République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieure et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2017/2018

Propriétés asymptotiques de l'estimateur à noyau de la densité d'un processus stationnaire à temps continu

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité : Analyse Stochastiques, Statistique des Processus et Applications

par

Bekhti Naima [1]

Sous la direction de

Dr. R. Rouane

Soutenue le 24/06/2018 devant le jury composé de

Mme : W. Benzatout	Université Dr Tahar Moulay - Saïda	Présidente
Mme : R. Rouane	Université Dr Tahar Moulay - Saïda	Encadreur
Mlle : S. Rahmani	Université Dr Tahar Moulay - Saïda	Examinateur
Mme : F. Benziadi	Université Dr Tahar Moulay - Saïda	Examinateur

1. e-mail : naimaninadz11@gmail.com

REMERCIEMENTS

En tout premier lieu, je remercie le bon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés .
Je désire tout d'abord, témoigner de ma reconnaissance envers ma directirce ***Dr. Rouane Rachida****, l'instigatrice de cette étude, qui m'a remis le pied à l'étrier de la recherche, pour sa présence, la constance de son soutien, son aide incommensurable, Sans elle, ce travail aurait été plus difficile, voire impossible.*
Je voudrais également remercier tous les membres de jury ; ***Mme. Benzatout Wahiba*** *de m'avoir honorée en acceptant de présider le jury,* ***Mme. Benziadi Fatima*** *et* ***Mlle. Rahmani Saadia*** *d'avoir accepté d'examiner ce modeste travail, merci pour toutes leurs remarques et critiques.*
J'exprime mes remerciements à tous mes enseignants du département de Mathématiques qui m'ont initié aux valeurs authentiques, en signe d'un profond respect.

dédicace

Je dédie ce modeste travail :

A ma chère mère ***Badra*** *que Dieu la protège pour m'avoir encouragée et permis d'entreprendre mes études. Sans elle, je ne serais pas là .*
A mon cher père ***Benyahia*** *qui avait toujours souhaité ma réussite, et son aide précieuse qu'il m'a apportée tout au long de ce travail.*
A mes chères sœurs ***Fatima****;* ***Imene****;* ***Ikram****;* ***Ikhlasse****;* ***Hadjer****; à mon chèr frère* ***Mohamed Elhabib*** *pour leur encouragement.*
A mon cher oncle ***Bekhti Abdelkarim****.*
A tous les membres de ma famille.
A ***M. Aimer*** *pour son soutien, conseils, et son aide.*
Une dédicace spéciale à mon adorable ***Imene Bouazza*** *pour son amitié, son soutien inconditionnel et son encouragement.*
Une spéciale dédicace à ma chère tante ***Ikhlef Jamila*** *, à mes chères sœurs* ***Nourhan ; Massilia****.*
A tous mes Ami(e)s que j'aime tant pour leur sincère amitié
A qui je dois ma reconnaissance et mon attachement, à tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

Table des matières

Introduction générale

L'estimation est un élément fondamental de la statistique mathématique qui développe des techniques pour décrire certaines caractéristiques d'ensembles d'observations. On y distingue

L'approche paramétrique, qui considère que les modèles sont connus, avec des paramètres inconnus. La loi de la variable étudiée est supposée appartenir à une famille de lois pouvant être caractérisée par une forme fonctionnelle connue (fonction de répartition F, densité f,...) qui dépend d'un ou plusieurs paramètres inconnus à estimer.

L'approche non paramétrique, qui ne fait aucune hypothèse sur la loi, ni sur ses paramètres. Nos connaissances sur le modèle ne sont pas précises, ce qui est souvent le cas dans la pratique. Dans cette situation, il est naturel de vouloir estimer une des fonctions décrivant le modèle, soit généralement la fonction de répartition ou la densité (pour le cas continu) : c'est l'objectif de l'estimation fonctionnelle. Si la fonction de répartition empirique F_n résout le problème statistique fondamental de la distribution de probabilité (associée à un échantillon $(T_1, T_2, ..., T_n)$ de variables aléatoires réelles indépendantes et de même loi) en fonction des valeurs numériques observées, elle est par contre limitée pour décrire visuellement les caractères de l'échantillon.

Les contributions à l'étude des paramètres fonctionnels ont beaucoup enrichi la littérature ces dernières décennies. Nous citons à titre d'exemples les travaux de Rosenblatt (1956)[57], Parzen (1962)[54], Nadaraya (1965)[49], Watson (1964)[63], Banon (1978)[6] ainsi que celui de Castellana- Leadbetter (1986)[20], présentant une condition portant sur les densités jointes et marginales d'un processus strictement stationnaire et utilisée pour atteindre des vitesses suroptimal dans le cas d'un processus à temps continu (voir Bosq (1997)[12]). L'estimation non paramétrique des fonctions de la densité, constitue un sujet qui a suscité une intense activité de recherche et permis l'introduction de nombreux outils et méthodes considérant des contextes multiples et variés.

La prise en compte de structure de dépendance pour l'estimation de paramètres rend le cadre de travail plus proche de ce qui est observé dans la réalité. L'étude des propriétés relatives aux processus à temps continu ne peut être envisagée en dehors d'une structure de dépendance des données. Les conditions de mélange sont les plus utilisées, à travers la littérature, dans l'analyse des propriétés de ce type de processus. Pour de nombreux processus bien connus, les propriétés de mélange ont déjà été établies. Cependant, la structure de dépendance de multiple processus utiles reste un problème ouvert. À titre d'exemple, Ibragimov et Linnik (1971)[38], Chernick (1981)[22] et Andrews (1984)[2] ont montré que dans certains cas le processus autorégressif linéaire du premier ordre en temps discret n'est pas fortement mélangeant.

En particulier, le processus stationnaire $(X_t)_{t\in\mathbb{Z}}$ vérifiant le modèle $AR(1)$ défini par $X_t = \theta X_{t-1}+\epsilon_t$, où les $(\epsilon_t)_{t\in\mathbb{Z}}$ est une suite d'innovation de Bernoulli i.i.d, n'est pas fortement mélangeant. Toutefois, l'ergodicité est conservée en prenant des fonctions mesurables d'un processus ergodique. Si $(\epsilon_t)_{t\in\mathbb{Z}}$ est un processus strictement stationnaire et ergodique, $Y_t = \mathcal{V}((..., \epsilon_{t-1}, \epsilon_t), (\epsilon_{t+1}, \epsilon_{t+2}, ...))$ pour une certaine fonction borélienne $\mathcal{V}(.)$, alors $(Y_t)_{t\in\mathbb{Z}}$ est également ergodique ; voir Proposition 2.10 à la page 54 de Bradley (2007)[13]. Comme le processus autorégressif, ci-dessus, peut être représenté comme une fonction linéaire des ϵ_t, il s'ensuit alors qu'il est aussi ergodique. Nous construisons maintenant un autre exemple de processus ergodique non mélangeant et cette fois-ci en temps continu. Il est bien connu que le mouvement brownien fractionnaire $(W_t^H; t \geq 0)$ de paramètre de Hurst $H \in (0,1)$ admet des incréments strictement stationnaires. Par ailleurs, le bruit gaussien fractionnaire, défini pour tout $s > 0$ par $(G_t^H)_{t\geq 0} = (W_{t+s}^H - W_t^H)_{t\geq 0}$, est un processus centré strictement stationnaire à longue mémoire lorsque $H \in (\frac{1}{2}, 1)$ (voir, par exemple, Beran (1994)[11] p. 55, Lu (2009)[46] p. 17). De ce fait, il contrevient alors à la condition de mélange fort. Pour un processus réel gaussien strictement stationnaire et centré $(G_t^H)_{t\geq 0}$, considérons la fonction de corrélation $R(t) = \mathbb{E}[G_0 G_t]$. En se reportant au travail de Maslowski et Pospì$\tilde{s}$il (2008)[48], Lemme 4.2, il en ressort que le processus $(G_t)_{t\geq 0}$ est ergodique lorsque $\lim_{t\to\infty} R(t) = 0$. Un calcul facile permet alors de voir que le processus $(G_t^H)_{t\geq 0}$ vérifie bien cette condition.

Par ailleurs, il est bien connu que les conditions de mélange sont difficiles à vérifier en pratique. Il est donc préférable de définir un cadre de travail général en considérant des processus satisfaisant des conditions d'ergodicité qui incluent à la fois des processus mélangeants et non-mélangeants. Il est bien connu que l'ergodicité est satisfaite par les processus mélangeants et elle est, en outre, plus facile à vérifier en pratique (Krengel (1985)[40], p. 24, Ash et Gardner (1975)[3], p. 120). La difficulté principale de l'ergodicité est qu'elle ne nous permet pas, sans conditions supplémentaires, d'obtenir des vitesses de convergence (Krengel (1985)[40]). Notre objectif est de construire un carde de travail relatif à l'estimation non paramétrique fonctionnelle de la densité en temps continu, vérifiant une condition d'ergodicité et d'établir des résultats de consistances ponctuelles et uniformes, avec des vitesses de convergence et la normalité asymptotique de l'estimateur.
Le présent manuscrit est composé de trois chapitres :

Dans le premier chapitre, on rappelle les principaux concepts ainsi que les propriétés les plus importantes qui nous permettent de mieux décrire les problèmes traités en citant au fur et à mesure les travaux antérieurs.

Dans le deuxième chapitre, nous nous intéressons à la question de l'estimation de la fonction de densité d'un processus stationnaire à temps continu. Sous des conditions de dépendance assez générales, nous établissons les convergences presque sûres, ponctuelles et uniformes avec des vitesses de convergence de l'estimateur à noyau de la densité du processus.

Dans le troisième chapitre et dans le même contexte ergodique, nous établissons la normalité asymptotique de l'estimateur considéré dans le chapitre précédent, puisqu'il ne reste plus qu'à montrer que les variances asymptotiques sont non nulles ce qui confirmera que les lois limitées sont non dégénérées.

Chapitre 1

Présentation

1.1 Historique de l'estimation non-paramétrique de la densité en temps continu

Soit $(X_i)_{i\geq 1}$ une suite stationnaire des variables aléatoires. On note $\mathbb{P}$ la loi de probabilité sur $\mathbb{R}$ régissant ce processus et f sa densité par rapport à la mesure de Lebesgue.

La méthode du noyau a d'abord été décrite en 1951 dans un rapport non publié par Fix et Hodges (voir Silverman and Jones (1989)[34]). Un peu plus tard, la première forme de l'estimateur à noyau était introduite dans les travaux de Rosenblatt (1956)[57] et Parzen (1962) [54] qui ont défini une application réelle K vérifiant la condition $\int_{\mathbb{R}} K(u)du = 1$ et connue depuis sous le nom du noyau de Parzen-Rosenblatt. L'estimateur de Parzen-Rosenblatt est défini, pour tout $x \in \mathbb{R}$, par

$$\hat{f}_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right),$$

où h_n, la fenêtre de lissage, est un paramètre tendant avec une certaine vitesse vers 0 lorsque n tend vers l'infini.

La première version de cet estimateur fut donnée par Rosenblatt(1956), en choisissant le noyau uniforme $K = \frac{\mathbb{1}_{[-1,1]}}{2}$, il donna l'erreur quadratique moyenne de l'estimateur de la densité pour des observations univariées indépendantes et identiquement distribuées. La généralisation de ce résultat fut obtenue par Parzen(1962) pour une vaste classe de noyaux qui établit aussi la normalité asymptotique. Calcoullos (1966)[19] traita le cas multivarié. Résultats de la section 6.3 du Bosq et Blanke (2008)[17] sont des améliorations des résultats parus dans Bosq (1998)[14]. Concernant la convergence presque sûre, ils ont délibérément choisi une condition de mélange assez forte (nommément une décroissance exponentielle du coefficient) pour présenter des résultats optimales (c'est-à-dire constantes asymptotiques similaires à ceux du cas i.i.d.). Pour les conditions de dépendance affaiblis, le lecteur intéressé pourra se référer aux ouvrages suivants (et les références qui y sont) : Liebscher (2001)[42] pour les processus arithmétiques fortement mélangeants, Doukhan et Louhichi (2001)[34] pour un nouveau concept de dépendance y compris les processus qui ne sont pas mélangeants en général. Pour des choix pratiques de h_n dans le cas dépendant.

Banon (1978)[6] fut le premier à s'intéresser à l'estimation de la densité en temps continu, dans le cas de processus de diffusion sous la condition G2 de Rosenblatt (voir Rosenblatt (1970)[59]) en utilisant des noyaux récursifs. L'estimateur de la densité par la méthode du noyau est donné, à partir de l'observation d'une partie $(X_t)_{0\leq t\leq T}$ d'un processus stationnaire $(X_t)_{t\geq 0}$. S'inspirant du cas discret, la forme de cet estimateur est donnée, pour tout $x \in \mathbb{R}$, par

$$\hat{f}_T(x) = \frac{1}{Th_T}\int_0^T K\left(\frac{x - X_t}{h_T}\right)dt.$$

Les premiers résultats de l'estimation de la fonction de densité pour les processus de diffusion sont apparus dans Banon (1978)[6] suivis de Banon et Nguyen ((1978)[7], (1981)[8]) ; Nguyen (1979)[50] et Nguyen et Pham (1980)[56].

Par la suite, Delecroix (1979)[27], étudia la convergence presque sûre et la convergence en moyenne quadratique des estimateurs de la densité et de la densité conditionnelle pour des processus strictement stationnaires et fortement mélangeants. Quelques années plus tard, Castellana et Leadbetter (1986)[20] démontrèrent qu'il était possible, sous certaines conditions portant sur la densité jointe du processus, d'atteindre une vitesse de convergence en moyenne quadratique de l'estimateur de la densité en T^{-1}, qu'on appelle vitesse "suroptimale" ou "paramétrique" puisqu'elle est identique à la vitesse obtenue usuellement en statistique paramétrique. Ce résultat historique a par la suite inspiré de nombreux travaux en temps continu. On pourra se référer par exemple aux ouvrages de Bosq (1998) [16] et Bosq et Blanke (2007)[19] pour une étude complète des estimateurs à noyau et par projection, et de Kutoyants (2004)[14] pour l'inférence des processus de diffusion ergodiques, ainsi qu'à leurs bibliographies respectives. Le cas particulier des ondelettes a également été examiné par Leblanc (1995)[45]. Enfin, Kutoyants (1995)[41] a prouvé que $(1/T)$ est la vitesse minimax lorsque les processus de diffusion sont observés. Bosq (1997)[12] et Bosq et al (1999)[15] ont obtenu la normalité asymptotique pour les estimateurs à noyaux basés sur des échantillons des processus α-mélangeants. Didi et Louani (2013a)[33] ont obtenu des résultats de convergence presque sûre, ponctuelle et uniforme, avec des vitesses de convergence non paramétrique dans un cadre de convergence assez général où des méthodes de preuve basées sur des différences de martingale et des projections successives sur une famille de σ-algèbre emboitées, comparables à celle définies dans Wu et al (2010)[65] dans le cas discret, la normalité asymptotique ainsi que la consistance avec des vitesses de convergence pour des estimateurs non paramétriques de la densité et de la fonction de régression pour une large classe de processus linéaires et non-linéaires utilisés dans les series temporelles. Finalement, le choix de la fenêtre de lissage est un problème qui a suscité beaucoup d'intérêt, Yondjé et al (1994)[68] a proposé une méthode pour sélectionner une fenêtre qui soit asymptotiquement "bonne". Cette méthode s'inspire des idées de validation croisée qui ont été proposées dans d'autres problèmes d'estimation fonctionnelle, citons par exemple, Marron (1987)[47] pour la densité, Härdle et Marron (1985)[37] pour la régression, Sarda et Vieu, (1991)[61] pour la fonction de hasard et Sarda (1991)[60] pour le cas de la fonction de répartition. Nous pouvons aussi nous référer aux travaux de Youndé et al (1993a)[66], Youndjé et Wells (2008)[68], Absava (1999)[1], Chacón et Tenreiro (2012)[21] qui traitent de la même question.

1.2 Définitions de base

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ un espace probabilisé filtré.

Définition 1.2.1. *(Processus en temps continu). Soit* $X = (X_t, t \geq 0)$ *un processus défini sur* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ *. Le processus* X *est dit mesurable si l'application*

$$X : [0, \infty[\times\Omega \to (\mathrm{E}, \varepsilon)$$

$$(t, \omega) \to X_t(\omega)$$

est mesurable par rapport à $\mathcal{B}([0, \infty[) \otimes \mathcal{F}_t$*. Le processus* X *est dit adapté si* $\forall t \geq 0, X_t$ *est* $\mathcal{F}_t$*-mesurable.*

Définition 1.2.2. *(Processus stationnaire). Un processus* $(X_t)_{t\in\mathbb{U}}, \mathbb{U} = \{\mathbb{Z}, \mathbb{R}^+\}$ *est dit strictement stationnaire ou stationnaire au sens strict si les lois jointes de* $(X_{t_1}, \cdots\cdots\cdots, X_{t_k})$ *et de* $(X_{t_{1+h}}, \cdots\cdots\cdots, X_{t_{k+h}})$ *sont identiques pour tout entier positif* k *et pour tous* $t_1, \ldots\ldots\ldots, t_k, h \in \mathbb{Z}$.

Définition 1.2.3. *(Ensemble invariant). Soit* $(X_t)_{t\in\mathbb{U}}, \mathbb{U} = \{\mathbb{Z}, \mathbb{R}^+\}$*, défini sur un espace de probabilité* $(\Omega, \mathcal{F}, \mathbb{P})$ *, l'ensemble* $\mathcal{B} \in \mathcal{F}$ *est appelé invariant si il existe un ensemble* $\mathcal{A}$*, tel que* $\mathcal{B} = \{(X_n, X_{n+1}, ...) \in \mathcal{A}\}$ *est vrai pour tout* $n \geq 1$.

Définition 1.2.4. *(Processus ergodique). Un processus* $X = \{X_t\}_{t\in\mathbb{U}}$*,* $\mathbb{U} = \{\mathbb{Z}, \mathbb{R}^+\}$ *est dit ergodique, si tout ensemble invariant* $\mathcal{B}$ *on a* $\mathbb{P}(\mathcal{B}) = 1 \, ou \, 0$.

Définition 1.2.5. *(Filtration). Une filtration est une suite croissante de tribus* $(\mathcal{F}_t)_{t\geq 0}$*, c'est-à-dire*

$$\mathcal{F}_t \subset \mathcal{F}_{t+s} \qquad \textit{pour tout} \qquad t, s \geq 0.$$

Définition 1.2.6. *(Martingales). Soit* M *un processus adapté avec* $\forall t \geq 0$*,* $M_t \in L^1$*. On dit que* M *est une* $(\mathcal{F}_t)_{t\geq 0}$*-martingale si* $M_t = \mathbb{E}[M_{t+s}|\mathcal{F}_t]$ *pour tout* $t, s \geq 0$.

Définition 1.2.7. *(Différence de martingale en temps continu). La variable aléatoire* $M = (M_t, t \geq 0)$ *est une différence de martingale par rapport à la filtration* $\mathcal{F} = (\mathcal{F}_t, t \geq 0)$ *si*

1- M_t *est* $\mathcal{F}_t$*-mesurable,*

2- $\mathbb{E}[M_t|\mathcal{F}_{t-s}] = 0, \quad t \geq 0, \qquad s \geq 0.$

1.3 Théorie ergodique

La théorie ergodique est apparue dans la mécanique statistique, notamment dans la théorie de Maxwell et dans la théorie de Gibbs. Il est nécessaire de faire une sorte de transition logique entre le comportement moyen de l'ensemble des systèmes dynamiques et la moyenne temporelle des comportements d'un système dynamique unique. Elle découle d'une hypothèse ingénieuse dont on s'est servi pendant une longue période sans la justifier, et sous des formes variées. Le premier résultat rigoureux pour le mathématicien, est le célèbre théorème de récurrence de Poincaré datant de l'année 1890 (voir Peskir (2000)[57], théorème 1.3, p.6) mais l'essor de la théorie date de 1931 avec les théorèmes de Neumann et de Birkhoff, elle entre alors définitivement dans le cadre de l'analyse fonctionnelle. Les théorèmes fondamentaux sont dus à Koopman, von Neumann et Carlman, d'une part, et à Birkhoff d'autre part, nous nous référons au Peskir (2000)[57] pour le théorème ergodique ponctuelle de Birkhoff (théorème 1.6, p.8), le théorème ergodique de Von Neumann (théorème 1.7, p.10) et le théorème ergodique sous-additif de Kingman (théorème 1.8, p.11).

La question naturelle qui se pose est : quand les moyennes des grandeurs générés de façon stationnaire convergent ? Dans la situation classique la stationnarité est décrite par une mesure préservant la transformation τ, et l'on considère les moyennes prises sur une séquence $f, f\ o\ \tau, f\ o\ \tau^2, \ldots..$ pour une fonction f intégrable. Cela correspond au concept probabiliste de la stationnarité.

Nous nous sommes intéressés au théorème ergodique, qui s'applique aux processus stationnaires et ergodiques. Le théorème ergodique de Birkhoff, défini pour les processus stationnaires au lieu d'endomorphismes de τ, a maintenant la forme suivante (voir Krengel (1985)[40], p. 26, théorème 4.4.)

Théorème 1.3.1. *Si $Y_0, Y_1, \ldots..$ est un processus réel stationnaire et Y_0 est intégrable, alors*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Y_k = \mathbb{E}[Y_0|\mathcal{F}], \qquad p.s. \tag{1.1}$$

où $\mathcal{F}$ est le σ-algèbre des ensembles invariants. Si de plus, le processus est ergodique la limite coïncide avec l'espérance de la variable Y_0.

Pour la convergence en L^2 de (1.1) et de sa version uniforme (en temps discret et continu) nous renvoyons au papier du Peskir (1998)[56].

Nous nous servirons d'une version de ce théorème pour les processus strictement stationnaires et ergodiques à temps continu.

Par souci de clarté et de compréhension, nous présentons quelques détails qui définissent la propriété ergodique des processus en temps continu.

Soit $X = \{X_t\}_{t\geq 0}$ un processus à temps continu prenant ses valeurs dans un espace mesurable $(E, \mathcal{X})$ sur lequel est définie une mesure de probabilité μ. Pour $\delta > 0$, soit Υ^δ une transformation δ, i.e., $(\Upsilon^\delta(x))_s = x_{s+\delta}$.

Définition 1.3.1. *(δ-erogodicité). Un processus en temps continu $X = \{X_t\}_{t\geq 0}$ est dit δ-ergodique, si chaque ensemble mesurable δ-invariant lié au processus X a une probabilité soit 1 ou 0. Autrement dit, pour tout ensemble δ-invariant $\mathcal{A}$, $\mu(A) = (\mu(\mathcal{A}))^2$.*

Définition 1.3.2. *Un processus en temps continu $X = \{X_t\}_{t\geq 0}$ est ergodique s'il est δ-ergodique pour tout $\delta > 0$.*
Alors le théorème ergodique ponctuel est donné sous la forme

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T Y_t dt = \mathbb{E}[Y_0], \qquad p.s. \tag{1.2}$$

Dans notre travail nous utilisons une version fonctionnelle de l'équation (1.2). Delecroix (1987)[30] fut le premier à introduire cette expression sous la forme d'une hypothèse vérifiée par les processus stationnaires, on pourra consulter à ce sujet aussi Gyorfi et al (1989)([35], théorème 3.5.1), Delecroix et al (1991)[31] imposent l'existence des densités conditionnelles, $f_{X_i}^{\mathcal{F}_{i-1}}$, des variables de X_i par rapport aux tribu ($\mathcal{F}_{i-1} = \sigma(X_{i-j}, j \geq 1)$) et en posant $U_n = \frac{1}{n}\sum_{i=1}^n f_{X_i}^{\mathcal{F}_{i-1}}$, ils utilisent les hypothèses suivantes pour établir les convergences souhaitées

- (1) $f \in \mathcal{C}^0(\mathbb{R}^d), \quad f_{X_i}^{\mathcal{F}_{i-1}} \in \mathcal{C}^0(\mathbb{R}^d)$
- (2) $\sup\limits_{x\in\mathbb{R}^d} |U_n(x) - f(x)| \underset{n\to\infty}{\overset{p.s}{\to}} 0.$

L'hypothèse (2) implique, pour tout $x \in \mathbb{R}^d$, que
$(2)^* \; U_n(x) - f(x) \underset{n\to\infty}{\overset{p.s}{\to}} 0.$
Notre approche consiste à introduire une version de $(2)^*$ dans le cas continu.

Parfois, on remplace les densités conditionnelles par d'autres fonctions aléatoires liées au processus et obtenues par le calcul des probabilités des petites boules.

La variété des phénomènes modélisés par des processus ergodiques motive notre travail et montre son importance. Les processus ergodiques sont largement utilisés dans la modélisation de problématiques appliqués. Sans être exhaustif, nous citons à titre d'exemples

- sciences biomédicales, voir Banks (1975)[5],
- économie, voir Bergstrom (1990)[9], Hale et Verduyn Lunel (1993)[36],
- analyse génétique, voir Lange (2002)[41],
- mécaniques, voir Arnold (1973)[4],
- physiques, voir Papanicolaou (1995)[53],
- mathématiques financières, voir Karatzas et Shreve (1991)[39],
- démographie, voir Cohen (1979)[24].

La description de certains phénomènes par des processus à temps discret peut provoquer une importante perte d'information. Le préjudice subi alors peut remettre en cause toute la modélisation mathématique. De ce fait, un choix de modèle en temps continu peut être déterminant pour la bonne description des phénomènes et les applications à la prévision de leurs évolutions.

1.4 Convergence

Définition 1.4.1. *(Convergence en probabilité). Soit I un intervalle de $\mathbb{R}$, (X_n) une suite de fonctions définies sur I, et X définie sur I. On dit que (X_n) converge en probabilité $(\mathbb{P})$ vers X sur I,*

$$\forall \epsilon > 0 \qquad \lim_{n\to\infty} \mathbb{P}\{|(X_n, X)| \geq \epsilon\} = 0.$$

Définition 1.4.2. *(Convergence presque sûre). Soit I un intervalle de $\mathbb{R}$, (X_n) une suite de fonctions définies sur I, et X définie sur I. On dit que (X_n) converge presque sûrement vers X sur I, si*

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

Remarque 1.4.1. *La convergence presque sûre entraîne la convergence en probabilité.*

Définition 1.4.3. *(Convergence uniforme). Soit I un intervalle de $\mathbb{R}$, (X_n) une suite de fonctions définies sur I, et X définie sur I. On dit que (X_n) converge uniformément vers X sur I si*

$$\forall \epsilon > 0, \quad \exists\, \omega \in I, \forall n \geq n_0, \, |X_n(\omega) - X(\omega)| < \epsilon.$$

Définition 1.4.4. *Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements on note*

$$\limsup_{n\to\infty} A_n = \cap_{n=0}^{\infty}(\cup_{k=n}^{\infty} A_k),$$

$$\liminf_{n\to\infty} A_n = \cup_{n=0}^{\infty}(\cap_{k=n}^{\infty} A_k).$$

Lemme 1.4.1. *(Lemme de Borel-Cantelli). Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.*

(i) *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \leq \infty$,alors*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0,$$

ou de manière équivalente,

$$\text{presque sûrement,} \qquad \{n \in \mathbb{N}; \omega \in A_n\} \quad \text{est fini.}$$

(ii) *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$ et si les événements A_n sont indépendants, alors*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 1,$$

ou de manière équivalente,
presque sûrement, $\{n \in \mathbb{N}; \omega \in A_n\}$ *est infini.*

Le lemme suivant est un type lemme de Borel-Cantelli pour les processus en temps continu.

Lemme 1.4.2. *Soit $(U_t, t \geq 0)$ un processus réel à temps continu de telle sorte que nous avons*

a) *Pour tout $\mu > 0$, il existe une fonction réelle décroissante ψ_μ intégrable sur $[0, +\infty)$ et satisfaisante*

$$\mathbb{P}(|U_t| \geq \mu) \leq \psi_\mu(t), \qquad t \geq 0,$$

b) *Les trajectoires de U_t sont uniformément continue avec une probabilité 1, alors*

$$\lim_{t\to\infty} U_t = 0, \qquad p.s.$$

1.5 Inégalités exponentielles

Lemme 1.5.1. *(Peña, V. H. and Giné, E. (1999).[55]). Soit $(Z_n)_{n\geq 1}$ une suite de différences de martingales réelles par rapport à la séquence de σ-algèbre $(\mathcal{F}_n = \sigma(Z_1, \cdots\cdots, Z_n))_{n\geq 1}$, où $\sigma(Z_1, \cdots\cdots, Z_n)$ est la tribu engendrée par les variables aléatoires $Z_1, \cdots\cdots, Z_n$. On pose $S_n = \sum_{i=1}^n Z_i$. Pour tout $p \geq 2$ et tout $n \geq 1$, supposons qu'il existe certaines constantes non négatif C et d_n telles que*

$$\mathbb{E}(Z_n^p|\mathcal{F}_{n-1}) \leq C^{p-2} p! d_n^2, \qquad \textit{presque sûrement.} \tag{1.3}$$

Alors, pour tout $\epsilon > 0$, on a

$$\mathbb{P}(|S_n| > \epsilon) \leq 2\exp\left\{ -\frac{\epsilon^2}{2(D_n + C\epsilon)} \right\},$$

où $D_n = \sum_{i=1}^n d_i^2$.

Le prochain lemme nous donne une inégalité exponentielle similaire pour des sommes partielles de différences de martingales bornées.

Lemme 1.5.2. *(Laïb(1999)[43]). Soit $\{Z_n : n \in \mathbb{N}\}$ une suite des différences de martingale par rapport à la filtration $\{\mathcal{F}_n = \sigma(Z_1, \cdots\cdots, Z_n) : i \in \mathbb{N}\}$, où $\sigma(Z_1, \cdots\cdots, Z_n)$ est la tribu engendrée par les variables aléatoires $Z_1, \cdots\cdots, Z_n$, telles que $\| Z_n \| \leq B$ p.s. pour $1 \leq i \leq n$. Pour tout $\epsilon > 0$, on a*

$$\mathbb{P}\left\{ \max_{1\leq i\leq n} \left\| \sum_{j=1}^{i} Z_j \right\| > \epsilon \right\} \leq 2\exp\left\{ -\frac{\epsilon^2}{2nB^2} \right\}. \tag{1.4}$$

Chapitre 2

Estimation de la fonction de densité d'un processus stationnaire à temps continu : Consistance ponctuelle et uniforme presque sûre

Dans ce chapitre, nous nous intéressons au comportement asymptotique de l'estimateur à noyau de la densité d'un processus stationnaire à temps continu. Sous des conditions de dépendance assez générales, nous établissons les convergences presque sûres, ponctuelles et uniformes avec des vitesses de convergence. Ce chapitre est divisé en quatre sections ; nous introduisons quelques notations générales et hypothèses dans la première section. La deuxième et troisième sections sont consacrées à la consistance ponctuelle presque sûre et la consistance uniforme presque sûre. Les démonstrations détaillées de ces résultats sont données dans la dernière section.

2.1 Notations générales et hypothèses

Soit $X = (X_t)_{t\geq 0}$ un processus réel stationnaire à temps continu. Sur la base de la partie $(X_t)_{0\leq t\leq T}$ observée du processus X, admettant une densité

f. Pour un T fixé, l'estimateur à noyau de la densité f de type Parzen-Rosenblatt est défini, pour tout $x \in \mathbb{R}$, par

$$\hat{f}_T(x) = \frac{1}{Th_T} \int_0^T K\left(\frac{x - X_t}{h_T}\right) dt,$$

où K est un noyau et h_T est le paramètre de lissage, c'est-à-dire, une suite de réels positifs ayant tendant au zéro avec un certain taux qui sera donné plus tard.

Afin d'afficher nos résultats, certaines notations sont nécessaires. Dans ce qui suit, pour un nombre réel positif δ tel que $n = \frac{T}{\delta} \in \mathbb{N}$, considérons la partition σ-algèbre $(T_j)_{0 \leq j \leq n}$ de l'intervalle $[0, T]$. De plus, pour $t > 0$ et $1 \leq j \leq n$, considérons les σ-algèbre $\mathcal{F}_t = \sigma((X_s), 0 \leq s < t)$ et $\mathcal{G}_j = \sigma((X_s), 0 \leq s \leq T_j)$. Pour tout $t > s$ définir $f_{X_t}^{\mathcal{F}_s}$ comme la densité conditionnelle de X_t étant donné le σ-algèbre $\mathcal{F}_s$ et l'ensemble $f^{\mathcal{F}_{t-\delta}} = f_{X_t}^{\mathcal{F}_{t-\delta}}$. Quand $s < 0$, $\mathcal{F}_s$ est le trivial σ-algèbre. Pour une variable aléatoire réelle ξ et tout $k \in \mathbb{N}$, on définit la projection $\mathcal{P}_k$ par $\mathcal{P}_k\xi = \mathbb{E}[\xi \mid \mathcal{F}_k] - \mathbb{E}[\xi \mid \mathcal{F}_{k-1}]$, où $\mathbb{E}[\xi \mid \mathcal{F}_k]$ est l'espérance conditionnelle de ξ sachant σ-algèbre $\mathcal{F}_k$.

En suite, nous supposons que notre modèle satisfait les hypothèses suivantes :

(A.1) **(i)** K est un noyau positif borné et $\int K(z)dz = 1$,

(ii) $\int \mid z \mid K(z)dz < \infty$,

(iii) Le noyau K est une fonction Lipschitzienne.

(A.2) **(i)** La densité f est différentiable avec une dérivée bornée.

(ii) Pour tout $\delta > 0$, la densité conditionnelle $f^{\mathcal{F}_{t-\delta}}$ est différentiable avec presque sûrement des dérivées bornées .

(iii) Pour tout $t \in [0, T]$, tout $\delta > 0$ et tout $x \in \mathbb{R}$, la fonction $f^{\mathcal{F}_{t-\delta}}(x)$ est presque sûrement bornée par une fonction déterministe $b_{t,\delta}(x)$.

(iv) Pour tout $\delta > 0, T^{-1} \int_0^T b_{t,\delta}(x)dt \to D(x) \neq 0$ quand $T \to \infty$.

(A.3) Pour tout δ, Nous avons

$$\sup_y \int_{\mathbb{R}^+} \|\mathcal{P}_1 f^{\mathcal{F}_{t-\delta}}(y)\|^2 dt < \infty.$$

Commentaires sur les hypothèses

Les hypothèses (A.1) sont très standard dans le cadre non paramétrique. Les hypothèses (A.2) imposent la régularité nécessaire sur le marginal et la densité conditionnelle pour atteindre les taux de convergence donnés ci-dessous. Approcher l'intégrale $\int_{\mathbb{R}^+} \|\mathcal{P}_1 f^{\mathcal{F}_{t-\delta}}(y)\|^2 dt$ par sa somme de Riemann, l'hypothèse (A.3) est similaire à la condition supposée par Wu (2003)[64] dans le cas discret qui est satisfaite par divers processus incluant linéaire comme ainsi que beaucoup de non linéaires. Pour plus détails et exemples, voir Wu (2003) [64].

2.2 Consistance ponctuelle presque sûre

Le théorème suivant traite la convergence presque sûre simple de $\hat{f}_T$,

Théorème 2.2.1. *On suppose que le processus $(X_t)_{t\geq 0}$ est strictement stationnaire et ergodique, autrement dit, pour tout $\delta > 0$ suffisamment petit, on a*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f_t^{\mathcal{F}_{t-\delta}}(x)dt = f(x), \qquad p.s. \tag{2.1}$$

où $f_t^{\mathcal{F}_{t-\delta}}$ est la densité conditionnelle de la v.a X_t sachant la tribu $\mathcal{F}_{t-\delta}$. En outre, supposons que les hypothèses (A.1)(i)-(ii) soient satisfaites, la densité $f_t^{\mathcal{F}_{t-\delta}}$ est une fonction Lipschitzienne et qu'il existe une constante L telle que

$$0 < L \leq \frac{Th_T^2}{\log T}, \tag{2.2}$$

en suite nous avons

$$\lim_{T\to\infty} \hat{f}_T(x) = f(x), \quad p.s.$$

Remarque 2.2.1. *Ici, l'hypothèse principale est que (X_t) est strictement stationnaire et ergodique représenté par l'equation (2.1), elle prend la forme du théorème ergodique de Birkhoff, épelé pour les processus stationnaires en temps continu, voir Krengel (1985) [40]. La condition de Lipschitzienne supposée sur $f_t^{\mathcal{F}_{t-\delta}}$ est la seule condition de régularité de la densité conditionnelle.*

La vitesse de convergence est donnée par le théorème suivant,

Théorème 2.2.2. *Supposons que les hypothèses (A.1)(i)-(ii), (A.2) et (A.3) soient satisfaites. Si de plus $Th_T/\log T \to \infty$ lorsque $T \to \infty$, alors, pour tout $x \in \mathbb{R}$, nous avons*

$$\hat{f}_T(x) - f(x) = \mathrm{O}(h_T) + \mathrm{O}_{p.s}\left(\left(\frac{\log T}{Th_T}\right)^{1/2}\right), \qquad \textit{lorsque} \quad T \to \infty.$$

2.3 Consistance uniforme presque sûre

À partir de maintenant, soit γ une constante prenant des valeurs dans l'intervalle $[0,1]$ et considérer l'ensemble $\mathcal{B}_T = \{x : |x| < T^\gamma\}$.
Notre résultat uniforme est indiqué et prouvé sous les hypothèses supplémentaires suivantes,

(A.4) Pour toute $T > 0$, $\sup_{x \in \mathcal{B}_T} D(x) = D_0 < \infty$.

(A.5) Il existe une constante $C > 0$ telle que pour T suffisamment grand,

$$\frac{|h'_T| T^{1/2}}{h_T^{3/2} (\log T)^{1/2}} \leq C,$$

où h'_t est la dérivée de h_T.

(A.6) Le noyau K est de support borné.

Le théorème suivant montre la convergence uniforme presque sûre,

Théorème 2.3.1. *Soit $(X_t)_{t\geq 0}$ un processus strictement stationnaire et ergodique tel que, pour tout $\delta > 0$ suffisamment petit,*

$$\lim_{T\to\infty} \sup_{x\in\mathbb{R}} \left| \frac{1}{f(x)} \frac{1}{T} \int_0^T f_t^{\mathcal{F}_{t-\delta}}(x)dt - 1 \right| = 0, \quad p.s. \tag{2.3}$$

En outre, on suppose que les hypothèses (A.1) et (A.6) soient satisfaites, et que la densité conditionnelle $f_t^{\mathcal{F}_{t-\delta}}$ est une fonction de Lipschitzienne et il existe deux constantes L et A de telle sorte que

$$(i) \quad 0 < L \leq \frac{Th_T^2}{\log T}, \qquad (ii) \quad \frac{h'_T}{h_T^2} \leq A < \infty, \tag{2.4}$$

avec h'_T est la dérivée de h_T par rapport à T, alors,

$$\lim_{T\to\infty} \sup_{x\in\mathbb{R}} \left| \hat{f}_T(x) - f(x) \right| = 0, \quad p.s.$$

Remarque 2.3.1. *Comme dans le théorème 2.3.1, le processus* (X_t) *est supposé être strictement stationnaire et ergodique. Dans le temps discret, au meilleur de nos connaissances, ce genre d'hypothèses, donné dans l'équation (2.3), est apparu pour la première fois dans le Delecroix (1980) [28]. il a été utilisé dans plusieurs documents parmi lesquels nous citons Györfi et al (1989) [35], Ould Said (1997) [52] et Laib (2005) [44]. La discussion de cette condition est effectuée dans Delecroix (1987) [30].*

Le dernier résultat de ce chapitre se présente sous la forme suivante

Théorème 2.3.2. *Supposons que les hypothèses (A.1)-(A.6) soient satisfaites. Si de plus* $Th_T/\log T \to \infty$ *quand* $T \to \infty$*, on a*

$$\sup_{x\in\mathbb{R}} |\hat{f}_T(x) - f(x)| = \mathrm{O}(h_T) + \mathrm{O}_{p.s}\left(\left(\frac{\log T}{Th_T}\right)^{1/2}\right), \quad \textit{lorsque} \quad T \to \infty.$$

Cette vitesse de convergence peut être améliorée comme le montre la remarque suivante

Remarque 2.3.2. *Sous les hypothèses du théorème 2.3.2, nous supposons, en outre, que la densité* f *est* s *fois dérivable et que sa dérivée d'ordre* s *est bornée, le noyau* K *est d'ordre* s*, alors nous obtenons le résultat suivant*

$$\sup_{x\in\mathbb{R}} |\hat{f}_T(x) - f(x)| = \mathrm{O}(h_T^s) + \mathrm{O}_{p.s}\left(\left(\frac{\log T}{Th_T}\right)^{1/2}\right) \quad \textit{lorsque} \quad T \to \infty.$$

2.4 Preuves

A partir de maintenant, pour tout $1 \leq j \leq n$, tout $T \geq 0$ et tout $\delta > 0$, on a

$$\bar{f}_T(x) = \frac{1}{T}\int_0^T \mathbb{E}\left[\frac{1}{h_T}K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt,$$

$$Z_j = Z_j(x) = \int_{T_{j-1}}^{T_j} K\left(\frac{x-X_t}{h_T}\right)dt, \tag{2.5}$$

$$Y_j = Y_j(x) = \int_{T_{j-1}}^{T_j} \left(K\left(\frac{x-X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt. \tag{2.6}$$

Remarquons, pour tout $\delta > 0$, que $(Y_j)_{1\leq j\leq n}$ est une suite de différence de martingale par rapport à σ-algèbre $(\mathcal{G}_j)_{1\leq j\leq n}$. En effet, puisque, pour tout $t \in [T_{j-1}, T_j]$, $\mathcal{G}_{j-2} \subseteq \mathcal{F}_{t-\delta} \subseteq \mathcal{G}_{j-1}$, il est clair que Y_j est $\mathcal{G}_j$ -mesurable et satisfait

$$\mathbb{E}[Y_j|\mathcal{G}_{j-2}] = \mathbb{E}\bigg[\int_{T_{j-1}}^{T_j}\bigg(K\bigg(\frac{x-X_t}{h_T}\bigg) - \mathbb{E}\bigg[K\bigg(\frac{x-X_t}{h_T}\bigg)|\mathcal{F}_{t-\delta}\bigg]\bigg)dt|\mathcal{G}_{j-2}\bigg] = 0.$$

Pour établir des résultats ponctuels et uniformes presque sûrs, nous avons besoin d'une inégalité exponentielle pour des sommes partielles de différences de martingale non bornées. Cette inégalité est donnée dans les lemmes 1.5.1, 1.5.2

Preuve du Théorème 2.2.1

Observez maintenant que, pour tout $x \in \mathbb{R}$, on peut écrire

$$\begin{aligned}\hat{f}_T(x) - f(x) &= \frac{1}{Th_T}\int_0^T \bigg[K\bigg(\frac{x-X_t}{h_T}\bigg) - \mathbb{E}\bigg[K\bigg(\frac{x-X_t}{h_T}\bigg)|\mathcal{F}_{t-\delta}\bigg]\bigg]dt \\ &+ \frac{1}{Th_T}\int_0^T \mathbb{E}\bigg[K\bigg(\frac{x-X_t}{h_T}\bigg)|\mathcal{F}_{t-\delta}\bigg]dt - f(x) \\ &= R_{1,T_n}(x) + R_{2,T_n}(x).\end{aligned} \tag{2.7}$$

Utilisant la décomposition (2.7) et (2.6), notons que $T = T_n = \delta n$ et $h_{T_n} = h_T$, nous pouvons écrire

$$R_{1,T_n}(x) = \frac{1}{T_n h_{T_n}}\int_0^T \bigg[K\bigg(\frac{x-X_t}{h_{T_n}}\bigg) - \mathbb{E}\bigg[K\bigg(\frac{x-X_t}{h_{T_n}}\bigg)|\mathcal{F}_{t-\delta}\bigg]\bigg]dt \tag{2.8}$$

$$= \frac{1}{T_n h_{T_n}}\sum_{k=1}^{n} Y_k, \tag{2.9}$$

puisque le noyau K est borné, alors

$$\begin{aligned}\frac{1}{T_n h_{T_n}}|Y_k| &\leq \frac{1}{T_n h_{T_n}}\int_{T_{j-1}}^{T_j} \bigg[K\bigg(\frac{x-X_t}{h_{T_n}}\bigg) - \mathbb{E}\bigg[K\bigg(\frac{x-X_t}{h_{T_n}}\bigg)|\mathcal{F}_{t-\delta}\bigg]\bigg]dt \\ &\leq \frac{2\delta\tilde{K}}{T_n h_{T_n}},\end{aligned}$$

avec $\tilde{K} = \max_{u\in\mathbb{R}} K(u)$. Observons que la suite $\{Y_k\}_{1\leq k\leq n}$ est une suite de différence de martingale par rapport à σ-algèbre $\mathcal{G}_{k-1}$. Rappelons $T = \delta n = T_n$ et $h_T = h_{T_n}$, en appliquant le lemme 1.5.2 pour n suffisamment grand et tout $\epsilon > 0$, on obtient

$$\begin{aligned}
\mathbb{P}\Big\{\Big|\sum_{k=1}^{n} Y_k\Big| > \epsilon(T_n h_{T_n})\Big\} &\leq 2\exp\Big\{-\frac{\epsilon^2(T_n h_{T_n})^2}{8n\delta^2\tilde{K}^2}\Big\}\\
&= 2\exp\Big\{-\frac{\epsilon^2 T_n h_{T_n}^2}{8\delta\tilde{K}^2\log T_n}\log T_n\Big\}\\
&= 2\exp\Big\{\log T_n^{-C_k\frac{T_n h_{T_n}^2}{\log T_n}}\Big\}\\
&= \frac{2}{T_n^{C_k u(T_n)}}.
\end{aligned}$$

Où $C_k = \frac{\epsilon^2}{8\delta\tilde{K}^2}$ est une constante positive, selon la condition (2.2) la fonction $u(T_n) = \frac{T_n h_{T_n}^2}{\log T_n}$ est bornée par une constante L strictement positive. Par le lemme Borel-Cantelli, on obtient

$$R_{1,T_n}(x) \overset{p.s}{=} 0, \qquad \text{lorsque} \qquad T_n \to \infty. \tag{2.10}$$

Nous nous tournons notre attention sur le deuxième terme de la décomposition (2.7), sous la condition (2.1) et comme $f_t^{\mathcal{F}_{t-\delta}}$ est une fonction de Lipschitzienne, par un changement de variable, on obtient

$$\begin{aligned}
R_{2,T_n}(x) &= \frac{1}{Th_T}\int_0^T \mathbb{E}\Big[K\Big(\frac{x-X_t}{h_T}\Big)|\mathcal{F}_{t-\delta}\Big]dt - f(x)\\
&= \frac{1}{Th_T}\int_0^T\int_{\mathbb{R}} K\Big(\frac{x-z}{h_T}\Big) f_t^{\mathcal{F}_{t-\delta}}(z)dzdt - f(x)\\
&= \frac{1}{T}\int_0^T\int_{\mathbb{R}} K(u)\Big(f_t^{\mathcal{F}_{t-\delta}}(x-h_T u) - f_t^{\mathcal{F}_{t-\delta}}(x)\Big)dudt\\
&\quad + \frac{1}{T}\int_0^T f_t^{\mathcal{F}_{t-\delta}}(x)dt - f(x)\\
&\leq \frac{1}{T}\int_0^T f_t^{\mathcal{F}_{t-\delta}}(x)dt - f(x) + \mathrm{O}(h_T) = o(1).
\end{aligned}$$

Alors,

$$R_{2,T_n}(x) \overset{p.s}{=} 0, \qquad \text{quand} \qquad T \to \infty. \tag{2.11}$$

En conjonction avec les équations (2.10) et (2.11), la preuve de Théorème 2.2.1 est achevée. ■

Preuve du Théorème 2.2.2

Théorème 2.2.2 est une conséquence directe des lemmes suivants :

Lemme 2.4.1. *Sous les hypothèses (A.1)-(A.2), pour $x \in \mathbb{R}$, on a*

$$R_{1,T}(x) = \mathrm{O}\left(\left(\frac{\log T}{Th_T}\right)^{\frac{1}{2}}\right) \qquad p.s. \tag{2.12}$$

Lemme 2.4.2. *Sous des hypothèses (A.1)(i)-(ii), (A.2)(i) et (A.3), pour $x \in \mathbb{R}$, on a*

$$R_{2,T_n}(x) = \mathrm{O}_{p.s}(T^{-1/2}) + \mathrm{O}(h_T). \tag{2.13}$$

Preuve du lemme2.4.1

Pour la démonstration, on a besoin d'utiliser le lemme 1.5.1 et en appliquant les inégalités de Minkowski et Jensen, on a

$$|\mathbb{E}[Y_j^p|\mathcal{G}_{j-2}]| \leq \mathbb{E}\left[\left|\int_{T_{j-1}}^{T_j}\left(K\left(\frac{x-X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right|^p|\mathcal{G}_{j-2}\right]$$

$$\begin{aligned}
&\leq \int_{T_{j-1}}^{T_j} \mathbb{E}\left[\left|K\left(\frac{x-X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right|^p|\mathcal{G}_{j-2}\right]dt \\
&\leq \int_{T_{j-1}}^{T_j}\left(\mathbb{E}\left[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{G}_{j-2}\right]^{1/p} + \mathbb{E}\left[\mathbb{E}[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}]^p|\mathcal{G}_{j-2}\right]^{1/p}\right)^p dt \\
&\leq \int_{T_{j-1}}^{T_j}\left(\mathbb{E}\left[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{G}_{j-2}\right]^{1/p} + \mathbb{E}\left[\mathbb{E}[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}]|\mathcal{G}_{j-2}\right]^{1/p}\right)^p dt \\
&= \int_{T_{j-1}}^{T_j}\left(2\mathbb{E}\left[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{G}_{j-2}\right]^{1/p}\right)^p dt = 2^p\int_{T_{j-1}}^{T_j}\mathbb{E}\left[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{G}_{j-2}\right]dt.
\end{aligned}$$

En utilisant le changement de variable $u = (x-y)/h_T$, sous les hypothèses (A.2)(ii)-(iii), avec $b_{j-2,\delta} = b_{T_{j-2},\delta}$, on a

$$\begin{aligned}\mathbb{E}\left[K^p\left(\frac{x-X_t}{h_T}\right)|\mathcal{G}_{j-2}\right]dt &= \int K^p\left(\frac{x-y}{h_T}\right)f_t^{\mathcal{G}_{j-2}}(y)dy \\ &= h_T\int K^p(u)f_t^{\mathcal{G}_{j-2}}(x-h_Tu)du \\ &\leq h_T \parallel K \parallel_\infty^{p-1}\int K(z)\left(f_t^{\mathcal{G}_{j-2}}(x)-h_Tz(f_t^{\mathcal{G}_{j-2}})'(x_T^*(z))\right)du \\ &= h_T \parallel K \parallel_\infty^{p-1}\,(f_t^{\mathcal{G}_{j-2}}(x)+o(1)) \\ &\leq h_T \parallel K \parallel_\infty^{p-1}\,(b_{j-2,\delta}(x)+o(1)),\end{aligned} \tag{2.14}$$

avec $x^* \in [x-h_Tz, x]$.
Par conséquent, on a

$$|\mathbb{E}[Y_j^p|\mathcal{G}_{j-2}]| \leq 2^p\delta h_{T_n}\|K\|_\infty^{p-1}(b_{j-2,\delta}(x)+o(1)) \leq p!C^{p-2}d_j^2,$$

oú $C = 2\|K\|_\infty$, $d_j^2 = 2^2\delta h_{T_n}\|K\|_\infty b_{j-2,\delta}(x)$. Notez que l'intégrale $\int_0^T b_{t,\delta}(x)dt$ peut être approché par la somme de Riemann $\delta\sum_{j=2}^n b_{j-2,\delta}(x)$. Sous l'hypothèse (A.2)(iv), on a

$$\begin{aligned}\mathcal{D}_n &= \sum_{i=2}^n d_j^2 = 4h_{T_n} \parallel K \parallel_\infty \left(\delta\frac{1}{n}\sum_{j=2}^n b_{j-2,\delta}(x)\right) \\ &\leq 4T_nh_{T_n} \parallel K \parallel_\infty \left(\frac{1}{T}\int_0^T b_{t,\delta}(x)dt\right) = 4T_nh_{T_n} \parallel K \parallel_\infty \mathrm{O}\big(D(x)\big).\end{aligned}$$

D'après le lemme 1.5.1, pour tout $\epsilon > 0$, on a

$$\begin{aligned}
&\mathbb{P}\left(|R_{1,T_n}(x)| > \epsilon\left(\frac{\log T_n}{T_n h_{T_n}}\right)^{\frac{1}{2}}\right) = \mathbb{P}\left(\left|\sum_{j=1}^{n}\right| > (T_n h_{T_n})\epsilon\left(\frac{\log T_n}{T_n h_{T_n}}\right)^{\frac{1}{2}}\right)\\
&\leq 2\exp\left\{-\frac{\epsilon^2(T_n h_{T_n})^2(\log T_n/T_n h_{T_n})}{8T_n h_{T_n} \parallel K \parallel_\infty \mathrm{O}\big(D(x)\big) + 4 \parallel K \parallel_\infty T_n h_{T_n}\epsilon(\log T_n/T_n h_{T_n})^{\frac{1}{2}}}\right\}\\
&\leq 2\exp\left\{-\frac{\epsilon^2(T_n h_{T_n})^2(\log T_n/T_n h_{T_n})}{8D(x)T_n h_{T_n} \parallel K \parallel_\infty (1+\frac{\epsilon}{2D(x)}(\log T_n/T_n h_{T_n})^{\frac{1}{2}})}\right\}\\
&\leq 2\exp\left\{-\frac{\epsilon^2}{8C_1 D(x) \parallel K \parallel_\infty}\log T_n\right\}\\
&= 2\exp\left\{\log\left(T_n^{-\frac{\epsilon^2}{8C_1 D(x)\|K\|_\infty}}\right)\right\}\\
&= 2T_n^{-\frac{\epsilon^2}{8C_1 D(x)\|K\|_\infty}}
\end{aligned} \tag{2.15}$$

Il suffit alors de considérer un choix approprié de ϵ et d'utiliser le lemme Borel-Cantelli pour achever la preuve.

Preuve du lemme 2.4.2

Pour chaque $x \in \mathbb{R}$, considérons la décomposition suivante :

$$R_{2,T_n}(x) = \bar{f}_T(x) - \mathbb{E}[\hat{f}_T(x)] + \mathbb{E}[\hat{f}_T(x)] - f(x) = B_T^{(1)}(x) + B_T^{(2)}(x).$$

Soit $K_h(.) = 1/hK(./h)$, on a

$$\begin{aligned}
|TB_T^{(1)}(x)| &= \left|\int_0^T \Big(\mathbb{E}\big[K_h(x-X_t)|\mathcal{F}_{t-\delta}\big] - \mathbb{E}\big[K_h(x-X_t)\big]\Big)dt\right|\\
&= \left|\int_{\mathbb{R}} K_h(x-y)\Big(\int_0^T f^{\mathcal{F}_{t-\delta}}(y)dt - Tf(y)\Big)dy\right|\\
&= \left|\int_{\mathbb{R}} K_h(x-y)H_T(y)dy\right|\\
&\leq \int_{\mathbb{R}} K_h(x-y) \parallel H_T(y) \parallel dy.
\end{aligned}$$

Considérons la projection $\mathcal{P}_k$ définie ci-dessus, et on utilise l'inégalité de Cauchy-Schwarz, on obtient

$$\parallel H_T(y) \parallel^2 = \left\|\int_0^T f^{\mathcal{F}_{t-\delta}}(y)dt - Tf(y)\right\|^2 = T\Big(\int_0^T \parallel \mathcal{P}_1 f^{\mathcal{F}_{t-\delta}}(y) \parallel^2 dt\Big),$$

Donc

$$\begin{aligned}\sup_y \| H_T(y) \|^2 &\leq T\biggl(\sup_y \int_0^T \| \mathcal{P}_1 f^{\mathcal{F}_{t-\delta}}(y) \|^2 \, dt\biggr)\\ &\leq T\biggl(\sup_y \int_{\mathbb{R}^+} \| \mathcal{P}_1 f^{\mathcal{F}_{t-\delta}}(y) \|^2 \, dt\biggr).\end{aligned}$$

Ensuite, en supposant la condition (A.3), on obtient

$$\sup_y \| H_T(y) \|^2 = \mathrm{O}(T). \tag{2.16}$$

Alors

$$B_T^{(1)}(x) = \mathrm{O}_{p.s}(T^{-\frac{1}{2}}).$$

De plus, en supposant les hypothèses (A.1) (ii) et (A.2) (i), on voit facilement que

$$\begin{aligned}B_T^{(2)}(x) &= \mathbb{E}[\hat{f}_T(x)] - f(x) = \frac{1}{h_T}\int_{\mathbb{R}} K\left(\frac{x-y}{h_T}\right) f(y)dy - f(x)\\ &= \int_{\mathbb{R}} K(z)(f(x) - h_T z f'(x^*))dz - f(x)\\ &= \mathrm{O}(h_T),\end{aligned}$$

avec $x^* \in [x - h_T z, x]$

Par conséquent, on a

$$R_{2,T_n}(x) = \mathrm{O}_{p.s}(T^{-1/2}) + \mathrm{O}(h_T).$$

Preuve du Théorème 2.2.2 En utilisant la décomposition (2.7), la preuve suit comme une conséquence directe du Lemme 2.4.1 et du Lemme 2.4.2. ■

Preuve du Théorème 2.3.1

L'idée de la preuve est de considérer une couverture de

$$\mathcal{B}_T = \{x \in \mathbb{R} : |x| \leq T^\nu\}.$$

Soit ν_T une fonction entière non décroissante tendant vers l'infini lorsque $T \to \infty$. considérons la partition $\{B_{T,i}\}_{1\leq i \leq T}$ de l'ensemble B_T définie par

$$\mathcal{B}_{T,i} = \{x : \|x - x_i\| \leq T^\gamma \nu_T^{-1}\}$$

où $(x_i)_{1\le i\le \nu_T}$ est une suite d'éléments de $\mathcal{B}_T$. Rappelons que

$$\begin{aligned}\sup_{x\in\mathbb{R}}|\hat{f}_T(x)-f(x)| \le & \sup_{x\in\mathbb{R}}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ & +\sup_{x\in\mathbb{R}}\left|\frac{1}{Th_T}\int_0^T\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt-f(x)\right| \\ = & \mathrm{F}_{T,1}+\mathrm{F}_{T,2}.\end{aligned} \tag{2.17}$$

Lemme 2.4.3. *Sous les hypothèses (A.1) (i) et (A.1) (iii) combinées avec la condition (2.4) nous avons*

$$\lim_{T\to\infty}\mathrm{F}_{T,1}=0, \qquad p.s. \tag{2.18}$$

Preuve :

Rappelons que

$$\begin{aligned}\mathrm{F}_{T,1} = & \sup_{x\in\mathbb{R}}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ \le & \sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ & +\sup_{x\in\mathcal{B}_T^c}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ = & \mathrm{F}_{T,1}^{(1)}+\mathrm{F}_{T,1}^{(2)},\end{aligned} \tag{2.19}$$

en utilisant la couverture de l'ensemble compact B_T nous pouvons écrire

$$\begin{aligned}\mathrm{F}_{T,1}^{(1)} = & \sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ \le & \max_{1\le k\le\nu_T}\sup_{x\in\mathcal{B}_{T,k}}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-K\left(\frac{x_k-X_t}{h_T}\right)\right)dt\right| \\ & +\max_{1\le k\le\nu_T}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x_k-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x_k-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ & +\max_{1\le k\le\nu_T}\sup_{x\in\mathcal{B}_{T,k}}\left|\frac{1}{Th_T}\int_0^T\left(\mathbb{E}\left[K\left(\frac{x_k-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\ = & \mathrm{F}_{T,1}^{(11)}+\mathrm{F}_{T,1}^{(12)}+\mathrm{F}_{T,1}^{(13)}.\end{aligned} \tag{2.20}$$

Maintenant, par hypothèse il existe $l > 0$ tel que

$$|K(u) - K(u')| \leq l|u - u'|, \qquad (u, u') \in \mathbb{R}^2.$$

Par conséquent,

$$\begin{aligned}
\mathrm{F}_{T,1}^{(11)} &= \max_{1\leq k\leq \nu_T} \sup_{x\in\mathcal{B}_{T,k}} \left| \frac{1}{Th_T} \int_0^T \left(\mathbb{E}\left[K\left(\frac{x - X_t}{h_T}\right)\right] - \mathbb{E}\left[K\left(\frac{x_k - X_t}{h_T}\right)\right]\right) dt \right| \\
&\leq \frac{1}{h_T^2} \max_{1\leq k\leq \nu_T} \sup_{x\in\mathcal{B}_{T,k}} |x - x_k| \\
&\leq \frac{T^\gamma}{\nu_T h_T^2},
\end{aligned}$$

et de même,

$$\begin{aligned}
\mathrm{F}_{T,1}^{(13)} &= \max_{1\leq k\leq \nu_T} \sup_{x\in\mathcal{B}_{T,k}} \left| \frac{1}{Th_T} \int_0^T \left(\mathbb{E}\left[K\left(\frac{x_k - X_t}{h_T}\right) |\mathcal{F}_{t-\delta}\right] - \mathbb{E}\left[K\left(\frac{x - X_t}{h_T}\right) |\mathcal{F}_{t-\delta}\right]\right) dt \right| \\
&\leq \frac{1}{h_T^2} \max_{1\leq k\leq \nu_T} \sup_{x\in\mathcal{B}_{T,k}} |x - x_k| \\
&\leq \frac{T^\gamma}{\nu_T h_T^2},
\end{aligned}$$

puis en choisissant $\nu_T = T^{\gamma+1}/h_T$, il s'ensuit que

$$\mathrm{F}_{T,1}^{(11)} = \mathrm{O}\left(\frac{1}{Th_T}\right), \qquad p.s. \tag{2.21}$$

de plus

$$\mathrm{F}_{T,1}^{(13)} = \mathrm{O}\left(\frac{1}{Th_T}\right), \qquad p.s. \tag{2.22}$$

en utilisant la notation (2.6), nous remarquons

$$\begin{aligned}
\mathrm{F}_{T,1}^{(12)} &= \max_{1\leq k\leq \nu_T} \left| \frac{1}{Th_T} \int_0^T \left(K\left(\frac{x_k - X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{x_k - X_t}{h_T}\right) |\mathcal{F}_{t-\delta}\right]\right) dt \right| \\
&= \max_{1\leq k\leq \nu_T} \left| \frac{1}{Th_T} \sum_{i=1}^n Y_i(x_k) \right|.
\end{aligned}$$

Puisque l'hypothèse (A.6) est vraie, alors nous pouvons écrire

$$\begin{aligned}
\left| Y_i(x_k) \right| &= \left| \frac{1}{Th_T} \int_{T_{i-1}}^{T_i} \left(K\left(\frac{x_k - X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{x_k - X_t}{h_T}\right) |\mathcal{F}_{t-\delta}\right]\right) dt \right| \\
&= 2\,\delta \sup_{y\in\mathbb{R}} K(y) = 2\delta \tilde{K}.
\end{aligned}$$

Maintenant, pour tout $\epsilon > 0$, et par Lemme 1.5.2, on note $T = T_n = \delta n$, $h_T = h_{T_n}$ et $T = T_n$, puis à l'aide de la condition (2.4) on obtient

$$\mathbb{P}\Big(\max_{1\le k\le \nu_{T_n}}\Big|\sum_{i=1}^{n} Y_i(x_k)\Big| > \epsilon(T_n h_{T_n})\Big) \le \sum_{k=1}^{\nu_{T_n}} \mathbb{P}\Big(\Big|\sum_{i=1}^{n} Y_i(x_k)\Big| > \epsilon(T_n h_{T_n})\Big) \tag{2.23}$$

$$\begin{aligned}
&\le 2\nu_{T_n} \exp\left\{-\frac{\epsilon^2 T_n^2 h_{T_n}^2}{8n\delta^2\tilde{K}^2}\right\} \\
&= 2\nu_{T_n} \exp\left\{-\frac{\epsilon^2 T_n h_{T_n}^2}{8\delta\tilde{K}^2 \log T_n}\log T_n\right\} \\
&\le 2\nu_{T_n} \exp\left\{\log T_n^{-\frac{\epsilon^2 L}{8\delta\,\tilde{K}^2}}\right\}
\end{aligned}$$

$$= \frac{2}{T_n^{-\gamma-1+\epsilon^2 L/8\delta\tilde{K}^2} h_{T_n}}, \tag{2.24}$$

en prenant ϵ suffisamment grand pour avoir $T_n^{-\gamma-1+\epsilon^2 L/8\delta\tilde{K}^2} h_{T_n} \to \infty$ lorsque $T_n \to \infty$. Par conséquent, la partie droite de l'equation (2.23) converge, il suffit ensuite d'utiliser le Lemme de Borel-Cantelli pour en déduire que

$$\mathrm{F}_{T,1}^{(12)} = o(1), \qquad p.s. \tag{2.25}$$

En combinant les équations (2.21), (2.22) et (2.25) nous obtenons que

$$\mathrm{F}_{T,1}^{(1)} = o(1), \qquad p.s. \tag{2.26}$$

Nous tournons notre attention maintenant au le deuxième terme de la décomposition (2.19). Il reste à montrer que

$$\mathrm{F}_{T,1}^{(2)} = o(1), \qquad p.s. \tag{2.27}$$

Le fait que $\sup_{0\le t\le T}|X_t| \le T^{2\gamma}/2$ combiné avec la condition $|x| > T^{2\gamma}$ implique évidemment, pour $0 \le t \le T$, $\left|\dfrac{x - X_t}{h_T}\right| > \dfrac{T^{2\gamma}}{2h_T}$. En outre, étant donné que par hypothèse (A.6) le noyau K est de support borné, il s'ensuit qu'il existe

$T_0 > 0$ tel que, pour $T > T_0$, nous avons $K\left(\dfrac{x - X_t}{h_T}\right) = 0$. Par conséquent, nous pouvons écrire

$$\left\{\sup_{0\le t\le T} |X_t| \le \frac{T^{2\gamma}}{2}, |x| > T^{2\gamma}\right\} \subset \left\{\sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\sum_{j=1}^{n} Y_j\right| = 0\right\}.$$

Alors, pour T suffisamment grand, nous avons

$$\mathbb{P}\left\{\sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\sum_{j=1}^{n} Y_j\right| > \eta\right\} \le \mathbb{P}\left(\sup_{0\le t\le T} |X_t| > \frac{T^{2\gamma}}{2}\right) \le \mathbb{P}\left(\sup_{0\le t\le T} |X_t| > T^{\gamma}\right).$$

Maintenant, nous devons montrer que l'application $T \mapsto \sup \mid \dfrac{1}{Th_T}\sum_{j=1}^{n} Y_j \mid$ est uniformément continue. Cette propriété est maintenue à condition qu'il existe une constante $\Lambda > 0$ telle que

$$\sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\sum_{i=1}^{n} Y_i - \frac{1}{Sh_S}\sum_{j=1}^{n'} Y_j'\right| \le \Lambda \mid T - S \mid, \tag{2.28}$$

quand $S = \delta n'$, $S_j = j\delta$ et $Y_j' = Y_j'(x) = \int_{S_{j-1}}^{S_j} \left(K\left(\frac{x-X_s}{h_s}\right) - \mathbb{E}\left[K\left(\frac{x-X_s}{h_s}\right)|\mathcal{F}_{s-\delta}\right]\right)ds$. on prend $Z_j' = \int_{S_{j-1}}^{S_j} K\left(\frac{x-X_s}{h_s}\right)ds$, Observons que

$$\begin{aligned}\sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\sum_{j=1}^{n} Y_j - \frac{1}{Sh_S}\sum_{j=1}^{n'} Y_j'\right| &= \sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\sum_{j=1}^{n} Z_j - \frac{1}{Sh_S}\sum_{j=1}^{n'} Z_j'\right| \\ &\quad + \sup_{|x|>T^{2\gamma}} \left|\frac{1}{Th_T}\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right. \\ &\quad \left. - \frac{1}{Sh_s}\int_0^S \mathbb{E}\left[K\left(\frac{x-X_s}{h_S}\right)|\mathcal{F}_{s-\delta}\right]ds\right| \\ &= \sup_{|x|>T^{2\gamma}} \mid \hat{f}_T(x) - \hat{f}_S(x) \mid \\ &\quad + \sup_{|x|>T^{2\gamma}} \mid \bar{f}_T(x) - \bar{f}_S(x) \mid . \end{aligned} \tag{2.29}$$

Notons que

$$\log(\hat{f}_T(x)) = -\log(Th_T) + \log\left(\int_0^T K\left(\frac{x-X_t}{h_T}\right)dt\right),$$

la différenciabilité par rapport à T, nous donne

$$\begin{aligned}\frac{d\log(\hat{f}_T(x))}{dT} &= -\frac{h_T - Th'_T}{Th_T} - \frac{h'_T}{h_T^2}\bigg(\int_0^T K\Big(\frac{x-X_t}{h_T}\Big)dt\bigg)^{-1}\bigg(\int_0^T (x-X_t)\\ &K'\Big(\frac{x-X_t}{h_T}\Big)dt\bigg) + \bigg(\int_0^T K\Big(\frac{x-X_t}{h_T}\Big)dt\bigg)^{-1} K\Big(\frac{x-X_T}{h_T}\Big).\end{aligned}$$

Rappelons la propriété de la fonction logarithme log :

$$\frac{d\log g(T)}{dT} = \frac{dg(T)/dT}{g(T)},$$

où $g(T)$ est une fonction. Nous obtenons alors

$$\begin{aligned}\frac{d\hat{f}_T(x)}{dT} &= \hat{f}_T(x)\frac{d\log \hat{f}_T(x)}{dT}\\ &= \frac{1}{T^2h_T}\bigg(\int_0^T K\Big(\frac{x-X_t}{h_T}\Big)dt\bigg) + \frac{h'_T}{Th_T^2}\bigg(\int_0^T K\Big(\frac{x-X_t}{h_T}\Big)dt\bigg)\\ &\quad -\frac{h'_T}{Th_T^3}\bigg(\int_0^T (x-X_t)K'\Big(\frac{x-X_t}{h_T}\Big)dt\bigg) + \frac{1}{Th_T}K\Big(\frac{x-X_T}{h_T}\Big).\end{aligned}$$

L'hypothèse (A.6) permet de supposer l'existence d'un constante positive $c_K > 0$, $K'(u) = 0$ chaque fois que $|u| \le c_K$,

$$\begin{aligned}\left|\frac{d\hat{f}_T(x)}{dT}\right| &= \left|\hat{f}_T(x)\frac{d\log \hat{f}_T(x)}{dT}\right|\\ &\le \frac{1}{Th_T}\tilde{K} + \frac{h'_T}{h_T^2}\tilde{K} + \frac{h'_T}{h_T^2}c_K\tilde{K}' + \frac{1}{Th_T}\tilde{K},\end{aligned}$$

Où $\tilde{K} = \sup\limits_{y\in\mathbb{R}} K(y)$ et $\tilde{K}' = \sup\limits_{y\in\mathbb{R}} K(y)$. Par conséquent

$$\sup_{|x|>T^{2\gamma}} |\hat{f}_T(x) - \hat{f}_S(x)| \le \bigg(\frac{2}{Th_T}\tilde{K} + \frac{h'_T}{h_T^2}\big(\tilde{K} + c_K\tilde{K}'\big)\bigg)|T-S|. \tag{2.30}$$

Considérons de la deuxième partie de la décomposition (2.29), d'abord, observons que

$$\log(\bar{f}_T(x)) = -\log(Th_T) + \log\bigg(\int_0^T \mathbb{E}\Big[K\Big(\frac{x-X_t}{h_T}\Big)|\mathcal{F}_{t-\delta}\Big]dt\bigg),$$

par conséquent,

$$\begin{aligned}\frac{d\log(\bar{f}_T(x))}{dT} &= -\frac{1}{T}-\frac{h'_T}{h_T}-\frac{h'_T}{h_T^2}\left(\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)^{-1}\\ &\times\left(\int_0^T \mathbb{E}\left[(x-X_t)K'\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)\\ &+\left(\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)^{-1}\mathbb{E}\left[K\left(\frac{x-X_T}{h_T}\right)|\mathcal{F}_{T-\delta}\right].\end{aligned}$$

En procédant comme ci-dessus, nous obtenons

$$\begin{aligned}\frac{d\bar{f}_T(x)}{dT} &= \bar{f}_T(x)\frac{d\log\bar{f}_T(x)}{dT}\\ &= -\frac{1}{T^2h_T}\left(\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)+\frac{h'_T}{Th_T^2}\left(\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)\\ &-\frac{h'_T}{Th_T^3}\left(\int_0^T \mathbb{E}\left[(x-X_t)K'\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right)+\frac{1}{Th_T}\mathbb{E}\left[K\left(\frac{x-X_T}{h_T}\right)|\mathcal{F}_{T-\delta}\right].\end{aligned}$$

Sous l'hypothèse (A.6), nous avons

$$\left|\frac{d\bar{f}_T(x)}{dT}\right| \leq \frac{2}{Th_T}\tilde{K}+\frac{h'_T}{h_T^2}(\tilde{K}+c_K\tilde{K}').$$

Par conséquent,

$$\sup_{|x|>T^{2\gamma}}|\bar{f}_T(x)-\bar{f}_S(x)| \leq \left(\frac{2}{Th_T}\tilde{K}+\frac{h'_T}{h_T^2}\left(\tilde{K}+c_K\tilde{K}'\right)\right)|T-S|. \tag{2.31}$$

L'inégalité (2.28) résulte de la décomposition (2.29) avec condition (2.4)(ii) et les équations (2.30) et (2.31). Prendre $h_T=T^{-1/5}$ on remarque que

$$\max\left\{\left|\frac{d(\hat{f}_T(x))}{dT}\right|,\left|\frac{d(\bar{f}_T(x))}{dT}\right|\right\} \leq CT^{-\frac{4}{5}},$$

où C est une constante positive. Une application directe du Lemme 1.5.2 combiné avec la condition (2.4) nous donne

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{i=1}^n Y_i\right| > Th_T\epsilon_0\right) &\leq 2\exp\left\{-\frac{T^2h_T^2\epsilon_0^2}{8n\delta^2\tilde{K}^2}\right\}\\ &= 2\exp\left\{\log T^{-\frac{\epsilon_0^2}{8\delta\tilde{K}^2}\frac{Th_T^2}{\log T}}\right\} \leq 2T^{-\frac{\epsilon_0^2 L}{8\delta\tilde{K}^2}}.\end{aligned}$$

Par conséquent, l'application $T \to \sup_{|x|>T^{2\gamma}} \left| \frac{1}{Th_T} \sum_{j=1}^{n} Y_j \right|$ est uniformément continu pour chaque ω et le résultat (2.27) découle de l'utilisation du Lemme de Borel-Cantelli pour processus à temps continue (voir, par exemple, Bosq (1998), Lemme 4.2). nous réalisons la preuve de Lemme 2.4.3 à partir les équations (2.26) et (2.27).

Lemme 2.4.4. *En supposant que la densité conditionnelle $f_t^{\mathcal{F}_{t-\delta}}$ est une fonction de Lipschitzienne avec l'hypothèse (A.1) (ii) et la condition (2.3), nous avons*

$$\lim_{T\to\infty} \mathrm{F}_{T,2} = 0, \qquad p.s. \tag{2.32}$$

Preuve. Observe que

$$\begin{aligned}
\lim_{T\to\infty} \mathrm{F}_{T,2} &= \sup_{x\in\mathbb{R}} \left| \frac{1}{Th_T} \int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right) |\mathcal{F}_{t-\delta}\right] dt - f(x) \right| \\
&= \sup_{x\in\mathbb{R}} \left| \frac{1}{Th_T} \int_0^T \int_{\mathbb{R}} K\left(\frac{x-u}{h_T}\right) f_t^{\mathcal{F}_{t-\delta}}(u) du dt - f(x) \right| \\
&= \sup_{x\in\mathbb{R}} \left| \frac{1}{T} \int_0^T \int_{\mathbb{R}} K(v) f_t^{\mathcal{F}_{t-\delta}}(x - h_T v) dv dt - f(x) \right| \\
&\leq \sup_{x\in\mathbb{R}} \left| \frac{1}{T} \int_0^T \int_{\mathbb{R}} K(v) \big(f_t^{\mathcal{F}_{t-\delta}}(x - h_T v) - f_t^{\mathcal{F}_{t-\delta}}(x) \big) dv dt \right| \\
&\quad + \sup_{x\in\mathbb{R}} \left| \frac{1}{T} \int_0^T \big(f_t^{\mathcal{F}_{t-\delta}}(x) - f(x) \big) dt \right| = \mathrm{F}_{T,2}^{(1)} + \mathrm{F}_{T,2}^{(2)},
\end{aligned}$$

comme $f_t^{\mathcal{F}_{t-\delta}}(x)$ est Lipschitzienne et sous l'hypothèse (A1)(ii) nous obtenons

$$\begin{aligned}
\mathrm{F}_{T,2}^{(1)} &= \frac{1}{T} \int_0^T \int_{\mathbb{R}} K(v) \sup_{x\in\mathbb{R}} \left| f_t^{\mathcal{F}_{t-\delta}}(x - h_T v) - f_t^{\mathcal{F}_{t-\delta}}(x) \right| dv dt \\
&\leq h_T \int_{\mathbb{R}} v K(v) dv = \mathrm{O}(h_T).
\end{aligned} \tag{2.33}$$

De l'autre côté, la condition (2.3) nous donne

$$\mathrm{F}_{T,2}^{(2)} \leq \sup_{x\in\mathbb{R}} \left| \frac{1}{Tf(x)} \int_0^T f_t^{\mathcal{F}_{t-\delta}}(x) dt - 1 \right| = o(1). \tag{2.34}$$

En combinant les inégalités (2.33) et (2.34) nous obtenons

$$\mathrm{F}_{T,2} \overset{\mathbb{P}}{=} o(1), \qquad \text{quand} \qquad T \to \infty.$$

Preuve du Théorème 2.3.1 : la preuve est achevée en combinant la décomposition (2.17) avec les Lemmes 2.4.3 et 2.4.4. ■

Preuve du Théorème 2.3.2

Nous avons besoin ici de la décomposition (2.7) et de quelques résultats intermédiaires et le comportement de chaque terme de cette décomposition. Observons d'abord que

$$\begin{aligned}
\sup_{x\in\mathcal{B}_T}|\hat{f}_T(x)-f(x)| &\leq \sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right| \\
&\quad+\sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt-f(x)\right| \\
&= A_T+B_T.
\end{aligned}$$

Où

$$\begin{aligned}
A_T &= \sup_{x\in\mathcal{B}_T}|A_T(x)|=\sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\left(K\left(\frac{x-X_t}{h_T}\right)-\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]\right)dt\right|. \\
B_T &= \sup_{x\in\mathcal{B}_T}|B_T(x)|=\sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T\mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt-f(x)\right|.
\end{aligned}$$

Soit ν_T une fonction entière non décroissante tendant vers l'infini lorsque $T\to\infty$. Considérons la partition $\{\mathcal{B}_{T,i}\}_{1\leq i\leq\nu_T}$ de l'ensemble $\mathcal{B}_T$ définie par

$$\mathcal{B}_{T,i}=\{x:\parallel x-x_i\parallel\leq T^{\gamma}\nu_T^{-1}\}.$$

Où $(x_i)_{1\leq i\leq T}$ est une suite d'éléments de $\mathcal{B}_T$. Par conséquent, nous avons

$$A_T\leq\max_{1\leq i\leq\nu_T}\sup_{x\in\mathcal{B}_{T,i}}|A_T(x)-A_T(x_i)|+\max_{1\leq i\leq\nu_T}|A_T(x_i)|=I_1+I_2. \tag{2.35}$$

En utilisant l'hypothèse (A.1)(iii), il s'ensuit qu'il existe une constante positive C_K telle que

$$\begin{aligned}
\left|K\left(\frac{x-X_t}{h_T}\right)-K\left(\frac{x_i-X_t}{h_T}\right)\right| &\leq C_K\left|\frac{x-X_t}{h_T}-\frac{x_i-X_t}{h_T}\right| \\
&= \frac{C_K}{h_T}|x-x_i|\leq\frac{C_KT^{\gamma}}{h_T\nu_T}.
\end{aligned}$$

Par conséquent, nous avons

$$\begin{aligned} I_1 &\leq \frac{1}{Th_T}\sum_1^n \int_{T_{j-1}}^{T_j} \Big\{ \max_{1\leq i\leq \nu_T} \sup_{x\in\mathcal{B}_T\bigcap\mathcal{B}_{T,i}} \Big| K\Big(\frac{x-X_t}{h_T}\Big) - K\Big(\frac{x_i-X_t}{h_T}\Big)\Big| \\ &\quad + \mathbb{E}\Big[\max_{1\leq i\leq \nu_T} \sup_{x\in\mathcal{B}_T\bigcap\mathcal{B}_{T,i}} \Big| K\Big(\frac{x-X_t}{h_T}\Big) - K\Big(\frac{x_i-X_t}{h_T}\Big)\Big| |\mathcal{G}_{j-1}\Big]\Big\} dt. \\ &\leq \frac{2\delta n}{Th_T}\times\frac{C_K T^\gamma}{h_T\nu_T} = \frac{2C_K T^\gamma}{h_T^2\nu_T}. \end{aligned}$$

Prenons $\nu_T = T^{\gamma+2}$ et $\epsilon_T = \Big(\frac{\log T}{Th_T}\Big)^{1/2}$, nous obtenons

$$\epsilon_T^{-1} I_1 \leq \Big(\frac{Th_T}{\log T}\Big)^{1/2} \frac{2C_K T^\gamma}{h_T^2 T^{\gamma+2}} = \frac{2C_K}{\big((Th_T)^3 \log T\big)^{1/2}}.$$

Donc,

$$I_1 = o\bigg(\Big(\frac{\log T}{Th_T}\Big)^{1/2}\bigg), \qquad p.s. \tag{2.36}$$

En utilisant la notation (2.6), nous observons que

$$\max_{1\leq i\leq\nu_T} |A_T(x_i)| = \max_{1\leq i\leq \nu_T} |\sum_{j=1}^n Y_j(x_i)|.$$

L'ordre du terme I_2 peut être énonçé de la même manière que dans l'équation (2.12) avec la seule différence que l'hypothèse (A.4) est vraie. Donc, notons $T_n = T = \delta n$, en prenant $\nu_{T_n} = \nu_T = T_n^{\gamma+2}$ et $\epsilon_{T_n} = \epsilon_0\Big(\frac{\log T_n}{T_n h_{T_n}}\Big)^{1/2}$ avec un constante positive ϵ_0, nous obtenons

$$\begin{aligned} &\mathbb{P}\bigg(\max_{1\leq k\leq \nu_{T_n}} \frac{1}{T_n h_{T_n}} \Big|\sum_{i=1}^n Y_i(x_k)\Big| > \epsilon_{T_n}\bigg) \\ &\qquad = \sum_{k=1}^{\nu_{T_n}} \mathbb{P}\bigg(\Big|\sum_{i=1}^n Y_i(x_k)\Big| > (T_n h_{T_n})\epsilon_0 \Big(\frac{\log T_n}{T_n h_{T_n}}\Big)^{1/2}\bigg) \\ &\qquad \leq 2\nu_{T_n} \exp\Big\{ - \frac{\epsilon_0^2 (T_n h_{T_n})^2 (\log T_n/T_n h_{T_n})}{\mathrm{O}(T_n h_{T_n}) + 2CT_n h_{T_n}\epsilon_0(\log T_n/T_n h_{T_n})^{\frac{1}{2}}}\Big\} \\ &\qquad \leq 2\nu_{T_n} \exp\Big\{ -\epsilon_0^2 \mathrm{O}(T_n h_{T_n}) \frac{\log T_n}{T_n h_{T_n}}\Big\} \\ &\qquad = 2\nu_{T_n}\exp\big\{-\epsilon_0^2 C\log T_n\big\} \leq 2T_n^{\gamma+2-\epsilon_0^2 C}, \end{aligned}$$

pour une certaine constante positive C. En procédant comme dans l'inégalité (2.15) et en considérant un choix de ϵ_0 qui permet d'utiliser le Lemme Borel-Cantelli, on obtient

$$I_2 = \mathrm{O}\left(\left(\frac{\log T_n}{T_n h_{T_n}}\right)^{1/2}\right) = \mathrm{O}\left(\left(\frac{\log T}{T h_T}\right)^{1/2}\right), \qquad p.s. \tag{2.37}$$

Il résulte alors les inégalités (2.35)-(2.37) que

$$A_T = \mathrm{O}\left(\left(\frac{\log T}{T h_T}\right)^{1/2}\right), \qquad p.s. \tag{2.38}$$

En outre, tenons compte les hypothèses (A.2) et (A.3) ainsi que l'équation (2.16). nous avons

$$\begin{aligned}
B_T &= \sup_{x\in\mathcal{B}_T}\left|\frac{1}{Th_T}\int_0^T \mathbb{E}\left[K\left(\frac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt - f(x)\right| \\
&\leq \sup_{x\in\mathbb{R}}\left|\frac{1}{T}\int_0^T \mathbb{E}\left[K_h(x-X_t)|\mathcal{F}_{t-\delta}\right]dt - \frac{1}{T}\int_0^T \mathbb{E}\big[K_h(x-X_t)\big]dt\right| \\
&\quad + \sup_{x\in\mathbb{R}}\left|\frac{1}{T}\int_0^T \mathbb{E}\big[K_h(x-X_t)\big]dt - f(x)\right| \\
&= \sup_{x\in\mathbb{R}}\left|\frac{1}{T}\int_{\mathbb{R}} K_h(x-y)H_T(y)dy\right| + \sup_{x\in\mathbb{R}}\left|\int_{\mathbb{R}} K(z)f(x-h_Tz)dz - f(x)\right| \\
&\leq \sup_{x\in\mathbb{R}}\frac{1}{T}\int_{\mathbb{R}} K_h(x-y)\parallel H_T(y)\parallel dy + \sup_{x\in\mathbb{R}}\left|\int_{\mathbb{R}} K(z)(f(x)-h_Tzf'(x^*))dz - f(x)\right| \\
&= \mathrm{O}(\sqrt{T})\sup_{x\in\mathbb{R}}\frac{1}{Th_T}\int_{\mathbb{R}} K\left(\frac{x-y}{h_T}\right)dy + \sup_{x\in\mathbb{R}}\left|h_Tf'(x^*)\int_{\mathbb{R}} K(z)dz\right|.
\end{aligned}$$

Ainsi,

$$B_T = \mathrm{O}_{p.s}(T^{-1/2}) + \mathrm{O}(h_T). \tag{2.39}$$

En combinant l'équation (2.38) avec (2.39), on obtient le résultat suivant

$$\sup_{x\in\mathcal{B}_T}|\hat{f}_T(x) - f(x)| = \mathrm{O}_{p.s}\left(\left(\frac{\log T}{Th_T}\right)^{1/2}\right) + \mathrm{O}_{p.s}(T^{-1/2}) + \mathrm{O}(h_T). \tag{2.40}$$

Notre tâche est maintenant, de gérer le terme $\sup_{x\in\mathcal{B}_T^c}|\hat{f}_T(x)-f(x)|$, où $\mathcal{B}_T^c$ est le complémentaire de l'ensemble de $\mathcal{B}_T$. En procédant comme pour l'équation (2.39), il est clair que

$$\sup_{x\in\mathcal{B}_T^c}|\bar{f}_T(x) - f(x)| = \mathrm{O}_{p.s}(T^{-1/2}) + \mathrm{O}(h_T), \qquad p.s. \tag{2.41}$$

Il reste ensuite à prouver que

$$\lim_{T\to\infty}\sqrt{\frac{\log T}{Th_T}}\sup_{x\in\mathcal{B}_T^c}|\hat{f}_T(x)-\bar{f}(x)|=0,\quad p.s. \tag{2.42}$$

Le fait que $\sup\limits_{0\le t\le T}|X_t|\le T^{2\gamma}/2$ combiné avec la condition $|x|>T^{2\gamma}$ implique évidemment, pour $0\le t\le T$, $\left|\dfrac{x-X_t}{h_T}\right|>\dfrac{T^{2\gamma}}{2h_T}$. De plus, puisque par l'hypothèse (A.6) le noyau K est de support borné, il s'ensuit qu'il existe $T_0>0$ tel que, pour $T>T_0$, nous avons $K\left(\dfrac{x-X_t}{h_T}\right)=0$. Par conséquent, nous pouvons écrire

$$\left\{\sup_{0\le t\le T}|X_t|\le\frac{T^{2\gamma}}{2},|x|>T^{2\gamma}\right\}\subset\left\{\sup_{|x|>T^{2\gamma}}\left|\frac{1}{Th_T}\sum_{j=1}^{n}Y_j\right|=0\right\}.$$

Par conséquent, pour T suffisamment grand, nous avons

$$\mathbb{P}\left\{\sup_{|x|>T^{2\gamma}}\left|\frac{1}{Th_T}\sum_{j=1}^{n}Y_j\right|>\eta\right\}\le\mathbb{P}\left(\sup_{0\le t\le T}|X_t|>\frac{T^{2\gamma}}{2}\right)\le\mathbb{P}\left(\sup_{0\le t\le T}|X_t|>T^{\gamma}\right).$$

Maintenant, nous devons montrer que l'application $T\mapsto\sup\mid\dfrac{1}{Th_T}\sum\limits_{j=1}^{n}Y_j\mid$ est uniformément continue. Cette propriété est maintenue à condition qu'il existe une constante $\Lambda>0$ telle que

$$\sup_{|x|>T^{2\gamma}}\left|\frac{\epsilon_T}{Th_T}\sum_{j=1}^{n}Y_j-\frac{\epsilon_S}{Sh_S}\sum_{j=1}^{n'}Y_j'\right|\le\Lambda\mid T-S\mid, \tag{2.43}$$

quand $S=\delta n'$,$S_j=j\delta$ et $Y_j'=Y_j'(x)=\int_{S_{j-1}}^{S_j}\left(K\left(\frac{x-X_s}{h_s}\right)-\mathbb{E}\left[K\left(\frac{x-X_s}{h_s}\right)|\mathcal{F}_{s-\delta}\right]\right)ds$. et $\epsilon_T=\left(\frac{Th_T}{\log T}\right)$ on prend $Z_j'=\int_{S_{j-1}}^{S_j}K\left(\frac{x-X_s}{h_s}\right)ds$, Observons que

$$\begin{aligned}\sup_{|x|>T^{2\gamma}}\left|\frac{\epsilon_T}{Th_T}\sum_{j=1}^{n}Y_j-\frac{\epsilon_S}{Sh_S}\sum_{j=1}^{n'}Y_j'\right|= &\quad\sup_{|x|>T^{2\gamma}}\left|\tfrac{\epsilon_T}{Th_T}\textstyle\sum_{j=1}^{n}Z_j-\tfrac{\epsilon_S}{Sh_S}\sum_{j=1}^{n'}Z_j'\right|\\ &+\sup_{|x|>T^{2\gamma}}\left|\tfrac{\epsilon_T}{Th_T}\int_0^T\mathbb{E}\left[K\left(\tfrac{x-X_t}{h_T}\right)|\mathcal{F}_{t-\delta}\right]dt\right.\\ &\left.-\tfrac{\epsilon_S}{Sh_S}\int_0^S\mathbb{E}\left[K\left(\tfrac{x-X_s}{h_S}\right)|\mathcal{F}_{s-\delta}\right]ds\right|\end{aligned}$$

$$\begin{aligned} &= \sup_{|x|>T^{2\gamma}} \mid \epsilon_T \hat{f}_T(x) - \epsilon_S \hat{f}_S(x) \mid \\ &\quad + \sup_{|x|>T^{2\gamma}} \mid \epsilon_T \bar{f}_T(x) - \epsilon_S \bar{f}_S(x) \mid . \end{aligned} \tag{2.44}$$

Notons aussi que

$$\begin{aligned} \log\Big(\epsilon_T \hat{f}_T(x)\Big) &= \log(\epsilon_T) - \log(Th_T) + \log I_T \\ &= \frac{1}{2}\log(Th_T) + \frac{1}{2}\log(\log T) - \log(Th_T) + \log I_T \\ &= \frac{1}{2}\log(\log T) - \frac{1}{2}\log(Th_T) + \log I_T, \end{aligned}$$

où $I_T = \int_0^T K\big(\frac{x-X_t}{h_T}\big)dt$. La différenciabilité par rapport à T, nous obtenons

$$\begin{aligned} \frac{d\log \epsilon_T \hat{f}_T(x)}{dT} &= \frac{1}{2T\log T} - \frac{h_T - Th'_T}{2Th_T} - \frac{h'_T}{h_T^2} I_T^{-1} \\ &\quad \int_0^T (x-X_t)K'\left(\frac{x-X_t}{h_T}\right)dt + I_T^{-1}K\left(\frac{x-X_T}{h_T}\right). \end{aligned}$$

Par l'hypothès (A.6), pour $c_K > 0$, $K'(u) = 0$ pour certains $|u| \geq c_K$, nous avons

$$\begin{aligned} \left|\frac{d\log \epsilon_T \hat{f}_T(x)}{dT}\right| &\leq \frac{1}{2T\log T} + \frac{1}{2T} + \frac{|h'_T|}{2h_T} + |I_T^{-1}|K\left(\frac{x-X_T}{h_T}\right) \\ &\quad + 2c_k \frac{|h'_T|}{h_T^2} |I_T^{-1}| Th_T \|K'\|_\infty. \end{aligned}$$

Comme $\frac{d\epsilon_T \hat{f}_T(x)}{dT} = \epsilon_T \hat{f}_T(x) \frac{d\log \epsilon_T \hat{f}_T(x)}{dT}$, on voit facilement que

$$\left|\frac{d\log \epsilon_T \hat{f}_T(x)}{dT}\right| \leq \frac{\|K\|_\infty}{(Th_T\log T)^{\frac{1}{2}}}\left(\frac{1}{2\log T} + \frac{3}{2}\right) + \frac{|h'_T|(Th_T)^{\frac{1}{2}}}{h_T^2(\log T)^{\frac{1}{2}}}\left(\frac{\|K\|_\infty}{2} + 2c_k\|K'\|_\infty\right).$$

Par conséquent, nous avons

$$\begin{aligned} &\sup_{|x|>T^{2\gamma}} |\epsilon_T \hat{f}_T(x) - \epsilon_S \hat{f}_S(x)| \leq \\ &\left(\frac{\|K\|_\infty}{(Th_T\log T)^{\frac{1}{2}}}\left(\frac{1}{2\log T} + \frac{3}{2}\right) + \frac{|h'_T|(Th_T)^{\frac{1}{2}}}{h_T^2(\log T)^{\frac{1}{2}}}\left(\frac{\|K\|_\infty}{2} + 2c_K\|K'\|_\infty\right)\right)|T-S|. \end{aligned} \tag{2.45}$$

Considérons maintenant la deuxième partie de (2.44), observons d'abord que

$$\begin{aligned}\log\left(\epsilon_T \bar{f}_T(x)\right) &= \log\left(\frac{\epsilon_T}{Th_T}\int_0^T \mathbb{E}\Big[K\Big(\frac{x-X_t}{h_T}\Big)|\mathcal{F}_{t-\delta}\Big]dt\right)\\ &= -\frac{1}{2}\log(\log T)-\frac{1}{2}\log(Th_T)+\log L_T,\end{aligned}$$

où $L_T=\int_0^T \mathbb{E}\big[K\big(\frac{x-X_t}{h_T}\big)|\mathcal{F}_{t-\delta}\big]dt$. Par conséquent, nous avons

$$\begin{aligned}\frac{d}{dT}\log\left(\epsilon_T \bar{f}_T(x)\right) &= \frac{d}{dT}\big(-\frac{1}{2}\log(\log T)-\frac{1}{2}\log(Th_T)+\log L_T\big)\\ &= -\frac{1}{2T\log T}-\frac{1}{2T}-\frac{h_T'}{2h_T}\\ &\quad -\frac{h_T'}{h_T^2}L_T^{-1}\int_0^T \mathbb{E}\Big[(x-X_t)K'\Big(\frac{x-X_t}{h_T}\Big)|\mathcal{F}_{t-\delta}\Big]dt\\ &\quad +L_T^{-1}\mathbb{E}\Big[K\Big(\frac{x-X_T}{h_T}\Big)|\mathcal{F}_{T-\delta}\Big].\end{aligned}$$

Ainsi, de même que ci-dessus, nous obtenons

$$\begin{aligned}\frac{d}{dT}\left(\epsilon_T \bar{f}_T(x)\right) &= \left(\epsilon_T \bar{f}_T(x)\right)\frac{d}{dT}\log\left(\epsilon_T \bar{f}_T(x)\right)\\ &= \epsilon_T \bar{f}_T(x)\left(-\frac{1}{2T\log T}-\frac{1}{2T}-\frac{h_T'}{2h_T}\right)\\ &\quad -\frac{h_T'}{h_T^2}\frac{\epsilon_T}{Th_T}\int_0^T \mathbb{E}\Big[(x-X_t)K'\Big(\frac{x-X_t}{h_T}\Big)|\mathcal{F}_{t-\delta}\Big]dt\\ &\quad +\frac{\epsilon_T}{Th_T}\mathbb{E}\Big[K\Big(\frac{x-X_T}{h_T}\Big)|\mathcal{F}_{T-\delta}\Big],\end{aligned}$$

et

$$\begin{aligned}\frac{d(\epsilon \bar{f}_T(x))}{dT} &\le \frac{\|K\|_\infty}{(Th_T\log T)^{1/2}}\left(\frac{1}{2\log T}+\frac{3}{2}\right)\\ &\quad +\frac{|h_T'|}{h_T^2}\frac{(Th_T)^{1/2}}{(\log T)^{1/2}}\left(\frac{1}{2}\|K\|_\infty+2c_K\|K'\|_\infty\right).\end{aligned}$$

Par conséquent, nous avons

$$\begin{aligned}&\sup_{|x|>T^{2\gamma}}|\epsilon_T\bar{f}_T(x)-\epsilon_S\bar{f}_S(x)|\le\\ &\left(\frac{\|K\|_\infty}{(Th_T\log T)^{\frac{1}{2}}}\left(\frac{1}{2\log T}+\frac{3}{2}\right)+\frac{|h_T'|(Th_T)^{\frac{1}{2}}}{h_T^2(\log T)^{\frac{1}{2}}}\left(\frac{\|K\|_\infty}{2}+\|K'\|_\infty\right)\right)|T-S|,\end{aligned}\tag{2.46}$$

L'inégalité (2.43) découle immédiatement de l'hypothèse (A.5) et les inégalités (2.44), (2.45) et (2.46). pour le choix particulier du lissage paramètre $h_T = T^{-1/5}$, on a

$$\max\left\{\left|\frac{d(\epsilon_T \hat{f}_T(x))}{dT}\right|, \left|\frac{d(\epsilon_T \bar{f}_T(x))}{dT}\right|\right\} \leq CT^{-\frac{2}{5}}\big(\log T\big)^{-1/2},$$

où C est une constante positive.
Une autre utilisation du Lemme 1.5.1 nous permet d'écrire

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Y_i\right| > Th_T\epsilon_T\epsilon_0\right) \;\; \leq \;\; 2\exp\left\{-\epsilon_0^2 \log T\right\} = 2T^{-\epsilon_0^2}.$$

Par conséquent, l'application

$$T \to \sup_{|x|>T^{2\gamma}}\left|\frac{\epsilon_T}{Th_T}\sum_{j=1}^{n} Y_j\right|,$$

est uniformément continue pour chaque ω et le résultat (2.42) s'ensuit en faisant usage de du Lemme de Borel-Cantelli pour les processus en temps continu (voir, par exemple, Bosq (1998), Lemme 4.2).
Preuve du Théorème 2.3.2 : la preuve est achevée en combinant les inégalités (2.40), (2.41) et (2.42). ■

Chapitre 3

Normalité asymptotique d'estimateurs à noyau de la densité d'un processus stationnaire à temps continu

La propriété asymptotique abordée dans ce chapitre est la normalité asymptotique, il s'agit d'un sujet très important en statistique. En effet, la normalité asymptotique nous permet de construire les intervalles de confiance et de faire des tests. Ce chapitre est divisé en deux sections, nous regroupons dans la première section l'ensemble des hypothèses utilisées pour établir La normalité asymptotique. Dans la deuxième section, nous énonçons le résultat principal de ce chapitre et la démonstration .

3.1 Hypothèses

Pour un nombre réel positif δ tel que $n = \frac{T}{\delta} \in \mathbb{N}$ considérons la δ-partition $(T_j)_{1\leq j\leq n}$ de l'intervalle $[0, T]$. De plus, pour $t > 0$ et $\delta > 0$, considérons les σ-algèbre $\mathcal{G}_t = \sigma((X_s) : 0 \leq s \leq t), \mathcal{F}_t = \sigma((X_s, Y_s) : 0 \leq s < t)$ et $\mathcal{S}_{t,\delta} = \sigma((X_s, Y_s); (X_r) : 0 \leq s < t, t \leq r \leq t + \delta)$. Chaque fois que $s < 0$, $\mathcal{G}_s$ et $\mathcal{F}_s$ se tenir comme σ-algèbre triviale. Pour $j \in \mathbb{N}$, désigner $\mathcal{G}_j = \mathcal{G}_{T_j}$ et $\mathcal{F}_j = \mathcal{F}_{T_j}$. Pour une variable aléatoire réelle ξ et tout $k \in \mathbb{N}$, nous définissons la projection $\mathcal{P}_k$ par $\mathcal{P}_k\xi = \mathbb{E}[\xi|\mathcal{G}_k] + \mathbb{E}[\xi|\mathcal{G}_{k-1}]$, où $\mathbb{E}[\xi|\mathcal{G}_k]$ est l'espérance conditionnelle de ξ étant donné le σ-algèbre $\mathcal{G}_k$.

Fixez maintenant quelques hypothèses nécessaires pour énonçer nos résultats :

(H.0) K est un noyau borné à symétrie sphérique sur son support compact, tel que $\int K(z)dz = 1$, et satisfaisant $\int z_1^{\alpha_1}...z_d^{\alpha_d}K(z)dz = 0$, où $\alpha_1...\alpha_d \in \mathbb{N}$; $\alpha_1 + ... + \alpha_d = j$, $j = 1, 2$. et $\int \|z\|^2 K(z)dz < \infty$.

(H.1) Pour tous $0 \leq s \ < t \ \leq T$, il existe une constante positive c_0, et une fonction aléatoire bornée continue et positive $f_{t,s}$ définie sur $\mathbb{R}^d$ telle que, presque sûrement,

$$\mathbb{P}^{\mathcal{F}_s}(X_t \in S_{r,x}) = \mathbb{P}(\| X_t - x \| \leq r|\mathcal{F}_s) = c_0 r^d f_{t,s}(x) + o(r^d),\ p.s\ r \to 0.$$

(H.2) **(i)** La fonction f est deux fois différentiable, avec une deuxième dérivée bornée.

(ii) La fonction $f_{t,s}$ est deux fois différentiable avec presque sûrement une dérivée seconde.

(ii) Pour tout r et tout $(s,t) \in [0,T]^2$ tels que $r \ < s$ et $r \ < t$, la densité conditionnelle du couple (X_s, X_t) sachant le σ-algèbre $\mathcal{G}_r$ existe.

(H.3) Pour tout δ, nous avons

$$\sup_y \int_{\mathbb{R}^+} \|\mathcal{P}_1 f_{t,t-\delta}(Y)\|^2 dt < \infty.$$

Commentaires sur les hypothèses

Hypothèse (H.0) sont très standard dans le cadre non paramétrique. Hypothèse (H.1) supposons que la probabilité conditionnelle de la sphère $S_{r,x}$ sachant le σ-algèbre $\mathcal{F}_s$ est asymptotiquement régie par une dimension locale lorsque le rayon r tend vers zéro sans supposer l'existence de densités marginales et conditionnelles pour atteindre les taux de convergence donnés ci-dessous. Remarque que l'hypothèse (H.1) est vraie alors que la distribution conditionnelle $\mathbb{P}_{X_t}^{\mathcal{F}_s}$ a une fonction de densité continue $f_{X_t}^{\mathcal{F}_s}$ dans $\mathbb{R}^d$.

3.2 Propriétes asymptotiques

Théorème 3.2.1. *Supposons que les hypothèses (H.0) et (H.2)-(H.3) soient vérifiées, et*

$$\lim_{T\to\infty} h_T^2 (Th_T^d)^{\frac{1}{2}} = 0, \tag{3.1}$$

alors, pour tout $x \in \mathbb{R}^d$, *on a*

$$(Th_T^d)^{\frac{1}{2}}(\hat{f}_T(x) - f(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)) \quad quand \quad T \to \infty.$$

où $\xrightarrow{\mathcal{D}}$ *est la convergence en distribution et* $\mathcal{N}$ *est la distribution gaussienne.*

Preuve du Théorème 3.2.1

Pour prouver notre résultat, introduisons d'abord la notation suivante

$$\bar{f}_{T,1}(\mathbf{x}) = \frac{1}{Th_T^d}\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i} \mathbb{E}\left[K\left(\frac{\mathbf{x}-X_t}{h_T}\right)|\mathcal{G}_{i-2}\right]dt.$$

Considérons maintenant la décomposition suivante

$$\begin{aligned}\sqrt{Th_T^d}\big(\hat{f}_T(\mathbf{x}) - f(\mathbf{x})\big) &= \sqrt{Th_T^d}\big(\hat{f}_T(\mathbf{x}) - \bar{f}_{T,1}(\mathbf{x})\big) + \sqrt{Th_T^d}\big(\bar{f}_T(\mathbf{x}) - f(\mathbf{x})\big)\\ &= (Th_T^d)^{1/2}\big(V_T(\mathbf{x}) + B_T(\mathbf{x})\big).\end{aligned}$$

Notre première préoccupation est d'étudier le comportement du biais conditionnel normalisé $(Th_T^d)^{1/2}B_T(\mathbf{x})$. Pour tout $\mathbf{x} \in \mathbb{R}^d$, nous avons

$$\begin{aligned}(Th_T^d)^{1/2}B_T(\mathbf{x}) &= (Th_T^d)^{1/2}\big(\bar{f}_{T,1}(\mathbf{x}) - \mathbb{E}[\hat{f}_T(\mathbf{x})]\\ &\quad +\mathbb{E}[\hat{f}_T(\mathbf{x})] - f(\mathbf{x})\big)\\ &= (Th_T^d)^{1/2}\big(B_T^{(1)}(\mathbf{x}) + B_T^{(2)}(\mathbf{x})\big).\end{aligned}$$

Ci-après, nous utilisons la notation $K_h(.) = \frac{1}{h}K(\frac{\cdot}{h})$. On observe que

$$\begin{aligned}|TB_T^{(1)}(\mathbf{x})| &= \left|\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i}\big(\mathbb{E}\big[K_h\big(\mathbf{x}-X_t\big)|\mathcal{G}_{i-2}\big] - \mathbb{E}\big[K_h\big(\mathbf{x}-X_t\big)\big]\big)dt\right|\\ &= \left|\int_{\mathbb{R}^d} K_h\big(\mathbf{x}-\mathbf{y}\big)\left(\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i} f_{t,T_{i-2}}(\mathbf{y})dt - Tf(\mathbf{y})\right)d\mathbf{y}\right|\\ &= \left|\int_{\mathbb{R}^d} K_h\big(\mathbf{x}-\mathbf{y}\big)H_T^{(1)}(\mathbf{y})d\mathbf{y}\right| \leq \int_{\mathbb{R}^d} K_h\big(\mathbf{x}-\mathbf{y}\big)\|H_T^{(1)}(\mathbf{y})\|d\mathbf{y}.\end{aligned}$$

En utilisant la partie inférieure de la somme de Riemann, pour $\delta > 0$ assez petit, il s'ensuit que

$$\sum_{i=1}^{n}\int_{T_{i-1}}^{Ti} f_{t,T_{i-2}}(\mathbf{y})dt \simeq \sum_{i=1}^{n}\delta f_{T_{i-1},T_{i-2}} \leq \int_0^T f_{t,t-\delta}(\mathbf{y})dt.$$

Considérons la projection $\mathcal{P}_k$ définie ci-dessus, suivant Wu (2003) et l'utilisation de l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned}\|H_T^{(1)}(\mathbf{y})\|^2 &= \left\|\sum_{i=1}^{n}\int_{T_{i-1}}^{Ti} f_{t,T_{i-2}}(\mathbf{y})dt - Tf(\mathbf{y})\right\|^2 \\ &\leq \left\|\int_0^T f_{t,t-\delta}(\mathbf{y})dt - Tf(\mathbf{y})\right\|^2 = T\bigg(\int_0^T \|\mathcal{P}_1 f_{t,t-\delta}(\mathbf{y})\|^2 dt\bigg).\end{aligned}$$

Ainsi,

$$\sup_{\mathbf{y}}\|H_T^{(1)}(\mathbf{y})\|^2 \leq T\bigg(\sup_{\mathbf{y}}\int_0^T \|\mathcal{P}_1 f_{t,t-\delta}(\mathbf{y})\|^2 dt\bigg) \leq T\left(\sup_{\mathbf{y}}\int_{\mathbb{R}^+} \|\mathcal{P}_1 f_{t,t-\delta}(\mathbf{y})\|^2 dt\right).$$

Par la suite, en supposant la condition (H.2), nous obtenons

$$\sup_{\mathbf{y}}\|H_T^{(1)}(\mathbf{y})\|^2 = \mathrm{O}(T), \tag{3.2}$$

ce qui implique que

$$B_T^{(1)}(\mathbf{x}) = \mathrm{O}(T^{-1/2}). \tag{3.3}$$

De plus, en supposant les conditions (H.0) (i) et (H.2) (i), on voit facilement que

$$\begin{aligned}B_T^{(2)}(\mathbf{x}) &= \mathbb{E}[\hat{f}_T(\mathbf{x})] - f(\mathbf{x}) \\ &= \int_{\mathbb{R}^d} K_h(\mathbf{x}-\mathbf{y})f(\mathbf{y})d\mathbf{y} - f(\mathbf{x}) \\ &= \int_{\mathbb{R}^d} K(\mathbf{z})\left(f(\mathbf{x}) - h_T\sum_{i=1}^{d} z_i\frac{\partial f}{\partial x_i}(\mathbf{x}^*) + \frac{h_T^2}{2}\sum_{j=1}^{d}\sum_{k=1}^{d} z_j z_k \frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}^*)\right) d\mathbf{z} \\ &\quad - f(\mathbf{x}) \\ &= \sup_{x_j,x_k}\left|\frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}^*)\right| \frac{h_T^2}{2}\int_{\mathbb{R}^d}\|\mathbf{z}\|^2 K(\mathbf{z})d\mathbf{z} \\ &= \mathrm{O}(h_T^2),\end{aligned}$$

où $\mathbf{x}^*$ varie entre $\mathbf{x}$ et $\mathbf{x} - \mathbf{z}h_T$. Ainsi,

$$B_T(\mathbf{x}) = \mathrm{O}(h_T^2) + \mathrm{O}(T^{-1/2}).$$

De toute évidence, nous avons

$$B_T(\mathbf{x}) = \mathrm{O}(h_T^2).$$

Considérons la condition (3.1), nous concluons que

$$(Th_T^d)^{1/2} B_T(\mathbf{x}) = \mathrm{O}(h_T^2 (Th_T^d)^{1/2}).$$

Par conséquent, pour énonçer Le Théorème 3.2.1, il suffit alors de prouver l'assertion suivante

$$(Th_T^d)^{1/2} V_T(\mathbf{x}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\mathbf{x})).$$

À cet égard, observons que

$$\begin{aligned}
(Th_T^d)^{1/2} V_T(\mathbf{x}) &= \sqrt{Th_T^d}\big(\hat{f}_T(\mathbf{x}) - \bar{f}_{T,1}(\mathbf{x})\big) \\
&= (Th_T^d)^{-1/2} \sum_{i=1}^{n} \Bigg(\int_{T_{i-1}}^{T_i} K\left(\frac{\mathbf{x} - X_t}{h_T}\right) dt \\
&\quad - \mathbb{E}\left[\int_{T_{i-1}}^{T_i} K\left(\frac{\mathbf{x} - X_t}{h_T}\right) dt | \mathcal{G}_{i-2} \right] \Bigg) \\
&= \sum_{i=1}^{n} \big(\eta_{T,i}(\mathbf{x}) - \mathbb{E}\big[\eta_{T,i}(\mathbf{x}) | \mathcal{G}_{i-2}\big] \big) \\
&= \sum_{i=1}^{n} \xi_{T,i}(\mathbf{x}),
\end{aligned}$$

où $\eta_{T,i}(\mathbf{x}) = (Th_T^d)^{-1/2} \int_{T_{i-1}}^{T_i} K(\frac{\mathbf{x}-X_t}{h_T}) dt$ et $\xi_{T,i}(\mathbf{x}) = \eta_{T,i}(\mathbf{x}) - \mathbb{E}[\eta_{T,i}(\mathbf{x}) | \mathcal{G}_{i-2}]$. De plus, on voit facilement que $(\xi_{T,i}(\mathbf{x}))_{1 \leq i \leq n}$ est une suite de différence martingale par rapport à la suite de σ-algèbre $(\mathcal{G}_{i-2})_{1 \leq i \leq n}$. Par conséquent, nous devons vérifiez les deux conditions suivantes

(a) $\sum_{i=1}^{n} \mathbb{E}[\xi_{T,i}^2(\mathbf{x}) | \mathcal{G}_{i-2}] \xrightarrow{\mathbb{P}} \sigma^2(x)$

(b) $n\mathbb{E}[\xi_{T,i}^2(\mathbf{x}) \mathbb{1}_{\{|\xi_{T,i}(\mathbf{x})| > \epsilon\}}] = o(1)$ détient, pour tout $\epsilon > 0$,

qui sont nécessaires pour établir la normalité asymptotique liée à temps discret les suite de différence de martingale (voir, par exemple, Hall et Heyde (1980)).
D'abord, par un changement de variable et le développement de Taylor de d'ordre 1 de $f_{t,T_{i-2}}(\mathbf{x}-h_T\mathbf{z})$ autour de $\mathbf{x}$ lorsque l'hypothèse (H.2) (ii) est vérifiée, observons que

$$\left|\sum_{i=1}^{n}\mathbb{E}[\eta_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}]-\sum_{i=1}^{n}\mathbb{E}[\xi_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}]\right|=\left|\sum_{i=1}^{n}(\mathbb{E}[\eta_{T,i}(\mathbf{x})|\mathcal{G}_{i-2}])^2\right|$$

$$\begin{aligned}
&= \left|\frac{1}{Th_T^d}\sum_{i=1}^{n}\left(\int_{T_{i-1}}^{T_i}\int_{\mathbb{R}^d}K\left(\frac{\mathbf{x}-\mathbf{y}}{h_T}\right)f_{t,T_{i-2}}(\mathbf{y})d\mathbf{y}dt\right)^2\right|\\
&= h_T^d\left|\frac{1}{T}\sum_{i=1}^{n}\left(\int_{T_{i-1}}^{T_i}\int_{\mathbb{R}^d}K(\mathbf{z})f_{t,T_{i-2}}(\mathbf{x}-h_T\mathbf{z})d\mathbf{z}dt\right)^2\right|\\
&= h_T^d\left|\frac{1}{T}\sum_{i=1}^{n}\left(\int_{T_{i-1}}^{T_i}f_{t,T_{i-2}}(\mathbf{x})dt\right)^2+\mathrm{O}(h_T)\right|\\
&= h_T^d\left|\frac{1}{\delta n}\sum_{i=1}^{n}g_i^{\mathcal{G}_{i-2}}(\mathbf{x})+\mathrm{O}(h_T)\right|\\
&= \mathrm{O}(h_T^d).
\end{aligned}$$

Où $g_i^{\mathcal{G}_{i-2}}(\mathbf{x})=(\int_{T_{i-1}}^{T_i}f_{t,T_{i-2}}(\mathbf{x})dt)^2$. En utilisant la somme de Riemann, la quantité $g_i^{\mathcal{G}_{i-2}}(x)$ peut être approchée, chaque fois que δ est suffisamment petit, par $\delta f_{T_{i-1},T_{i-2}}$. C'est alors clair de la discussion ci-dessus que le processus $(f_{T_{i-1},T_{i-2}})_{i\geq 1}$ est stationnaire et ergodique. Donc la somme $\frac{1}{n}\sum_{i=1}^{n}g_i^{\mathcal{G}_{i-2}}(\mathbf{x})$ a une limite finie, (voir Krengel (1985), Théorème 4.4), qui est

$$\mathbb{E}[g_1^{\mathcal{G}_{-\delta}}(\mathbf{x})]=\left(\int_0^{\delta}f(\mathbf{x})dt\right)^2=\delta^2f^2(\mathbf{x}). \tag{3.4}$$

En utilisant l'inégalité de Jensen, nous obtenons

$$\begin{aligned}
\mathbb{E}[\eta_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}] &= \frac{1}{Th_T^d}\mathbb{E}\left[\left(\int_{T_{i-1}}^{T_i}K\left(\frac{\mathbf{x}-X_t}{h_T}\right)dt\right)^2|\mathcal{G}_{i-2}\right]\\
&\leq \frac{1}{Th_T^d}\int_{T_{i-1}}^{T_i}\mathbb{E}\left[K^2\left(\frac{\mathbf{x}-X_t}{h_T}\right)|\mathcal{G}_{i-2}\right]dt.
\end{aligned}$$

Par un développement de Taylor de d'ordre 1 de la fonction $f_{t,T_{i-2}}$, pour $\mathbf{x}^*$ dans $[\mathbf{x}-h_T\mathbf{v},\mathbf{x}]$, nous obtenons

$$\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}[\eta_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}] &\leq \frac{1}{Th_T^d}\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i}\mathbb{E}\Big[K^2\Big(\frac{\mathbf{x}-X_t}{h_T}\Big)|\mathcal{G}_{i-2}\Big]dt\\
&= \frac{1}{Th_T^d}\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i}\int_{\mathbb{R}^d}K^2\Big(\frac{\mathbf{x}-\mathbf{u}}{h_T}\Big)f_{t,T_{i-2}}(\mathbf{u})d\mathbf{u}dt\\
&= \frac{1}{T}\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i}\int_{\mathbb{R}^d}K^2(\mathbf{v})f_{t,T_{i-2}}(\mathbf{x}-h_T\mathbf{v})d\mathbf{v}dt\\
&= \frac{1}{\delta}\Big(\frac{1}{n}\sum_{i=1}^{n}\int_{T_{i-1}}^{T_i}f_{t,T_{i-2}}(\mathbf{x})+\mathrm{O}(h_T)\Big)\int_{\mathbb{R}^d}K^2(\mathbf{v})d\mathbf{v}dt.
\end{aligned}$$

Procédons de même que pour l'equation (3.4), il est clair, à chaque fois que δ est assez petit, que les quantités $(\int_{T_{i-1}}^{T_i}f_{t,T_{i-2}}(\mathbf{x})dt)_{i\in\mathbb{N}}$ peuvent être approchées par $(\delta f_{T_{i-1},T_{i-2}}(\mathbf{x}))_{i\in\mathbb{N}}$. En conséquence, l'utilisation de la théorie ergodique et des propriétés de stationnarité du processus $(X_t)_{t\geq 0}$, il s'ensuit que

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\Big(\int_{T_{i-1}}^{T_i}f_{t,T_{i-2}}(\mathbf{x})dt\Big) &= \mathbb{E}\Big(\int_{T_0}^{T_1}f_t(\mathbf{x})dt\Big)+o(1)=\int_0^{\delta}\mathbb{E}(f_t(\mathbf{x}))dt+o(1)\\
&= \delta f(\mathbf{x})+o(1).
\end{aligned}$$

Il s'ensuit que

$$\sum_{i=1}^{n}\mathbb{E}[\eta_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}]\leq f(\mathbf{x})\int_{\mathbb{R}^d}K^2(\mathbf{y})d\mathbf{y}+\mathrm{O}(h_T),$$

ce qui donne

$$\sum_{i=1}^{n}\mathbb{E}[\xi_{T,i}^2(\mathbf{x})|\mathcal{G}_{i-2}]\xrightarrow{\mathcal{D}}\sigma^2(\mathbf{x})\leq f(\mathbf{x})\int_{\mathbb{R}^d}K^2(\mathbf{y})d\mathbf{y}\quad\text{quand }\ T\to\infty. \tag{3.5}$$

Preuve de la partie (b). Utilisons successivement les inégalités de Holder, Markov, Jensen puis Minkowski, nous obtenons pour tous $\epsilon>0$ et tout p et

q tels que $1/p + 1/q = 1$

$$\begin{aligned}
\mathbb{E}[\xi_{T,i}^2(\mathbf{x})\mathbb{1}_{\{|\xi_{T,i}(\mathbf{x})|>\epsilon\}}] &\leq (\mathbb{E}[\xi_{T,i}^{2q}(\mathbf{x})])^{1/q}(\mathbb{P}\{|\xi_{T,i}(\mathbf{x})| > \epsilon\})^{1/p} \leq \epsilon^{-2q/p}\mathbb{E}[|\xi_{T,i}(\mathbf{x})|^{2q}] \\
&\leq \frac{\epsilon^{-2q/p}}{(Th_T^d)^q}\mathbb{E}\left[\int_{T_{i-1}}^{T_i}\left|K\left(\frac{\mathbf{x}-X_t}{h_T}\right) - \mathbb{E}\left[K\left(\frac{\mathbf{x}-X_t}{h_T}\right)|\mathcal{G}_{t-\delta}\right]\right|^{2q} dt\right] \\
&\leq \frac{\epsilon^{-2q/p}}{(Th_T^d)^q}\int_{T_{i-1}}^{T_i}\mathbb{E}\left[K^{2q}\left(\frac{\mathbf{x}-X_t}{h_T}\right)\right]dt \\
&= \frac{\epsilon^{-2q/p}}{(Th_T^d)^q}\int_{T_{i-1}}^{T_i}\int_{\mathbb{R}}K^{2q}\left(\frac{\mathbf{x}-\mathbf{y}}{h_T}\right)f(\mathbf{y})d\mathbf{y}dt \\
&= \frac{\epsilon^{-2q/p}\delta}{T^q h_T^{d(q-1)}}\int_{\mathbb{R}}K^{2q}(\mathbf{z})f(\mathbf{x}-h_T\mathbf{z})d\mathbf{z}.
\end{aligned}$$

Par conséquent, puisque f est une densité, il s'ensuit que

$$n\mathbb{E}[\xi_{T,i}^2(\mathbf{x})\mathbb{1}_{\{|\xi_{T,i}(\mathbf{x})|>\epsilon\}}] = \frac{\epsilon^{-2q/p}\|K\|_\infty^{2q}}{(Th_T^d)^{q-1}} = o(1). \tag{3.6}$$

En combinant les equations (3.5) et (3.6), on obtient

$$(Th_T^d)^{1/2}V_T(\mathbf{x}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\mathbf{x})), \tag{3.7}$$

ce qui complète la preuve. ■

Conclusion

Dans ce mémoire, nous avons étudié les propriétés asymptotiques de l'estimateur à noyau de la densité d'un processus stationnaire à temps continu. La motivation essentielle est d'établir ces propriétés tout en considérant un cadre de dépendance des données assez général qui puisse être facilement utilisé en pratique. Nous avons obtenu nos résultats, en faisant usage d'une approche basée sur l'utilisation des différences de martingales en liaison avec des techniques de projection sur des tribus. Nous avons attaché une importance primordiale à l'étude de la vitesse de convergence presque sûre, ponctuelle et uniforme du " terme stochastique " qui a été obtenue par les estimateurs centrés (par des facteurs n'étant pas nécessairement égaux à l'espérance de l'estimateur). En effet, il peut être montré que, pour les tailles de la fenêtre h les plus appropriées aux applications statistiques, et sous des conditions générales de régularité, le " terme de biais " (i.e., le facteur de centrage moins la fonction à estimer) converge avec une vitesse vers zéro. Les résultats qui ont été obtenus sur l'estimation d'une densité de probabilité à noyau continu peuvent être généraliser au cas des densités multidimensionnelles.

Bibliographie

[1] Absava, R. M. (1999). About choice of the window width in the kernel non parametric estimate of probability density. Appl. Math. Inform. 4, p. 13-28, 93.

[2] Andrews, D.W.K. (1984). Non-strong mixing autoregressive processes. J. Appl. Probab., 21, 930-934.

[3] Ash, R.B. and Gardner, M.F. (1975). Topics in stochastic processes, Academic Press.

[4] Arnold, L. (1973) Stochastic Differential Equations : Theory and Practice. New York : Wiley.

[5] Banks, H.T. (1975) Modeling and Control in the Biological Sciences. Lect. Notes Biomath, 6, Berlin : Springer-Verlag.

[6] Banon, G. (1978). Non parametric identification for diffusion processes. SIAM J. Control Optim. 16(3), p. 380-395.

[7] Banon, G. Nguyen, H.T. (1978). Sur l'estimation récurrente de la densité et de sa dérivée pour un processus de Markov. C. R. Acad. Sci. Paris Sér. A Math. 286(16), p. 691-4.

[8] Banon, G. Nguyen, H.T. (1981). Recursive estimation in diffusion model.SIAM J. Control Optim. 19(5), p. 676-85.

[9] Bergstrom, A.R. (1990) Continuous Time Eonometrie Modelling. Oxford : Oxford University Press.

[10] Berlinet, A., Gannoun, A. et Matzner-Løber, E. (1998). Normalité asymptotique d'estimateurs convergent du mode conditionnel. Canad. J.Statist. 26, p. 365- 380.

[11] Beran, J. (1994).Statistics for long memory processes. Chapman and Hall, New York.

[12] Bosq, D. (1997). Parametric rates of nonparametric estimators and predictors for continuous time processes. Ann. Statist., 25, p. 982-1000.

[13] Bosq, D and Blanke, D. (1997). Accurate rates of density estimators for continuous-time processes, Statist. Probab. Lett., 33 p. 185-191.

[14] Bosq D. (1998). Nonparametric statistics for stochastic processes : Estimation and prediction. Springer New york.

[15] Bosq, D. and Merlevède, F. and Peligrad, M. (1999).Asymptotic normality for density kernel estimators in discrete and continuous time, J. Multivariate Anal., 68, p. 78-95.

[16] Bosq, D and Blanke, D. (2007). Inference and prediction in large dimensions, Wiley Series in Probability and Statistics, p. 316.

[17] Bosq, D and Blanke, D. (2008). Regression estimation and prediction in continuous time, Journal of the Japan Statistical Society (Nihon Tôkei Gakkai Kaihô), 38. p. 15-26.

[18] Bradley, R. C. (2007). Introduction to strong mixing conditions. Vol. 1 . Kendrick Press, Heber City, UT.

[19] Cacoullos, Theophilos, (1966). Estimation of a multivariate density, Ann. Inst. Statist. Math. 18, p. 179-189.

[20] Castellana, J. V. and Leadbetter, M. R. (1986). On smoothed probability density estimation for stationary processes. Stochastic Process. Appl., 21, p. 179-193.

[21] Chacón, J.E. and Tenreiro, C. (2012). Exact and asymptotically optimal bandwidths for kernel estimation of density functionals. Methodol. Comput. Appl. Probab. 3,p. 523-548.

[22] Chernick, M.R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. Ann. Probab., 9, 145-149.

[23] Collomb, G. (1984). Propriétés de convergence presque complète du prédicteur à noyau. Z. Wahrsch. Verw. Gebiete. 66, p. 441-460.

[24] Cohen, J. E. (1979). Ergodic theorems in demography. Bull. Amer. Math. Soc. (N.S.) 1. p. 275-295.

[25] Debbarh, M. and Maillot, B. (2008a). Additive regression model for continuous time processes, iComm. Statist. Theory Methods, 37. p. 2416-2432.

[26] de la Peãa, V.H. and Giné, E. (1999). Decoupling, from dependence to independence. Probability and its applications. Springer-Verlag, New York.

[27] Delecroix, M. (1979).Sur l'estimation des densités d'un processus stationnaire et mélangeant . Publ. U.E.R. Math. Pures Appl. IRMA, 1(4), exp. no. I, 24. Seminar on Mathematical Statistics.

[28] Delecroix, M. (1980). Sur l'estimation des densités d'un processus stationaire à temps continu. Publ. Inst. Statist. Univ. Paris, 25, p. 17-39.

[29] Delecroix, M. and Rosa, A. C. (1996). Nonparametric estimation of a regression function and its derivatives under an ergodic hypothesis. J. Nonparametr. Statist., 6, p. 367-382.

[30] Delecroix, M. (1987). Sur l'estimation et la prévision nonparamétrique des processus ergodiques. Doctorat d'état, Université des sciences et techniques de Lille Flandre-Artois.

[31] Delecroix, M. Nogueira, M.E. Rosa, A. C. (1991). Sur l'estimation de la densité d'observations ergodiques. Statist. Anal. Données., 16, p. 25-38.

[32] Devroye, L., Györfi, L. et Lugossi, G. (1986). A probability Theory of Pattern Recognition. Springer New-York.

[33] Didi, S. Louani, D. (2013a) Consistency results for the kernel density estimate on continuous time stationary and dependent data. Journal of Statistics et Probability Letters, 83, 4, p. 1262-1270.

[34] Doukhan, P. Louhichi, S. (2001). Functional estimation of a density under a new weak dependence condition. Scand. J. Statist. 28(2), p. 325-41.

[35] Györfi,L., Härdle, W., Sarda, P., and Vieu, P. (1989). Nonparametric curve estimation from time series. Lecture Notes in Statistics, 60. Springer-Verlag, Berlin.

[36] Hale, J.K., Verduyn Lunel, S.M. (1993) lntroduction to Functional-differential Equations. New York : Springer-Verlag.

[37] Härdle, W. and Marron, J. S. (1985). Optimal bandwidth selection in nonparametric regression function estimation. Ann. Statist. 13(4). p. 1465-1481.

[38] Ibragimov, I.A., Linnik, Yu.V., (1971). Independent and Stationary Sequences of Random Variables . Wolters-Nordhoff, Groningen.

[39] Karatzas, I., Shreve, S.E. (1991) Brownian Motion and Stochastic Calculus. New York : Springer-Verlag.

[40] Krengel, U. (1985). Ergodic Theorems, Walter de Gruyter et Co. Berlin.

[41] Lange, K. (2002) Mathematieal and Statistical Methods for Genetie Analysis. New York : Springer-Verlag.

[42] Liebscher, E. (2001). Estimation of the density and the regression function under mixing conditions.Statist. Decisions. 19(1), p. 9-26.

[43] Laib, N. (1999). Exponential-type inequalities for martingale difference sequences. Application to nonparametric regression estimation. Comm. Statist. Theory Methods, 28(7), 1565-1576

[44] Laïb, N. (2005). Kernel estimates of the mean and the volatility functions in a nonlinear autoregressive model with ARCH errors. J. Statist. Plann. Inference, 134, p. 116-139.

[45] Leblanc, F. (1995), Wavelet density estimation of a continuous-time process and application to diffusion process. C.R. Acad. Sci. Paris, série I 321, p. 345-350.

[46] LU, Z. (2009). Analyse des Processus Longue Mémoire Stationnaires et Nonstationnaires : Estimations, Applications et Prévisions. PhD Thesis.

[47] Marron, J. S. (1987). A comparison of cross-validation techniques in density estimation. Ann. Statist. 15(1). p. 152-162

[48] Maslowski, B,and $Posp\tilde{s}il$, J (2008). Ergodicity and parameter estimates for infinite-dimensional fractional Ornstein-Uhlenbeck process. Appl. Math. Optim. 3, p. 401-429.

[49] Nadaraya, E.N. (1965). On nonparametric estimates of density functions and regression curves. Teor. Verojatnost. i Primenen., 10, p. 199-203.

[50] Nguyen, H.T. (1979). Density estimation in a continuous-time stationary Markov process.Ann. Statist. 7(2), p. 341-348.

[51] Nguyen, H.T. Pham, T.D. (1980). Sur l'utilisation du temps local en statistique des processus. C. R. Acad. Sci. Paris SeÂ´r. A Math. 290(3), p. 165-168.

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

[53] Papanicolaou, G. (1995) Diffusion in random media.In Surveys in Applied Mathematics, J.B. Keller, D. McLaughlin and G. Papanicolaou (Eds), Plenum Press, p. 205-255.

[54] Parzen, E. (1962). On estimation of a probability density function and mode, Ann. Math. Statist, p. 1065-1076.

[55] Peña, V. H. and Giné, E. (1999). Decoupling. Probability and its Applications (New York). Springer-Verlag, New York. From dependence to independence, Randomly stopped processes. U-statistics and processes. Martingales and beyond.

[56] Peskir, G. (1998). The uniform mean-square ergodic theorem for wide sense stationary processes, Stochastic Anal. Appl. , 16, p. 697-720.

[57] Peskir, G. (2000). From uniform laws of large numbers to uniform ergodic theorems, Lecture Notes Series (Aarhus), 66.

[58] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function, Ann. Math. Statist. 27, p. 832-837,

[59] Rosenblatt, M. (1970). Density estimates and Markov sequences. In Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969), pages 199-213. Cambridge Univ. Press, London.

[60] Sarda, P. (1991). Estimating smooth distribution functions. In Nonparametric functional estimation and related topics (Spetses, 1990), volume 335 of NATO Adv.Sci. Inst. Ser. C Math. Phys. Sci. p. 261-270. Kluwer Acad. Publ., Dordrecht.

[61] Sarda, P. and Vieu, P. (1991). Smoothing parameter selection in hazard estimation. Statist. Probab. Lett. 11(5). p. 429-434.

[62] Silverman, B.W. and Jones, MC. 1989 E. Fix and J. L. Hodges (1951). an important contribution to nonparametric discriminant analysis and density estimation. Int. Stat. Rev. 57(3) 233-247.

[63] Watson, G. S., (1964) Smooth regression analysis 26 of Sankhy$\bar{a}$ Ser. A, Sankhy$\bar{a}$ (Statistics). The Indian Journal of Statistics. Series A, p. 359-372.

[64] Wu, W.B., (2003) Nonparametric Estimation For stationary Processes. Technical Report 536 University of Chicago.

[65] Wu, W. B., Huang, Y. and Huang, Y. (2010). Kernel estimation for time series : An asymptotic theory. Stochastic Processes and their Applications, 120, p. 2412-2431

[66] Youndjé, É. Sarda, P. and Vieu, P. (1993a) Estimateur à noyau d'une densité conditionnelle : choix de la fenêtre pour des observations dépendantes. C. R. Acad. Sci. Paris SÃ c r. I Math. 316(9). p. 935-938.

[67] Youndjé, É. Sarda, P. and Vieu, P. (1994). Validation croisée pour l'estimation non-paramétrique de la densité conditionnelle. Publ. Inst. Statist. Univ. Paris. 38(1). p. 57-80,

[68] Youndjé, É. Wells, M.T. (2008). Optimal bandwidth selection for multivariate kernel deconvolution density estimation.TEST,17(1). p. 138-162.