \in \vartheta(x) \times \vartheta(x)$,

$$|F(y_1|x_1) - F(y_2|x_2)| \leq C_x \left(\Big|\Big| x_1 - x_2 \Big|\Big|^{\beta} + \Big| y_1 - y_2 \Big|^{\nu} \right)$$

  où $\vartheta(x)$ est un voisinage de u, C_x est un constante dépend de x

  ii $\int_{\mathbb{R}} |z|^{\nu} H^{'}(z)dz < \infty, \qquad$ H$^{'}$ est bornée.

- $(H4)$ Sur le paramétre de lissage
  - (i) $\lim\limits_{n \to \infty} h_K = \lim\limits_{n \to \infty} h_H = 0$
  - (ii) $\lim\limits_{n \to \infty} \dfrac{\log n}{n\phi_x(h_K)} = 0$
  - (iii) $\lim\limits_{n \to \infty} nh_H \sqrt{\log n} = +\infty$

Sous ces hypothèses, on obtient le résultat suivant :

5.2.2 Résultat

Théorème 5.2.1. Sous les hypothèses $(H1)-(H4)$, on a :

$$\sup_{x\in S}\left|\widehat{\zeta}_{p,n}(x)-\zeta_p(x)\right| = \circ(h_K^{\beta})+\circ(h_H^{\gamma})+\circ\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.s$$

Démonstration : La preuve de ce théorème est basée sur la décomposition suivante :

$$\forall\epsilon>0,\quad \eta(\epsilon)=\min\{F(\zeta_p(x)+\epsilon|x)-F(\zeta_p(x)|x), F(\zeta_p(x)|x)-F(\zeta_p(x)-\epsilon|x)\}$$

En suite

$$\forall\epsilon>0,\forall y>0,\quad |\zeta_p(x)-y|\leq\epsilon\Rightarrow|F(\zeta_p(x)|x)-F(y|x)|\geq\eta(\epsilon).$$

Donc,

$$\begin{aligned}\sup_{x\in}|F(\zeta_{p,n}(x)|x)-F(\zeta_p(x)|x)| &= |F(\zeta_{p,n}(x)|x)-F_n(\zeta_{p,n}(x)|x)|\\ &\leq \sup_{x\in}\sup_{y\in[a,b]}|F_n(y|x)-F(y|x)|\end{aligned}$$

Donc, il suffit de montrer

$$\sup_{x\in S}\sup_{y\in[a,b]}\left|\hat{F}_n(y|x)-F_n(y|x)\right| = \circ(h_K^{\beta})+\circ(h_H^{\gamma})+\circ\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right), \qquad p.s$$

On a,

$$\begin{aligned}F_n(y|x)-F(y|x) &= \frac{\psi_n(x,y)}{g_n(x)}-F(y|x)\\ &= \frac{1}{g_n(x)}\left\{\psi_n(x,y)-\widetilde{\psi}_n(x,y)+\widetilde{\psi}_n(x,y)-\mathbb{E}\widetilde{\psi}_n(x,y)\right\}\\ &- F(y|x)\left\{g_n(x)-\widetilde{g}_n(x)+\widetilde{g}_n(x)-\mathbb{E}\widetilde{g}_n(x)\right\}\\ &+ \frac{1}{g_n(x)}\left\{\alpha a_1\psi_n(x,y)-\widetilde{\psi}_n(x,y)-F(y|x)\left(\alpha a_1 g_n(x)-\widetilde{g}_n(x)\right)\right\}\end{aligned}$$

Pour montrer les résultats précédents, nous avons besoin des lemmes suivants.

Lemme 5.2.1. sous les hypothèses (H1) et (H2), on a

$$\sup_{x\in S}\left|\frac{1}{\phi(h_K)}\mathbb{E}\left(K\left(\frac{||x-X_i||}{h_K}\right)\right)-a_1 g(x)\right|\longrightarrow 0 \quad \text{quand} \quad n\to\infty \tag{5.1}$$

Lemme 5.2.2. sous les hypothèses (H3) (ii), (H4)((ii) et (iii)), on a

$$\sup_{x\in S}|\widetilde{g}_n(x)-\mathbb{E}\widetilde{g}_n(x)| = 0\left(\frac{\log n}{n\phi(h_K)}\right)^{1/2} \tag{5.2}$$

Lemme 5.2.3. sous les hypothèses (H2), (H4), on a

$$\sup_{x\in S}|g_n(x)-\widetilde{g}_n(x)|=0\left(n^{-1/2}\right) \tag{5.3}$$

Lemme 5.2.4. Si les hypothèses (H1) et (H2) sont satisfaits, alors

$$\sup_{x\in S}|\mathbb{E}(g_n(x))-\alpha a_1 g(x)|=0\left(n^{-1/2}\right), \qquad n\to\infty \tag{5.4}$$

Lemme 5.2.5. Sous les hypothèses (H1), (H2), on a

$$\sup_{x\in S}\sup_{y\in[a,b]}\left|\psi_n(x,y)-\widetilde{\psi}_n(x,y)\right|=0(n^{-1/2}), \quad \text{P.s}, \quad n\to\infty \tag{5.5}$$

Lemme 5.2.6. sous les hypothèses (H1), (H2) et (H4 (i)), on a

$$\sup_{x\in S}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x,y)-\mathbb{E}\widetilde{\psi}_n(x,y)\right|=0\left(\frac{\log n}{n\phi_n(h_K)}\right)^{1/2}, \quad \text{p.s} \quad n\to\infty \tag{5.6}$$

Lemme 5.2.7. sous les hypothèses (A3) (ii), (A4) (ii) and (iii), on a

$$\exists v>0, \quad \sum_{n\geq 1}\mathbb{P}\left(\inf_{x\in S}\widetilde{g}_n(x)\leq v\right)<\infty \tag{5.7}$$

Lemme 5.2.8. sous les hypothèses (H1), (H4), on a

$$\sup_{x\in S}\sup_{y\in[a,b]}\left|\mathbb{E}\widetilde{\psi}_n(x,y)-\alpha a_1\psi(x,y)\right|=0(h_K^\beta)+0(h_H^\beta), \qquad n\longrightarrow\infty \tag{5.8}$$

Démonstration du lemme (5.1) :
On a

$$\begin{aligned}
\sup_{x\in S}\left|\frac{1}{\phi(h_K)}\mathbb{E}\left(K\left(\frac{||x-X_i||}{h_K}\right)\right)\right) &= \frac{1}{\phi(h_K)}\int_0^{h_K}K\left(\frac{u}{h_K}\right)dP^{||x-X_1||}(u)\\
&= \frac{K(1)F_x(h_K)}{\phi(h_K)}-\frac{1}{\phi(h_K)}\int_0^1 K'(u)F_x(h_K u)du\\
&= \left(K(1)-\int_0^1 K'(u)\xi_{h_K}(u)du\right)[g(x)-0(1)]
\end{aligned}$$

On trouve la dernière égalité à cause de l'hypothèse $(A1-ii)$ qui converge vers $a_1 g(x)$ quad $n\to\infty$

Démonstration du lemme (5.2.2) : Comme S est un compact de R, alors, on peut extraire un recouvrement fini $S\subset\bigcup_{k=1}^{l_n}B(x_k,r_n)$ où B est la boule centré en x_k de rayon r_n. En autre, on suppose qu'il existe une constante C positive tel que $l_n r_n=C$. Pour tout $x\in S$, $k(x)=\arg\min_{1\leq k\leq l_n}d(x_k,x)$. On a

$$\begin{aligned}
|\widetilde{g}_n(x)-E\widetilde{g}_n(x)| &= \left|\widetilde{g}_n(x)-\widetilde{g}_n(x_{k(x)})\right|\\
&+ \left|\widetilde{g}_n(x_{k(x)})-E\widetilde{g}_n(x_{k(x)})\right|+\left|E\widetilde{g}_n(x_{k(x)})-E\widetilde{g}_n(x)\right|\\
&= G_1(x)+G_2(x)+G_3(x)
\end{aligned}$$

D'aprés l'hypothèse $(A3, ii)$ et $r_n = \left(\dfrac{h_K^2\phi(h_K)}{n}\right)^{1/2}$, alors

$$\sup_{x\in s}|G_1| \leq \frac{\alpha||K'||_\infty||x-x_{k(x)}||}{h_K} = \frac{\alpha||K'||_\infty r_n}{h_K\phi(h_K)} = \frac{\alpha||K'||_\infty}{(nh_K^2\phi(h_K))^{1/2}}$$

Pour n est assez large, on a $T_1 = T_3 = 0$
Pour le terme G_2, on pose

$$V_i(x_k) = \frac{\alpha}{\phi(h_K)G(Y_i)}\left(K\left(\frac{||x_k - X_i||}{h_K}\right) - E\left(K\left(\frac{||x_k - X_i||}{h_K}\right)\right)\right)$$

Tel que,

$$G_1 = \frac{1}{n}\sum_{i=1}^{n} V_i(x_k)$$

Donc, on peut trouver facilement

$$|V_i(x_k)| \leq \frac{\alpha||K||_\infty}{\phi(h_K)G(a_F)} = d$$

Et

$$\mathbb{E}\left|V_i(x_k)\right|^2 \leq \frac{C\alpha||K||_\infty}{\phi(h_K)G(a_F)} = \delta^2$$

L'utilisation de l'inégalité de Bernstein([3]), permet d'écrire :

$$\begin{aligned}
\mathbb{P}\left\{\max_{1\leq k\leq l_n}\sup_{x\in B(x_k,r_n)}\left|G_2(x_{k(x))}\right| > \epsilon\right\} &\leq \sum_{k=1}^{l_n}\mathbb{P}\left\{\sup_{x\in B(x,r)}\left|G_2(x_{k(x))}\right| > \epsilon\right\}\\
&= l_n \max_{k=1,\cdots,l_n}\mathbb{P}\left\{\left|G_2(x_{k(x))}\right| > \epsilon\right\}\\
&\leq 2\frac{C}{r_n}\exp\left(-\frac{n\epsilon^2\phi(h_K)G(a_F)}{\alpha||K||_\infty}\right)
\end{aligned}$$

En choisissant, r et $\epsilon = \epsilon_0\left(\dfrac{\log n}{n\phi(h_K)}\right)^{1/2}$, pour tout $\epsilon_0 > 0$, on obtient

$$J_2 \leq 2C\left(\frac{\log n}{nh_K\phi(h_K)}\right)^{1/2} \times \frac{1}{(n^2h_K\log n)^{1/2}} \times n^{2-\frac{C'\epsilon_0^2G(a_F)}{\alpha||K||_\infty}}$$

Pour

$$\frac{C'\epsilon_0^2G(a_F)}{\alpha||K||_\infty} = 3+\gamma, \qquad \gamma > 0$$

On a,

$$J_2 \leq 2C\left(\frac{\log n}{nh_K\phi(h_K)}\right)^{1/2} \times \frac{1}{(n^2h_K\log n)^{1/2}} \times n^{-1-\gamma}$$

Les hypothèses $[A4(ii, iii)]$, la série de Riemann et par le lemme de Borel-Cantelli donne le résultat.

Démonstration du lemme (5.2.3)
On a,

$$\begin{aligned}
\sup_{x\in S}|g_n(x)-\widetilde{g}_n(x)| &= \sup_{x\in S}\left|\frac{\alpha_n}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G_n(Y_i)}K\left(\frac{||x-X_i||}{h_K}\right)-\right.\\
&- \left.\frac{\alpha}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{||x-X_i||}{h_K}\right)\right|\\
&= \left(\frac{|\alpha_n-\alpha|}{G_n(a_F)}+\frac{\alpha}{G_n(a_F)G(a_F)}\sup_{a\le y\le b}|G_n(y)-G(y)|\right)\sup_{x\in S}|v_n(x)|
\end{aligned}$$

Avec

$$v_n(x)=\frac{1}{n\phi(h_K)}K\left(\frac{||x-X_i||}{h_K}\right)$$

On la convergence de $G_n(a_F)\longrightarrow G(a_F)$ et $\sup_{a\le y\le b}|G_n(y)-G(y)|=0(n^{-1/2})$ (voir la remarque 6 dans [19]). Les hypothèses du lemme (10) de ([8]) permet d'ecrire $\sup_{x\in S}|v_n(x)|=0(1)$

on obtient

$$\begin{aligned}
\widehat{\alpha}_n-\alpha_0 &= \frac{G_n(1-F_n(x-))}{R_n(x)}-\frac{G_0(1-F_0(x-))}{R(x)}\\
&= \frac{FRG}{RR}\left\{-\frac{1}{n}\sum W_i(x)-\frac{1}{n}\sum_{i=1}^{n}Z_i(x)\right\}\\
&- \frac{1}{nR(x)}\sum_{i=1}^{n}(I\left(V_i\le x\le U_i\right)-R(x))+\circ\left(\frac{\log^3 n}{n}\right)\\
&= -\alpha_0\frac{1}{n}\sum_{i=1}^{n}\xi_i(x)+\circ\left(\frac{\log^3 n}{n}\right)
\end{aligned}$$

Où

$$\xi_i(x)=W_i(x)+Z_i(x)+\frac{1}{R(x)}(I(V_i\le x\le U_i))-R(x)),\qquad i=1,\cdots,n$$

est une suite des variables indépendants et identiquement distribuées tends vers zero.

Démonstration du lemme (5.2.2)
Nous avons

$$\mathbb{E}[\widetilde{g}_n(x)]=\frac{\alpha}{\phi(h_K)}\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_1||}{h_K}\right)\right]$$

$$\begin{aligned}
&= \frac{\alpha}{\phi(h_K)}\mathbb{E}\left[\mathbb{E}\left[\frac{1_{\{Y_1\ge T_1\}}}{G(Y_1)}K\left(\frac{||x-X_1||}{h_K}\right)|X_1,Y_1\right]\right]\\
&= \frac{\alpha}{\phi(h_K}\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_1||}{h_K}\right)\mathbb{E}\left[1_{\{Y_1\ge T_1\}}\middle|X_1,Y_1\right]\right]\\
&= \frac{\alpha}{\phi(h_K)}\mathbb{E}\left[K\left(\frac{||x-X_1||}{h_K}\right)\right]
\end{aligned}$$

À l'aide du lemme (5.2.1), on obtient le résultat.
Démonstration du lemme 5.2.5

$$\begin{aligned}\left|\psi_n(x,y)-\widetilde{\psi}_n(x,y)\right| &= \left|\frac{\alpha_n}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G_n(Y_1)}K\left(\frac{||x-X_i||}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)\right.\\ &- \left.\frac{\alpha}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G(Y_1)}K\left(\frac{||x-X_i||}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)\right|\\ &\leq \left\{\frac{|\alpha_n-\alpha|}{G_n(a_F)}+\frac{\alpha}{G_n(a_F)G(a_F)}\sup_{a\leq y\leq b}|G_n(y)-G(y)|\right\}|\Psi_n(x,y)|\end{aligned}$$

Avec

$$\Psi_n(x,y)=\frac{1}{n\phi(h_K)}\sum_{i=1}^{n}K\left(\frac{||x-X_i||}{h_K}\right)H\left(\frac{y-Y_i}{h_H}\right)$$

De la même manière que le lemme 5.2.3 avec le lemme (3.3) dans ([9]) pour $j=0$, on obtient

$$\sup_{x\in s}\sup_{a\leq y\leq b}\left|\psi_n(x,y)-\widetilde{\psi}_n(x,y)\right|=O\left(n^{-1/2}\right)$$

Démonstration du lemme 5.2.6
Pour s, on utilise le même recouvrement du lemme (5.2.2). Puisque $[a,b]$ est un sous compact réel de $\mathbb{R}$, on peut extraire un nombre fini j_n de longueur $[a,b]\subset\bigcup_{t=1}^{j_n}I_t$, où $I_t=(y_t-u_n,y_t+u_n)$, $j_n=\frac{Q_1}{u_n}$ et $y_t=\arg\min_{t\in(t_1,\ldots,t_{a_n})}|y-t|$, on a

$$\begin{aligned}\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x,y)-\mathbb{E}\left[\widetilde{\psi}_n(x,y)\right]\right| &\leq \underbrace{\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x,y)-\widetilde{\psi}_n(x_k,y)\right|}_{T_1}\\ &+ \underbrace{\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x_k,y)-\widetilde{\psi}_n(x_k,y_t)\right|}_{T_2}\\ &+ \underbrace{\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x_k,y_t)-\mathbb{E}\left[\widetilde{\psi}_n(x_k,y_t)\right]\right|}_{T_3}\\ &+ \underbrace{\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\mathbb{E}\left[\widetilde{\psi}_n(x_k,y_t)\right]-\mathbb{E}\left[\widetilde{\psi}_n(x_k,t)\right]\right|}_{T_4}\\ &+ \underbrace{\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\mathbb{E}\left[\widetilde{\psi}_n(x_k,t)\right]-\mathbb{E}\left[\widetilde{\psi}_n(x,t)\right]\right|}_{T_5}\end{aligned}$$

Comme la H est limitée par 1, puis nous revenons directement au Lemm (5.2.2) avec le même choix de ε, on a

$$\mathbb{P}\left\{T_1>\varepsilon_0\left(\frac{\log n}{n\phi(h_k)}\right)^{1/2}\right\}=\mathbb{P}\left\{T_5>\varepsilon_0\left(\frac{\log n}{n\phi(h_k)}\right)^{1/2}\right\}=0$$

D'aprés la condition liptshizienne

$$\sup_{x\in\Sigma}\sup_{y\in[a,b]}\left|\widetilde{\psi}_n(x_k,y)-\widetilde{\psi}_n(x_k,y_t)\right|$$

$$\begin{aligned}
&\leq \sup_{x\in\Sigma}\sup_{y\in[a,b]}\frac{\alpha}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G(Y_i)}\left|H\left(\frac{y-Y_i}{h_H}\right)-H\left(\frac{y_t-Y_i}{h_H}\right)\right|K\left(\frac{||x-X_i||}{h_K}\right)\\
&\leq \sup_{x\in\Sigma}\sup_{y\in[a,b]}\frac{C|y-y_t|}{h_H}\left(\frac{\alpha}{n\phi(h_K)}\sum_{i=1}^{n}\frac{1}{G(Y_i)}K\left(\frac{||x-X_i||}{h_K}\right)\right)\\
&\leq C\frac{\alpha u_n}{h_H G(a_F)}\sup_{x\in\Sigma}\upsilon_n(x_k)
\end{aligned}$$

D'aprés lemme on obtient

$$O\left(\frac{Cu_n}{h_H G(a_F)}\alpha a_1 g(x)\right)=O\left(\frac{u_n}{h_H}\right)$$

En choisissant $u_n = n^{-\nu-1|2}$ avec $\nu > 0$ et avec $(A4, iii)$ nous obtenons $\frac{u_n}{h_H} = O\left(\left(\frac{\log n}{n\phi(h_K}\right)^{1/2}\right)$.
pour n est assez grand et pourtout les $\varepsilon > 0$, nous obtenons

$$\mathbb{P}\left\{T_2>\varepsilon_0\left(\frac{\log n}{n\phi(h_k)}\right)^{1/2}\right\}=\mathbb{P}\left\{T_4>\varepsilon_0\left(\frac{\log n}{n\phi(h_k)}\right)^{1/2}\right\}=0$$

Pour le terme T_3, il est clair que

$$\begin{aligned}
\mathbb{P}(T_3>\varepsilon) &< \sum_{k=1}^{l_n}\sum_{k=1}^{t_{s_n}}\mathbb{P}\left(\sup_{x\in B(x_k,r_n)}\sup_{y_t\in\{t_1,\dots,t_{\alpha_n}\}}\left|\widetilde{\psi}_n(x_k,y_t)-\mathbb{E}\left[\widetilde{\psi}_n(x_k,y_t)\right]\right|\varepsilon\right)\\
&\leq s_n l_n \max_{k=1,\cdots,l_n}\max_{y_t\in\{t_1,\dots,t_{\alpha_n}\}}\mathbb{P}\left(\left|\widetilde{\psi}_n(x_k,y_t)-\mathbb{E}\left[\widetilde{\psi}_n(x_k,y_t)\right]\right|>\varepsilon\right)
\end{aligned}$$

Soit

$$\lambda(x,y)=K\left(\frac{||x-X_i||}{h_K}\right)H\left(\frac{y_t-Y_i}{h_H}\right)-\mathbb{E}\left[K\left(\frac{||x-X_i||}{h_K}\right)H\left(\frac{y_t-Y_i}{h_H}\right)\right]$$

Les mêmes arguments du lemme (5.2.2) utilisées et le fait que $H \leq 1$, on déduit que

$$\mathbb{E}|\lambda|\leq 2||K||_\infty|H|\leq C,\quad \text{et}\quad \mathbb{E}|\lambda|^2\leq\frac{C'\alpha||K||_\infty}{G(a_F)\phi(h_K)}$$

Maintenant,nous appliquons l'inégalité exponentielle de Bernstein([3]) pour obtenir

$$\mathbb{P}(T_3>\epsilon)\leq 2s_n l_n\exp\left\{\frac{-C'n\varepsilon^2\phi(h_K)G(a_F)}{2\alpha||K||_\infty}\right\}$$

$$=\frac{2C}{r_n u_n}\exp\left\{\frac{-C'n\varepsilon^2\phi(h_K)G(a_F)}{\alpha||K||_\infty}\right\}$$

le méme choix de r_n, u_n et ε donnes

$$\frac{2C}{(n^2h_K^2\phi(h_K))^{1/2}} \times n^{\nu+2-\frac{C^{'}n\varepsilon^2\phi(h_K)G(a_F)}{\alpha||K||_\infty}}$$

Ce qui nous permet de conclure, en choisissant $\frac{C^{'}n\varepsilon^2\phi(h_K)G(a_F)}{\alpha||K||_\infty} = 3+\nu+\gamma$ avec $\gamma > 0$

$$\mathbb{P}(T_3 > \epsilon) \leq 2C\left(\frac{\log n}{nh_K\phi(h_K)}\right)^{1/2} \times \left(\frac{1}{nh_K \log n}\right)^{1/2} \times n^{-1-\gamma}$$

Par les deux hypothèses $(A4,(i),(iii))$, on déduit le résultat du lemme.

Démonstration du lemme 5.2.7

D'aprés l'égalité suivante

$$\inf_{x\in\Sigma} \widetilde{g}_n(x) \geq \inf_{x\in\Sigma} \mathbb{E}[\widetilde{g}_n(x)] - \sup_{x\in\Sigma} |\widetilde{g}_n(x) - \mathbb{E}[\widetilde{g}_n(x)]|$$

et lemme , on obtient le résultat.

Démonstration du lemme 5.2.8 On a,

$$\begin{aligned}
\mathbb{E}\big[\widetilde{\psi}_n(x,y)\big] &= \frac{\alpha}{\phi(h_K)}\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_1||}{h_K}\right)H\left(\frac{y-Y_1}{h_H}\right)\right] \\
&= \frac{\alpha}{\phi(h_K)}\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_1||}{h_K}\right)\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)\Big|X_1\right]\right]
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\mathbb{E}\left[H\left(\frac{y-Y_i}{h_H}\right)\Big|X_1\right] &= \int_R H\left(\frac{y-u}{h_H}\right)f(u|X_1)du \\
&= \int_R H^{'}(z)F(y-zh_H|X_1)dz \\
&= \int_R H^{'}(z)[F(y-zh_H|X_1)-F(y|x)]dz + F(y|x)
\end{aligned}$$

Puisque la fonction $H^{'}(.)$ est la densité de probabilité

$$\begin{aligned}
\mathbb{E}\big[\widetilde{\psi}_n(x,y)\big] &= \frac{\alpha}{\phi(h_K)}\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_i||}{h_K}\right)\int_R H^{'}(z)[F(y-zh_H|X_1)-F(y|x)]dz\right] \\
&+ \frac{\alpha}{\phi(h_K)}F(y|x)\mathbb{E}\left[\frac{1}{G(Y_1)}K\left(\frac{||x-X_i||}{h_K}\right)\right] \\
&= L_1 + L_2
\end{aligned}$$

A l'aide du lemme , L_2 tend vers à $\alpha a_1\psi(x,y)$ quand $n\to\infty$.

D'aprés l'hypothèse $(H3)$, on a

$$\begin{aligned}
\int_R H^{'}(z)|F(y-zh_H|X_1)-F(y|x)|dz &\leq C_x\int_R H^{'}(z)(h_K^\beta + |z|^\gamma h_H^\gamma)dz \\
&\leq C_xh_K^\beta + C_xh_H^\gamma\int_R |z|^\gamma H^{'}(z)dz
\end{aligned}$$

Il est clair d'aprés $(H3,ii)$ et Lemme (5.2.2), L_1 tend vers zero quand $n\longrightarrow\infty$.

Conclusion

On peut mensionner les remarques suivantes :

- ce problème trouvera une solution original, en effet le fléau de la dimension est la dégradation de la vitesse de convergence par rapport à l'augmentation de la dimension.
- Dans le cas de dimension infini, nous avons une dégradation de la vitesse de convergence par rapport à la faible concentration de la mesure de probanilité de la variable explicative fonctionnelle.
- Ainsi, une solution originale de ce problème est le choix de la semi-métrique qui donne une forte concentration de la mesure de probabilité de la variable fonctionnelle.
- Les vitesses des convergences obtenues sont insensibles aux corrélations des observations. Cependant, certaines hypothţeses suppléementaires ont été introduites afin de prendre en compte la dépendance des observations.
- Cet estimateur peut utiliser pour construire des intérvalles de confiance et de faire de la prédiction en séries temprelles.
- Les hypothéses de la dépendance sont ajoutées pour completer les données indépendantes qui n'a pas toujours raisonnable en pratique.

Bibliographie

[1] Anderson PK, Borgan, Gill RD et Keiding N (1993) Statistical model based on counting procecesse. New-York :Springer-Velrag.

[2] Berlinet,A,.Gannoun,A.,Matzner-Lober,E.(2001) Asymptotic normality of convergent estimaes of conditional quantiles.Statistics,vol. 35, pp. $139 - 169$.

[3] Bernstein,S.Probability Theory,4 thed.(in russian),ed.M.L.Gostechizdat (1946).

[4] Bosq, D(1998) Nonparametrice statistics for stochastic processus :Estimation and Prediction (Second Edition), Lecture Note in Statistics, 110, Springer Velrage ,New York.

[5] Cardot,H.Crambes,C.Sarda,P.(2004) Spline estimation of conditional quantities for functional covariates.C.R.,Math,Acad.Sci.Paris 339,No,$2, 227 - 236$.

[6] Dabo-Niag,S.,Laksaci,A.,(2006). Estimation non paramétrique des quantiles conditionnels pour variable explicative fonctionnelle.

[7] E.Giné,A. Guillou, On consistency of kernel density estimators forrandomly censored data : Rates holding uniformly over adaptive intervals, Ann.I. H.Poincaré 37(2001) 503-522.

[8] F. Ferraty, A. Laksaci, A. Tadj, P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, J. of Statist. Plann. and Inference 140 (2010), 335Ű352

[9] F. Ferraty, A. Laksaci, P. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models, Statist. Inf. Stoch. Processes 9 (2006), 47Ű76

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

[11] Ferraty,F.,Rabhi,A.Vieu,Ph,(2005). Conditional quantiles for functionally dependent data with application to the climatic El Nio Phenomenon.Sankhya,67, No, $2378 - 399$.

[12] Gannoun,A.(1989) Estimation de la médiane conditionnelle, Thèse de doctorat de l'université de Paris.

[13] Gannoun,A.,Saracco,J. Yu, K. (2003). Nonparametric prediction by conditional median and quantiles .J.Stat.Plann.Inference,117, No.$2, 207 - 223$.

[14] Gürler, U. Stute,W. Wang,J.L.(1993) Weak and strong quantile representations for randomly truncated data with applications,Statist. Probab. Lett.$17, 139 - 148$.

[15] He,S.Yang,G.(1998) Estimation of the truncation probability in the random truncation model,Ann.Statist.$26, 1011 - 1027$.

[16] Lemdani,M.Ould Said,E.(2007) ,Asymptotic behavior of the hazard rate kernel estimator under truncated and censored data.Comm in Statisti.Theory and Methods $36, 155 - 174$.

[17] Lemdani,M.Ould Said,E.Poulin,N. (2005) Strong representation of the quantile function for left truncated and dependent data,Math.Statis.14.$332-345$.

[18] Liange H.Y.Li D,Oi Y.(2009) Strong convergence in nonparametric regression with truncated dependant data ,J.Multivariat.Anal.$100, 162174$.

[19] M. Woodroofe, Estimating a distribution function with truncated data, Ann. Statist. 13

[20] Ould-Said,E.Tatachak,A (2009) Strong uniform consistency rate for the kernel mode under strong mixing hypothesis and left truncation ,Comm.in Statist.Theory and Methods $37, 2735-2759$.

[21] Roussas,G. (1969) Nonparametric estimation of the transition distribution function of a Markov process.Ann.Math.Statist.,$40, 1386-1400$.

[22] Stone,C.J.Consistent nonparametric regression.Discussion.Ann.Stat.$5, (1977), 595-645$.

[23] Samanta,M., (1989),Nonparametric estimation of conditional quantiles.Stat.Probab.Lett. 7, No. 5,407-412.

[24] Woodroofe,M.(1985) Estimation a distribution function with truncated data,Ann.Statist. $13, 477-480$.

[25] Y.S. Chow, H. Teicher, Probability Theory. Independence, Interchangeability, Martingales, Springer, New York, 1997.
(1985), 163Ű177.