

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediène

Faculté de Mathématiques



THÈSE

Présentée pour l'obtention du diplôme de DOCTORAT 3^{ème} Cycle

En : Mathématiques

Spécialité : Mathématiques Financières et Actuariat

Par : HAMRANI Farida

Sujet

**Estimation non paramétrique pour
les données incomplètes et associées**

Soutenue publiquement le 09/05/2017 à 13h30, devant le jury composé de :

Mr. Kamal BOUKHETALA	Professeur	USTHB, Algérie	Président
Mme. Zohra GUESSOUM	MCA	USTHB, Algérie	Directrice de Thèse
Mme. Karima BELAIDE	MCA	Univ. Bejaia, Algérie	Examinatrice
Mr. Ali LAKSACI	Professeur	Univ. S.B.A., Algérie	Examineur
Mr. Elias OULD SAÏD	Professeur	Univ. du Littoral, France	Examineur
Mr. Abdelkader TATACHAK	Professeur	USTHB, Algérie	Examineur
Mme. Ourida SADKI	Professeur	USTHB, Algérie	Invitée

*Cette thèse est dédiée à mon Père et à mon Grand Frère,
Mohamed, ceux qui rêvaient de me voir obtenir mon doctorat
et ne pouvaient pas terminer le long chemin avec moi.
Que Dieu ait vos âmes en paix.*

Remerciements

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je voudrais, en tout premier lieu, exprimer ma profonde reconnaissance à ma directrice de thèse, Madame Zohra GUESSOUM, d'avoir accepté de m'initier à la recherche et diriger ce travail. Comme tous ceux qui ont eu la chance d'être sous sa direction, j'ai pu constater à quel point elle est attachante, attentionnée et profondément gentille. Elle restera pour moi un modèle, pour ses grandes compétences scientifiques, pour son exigence de la clarté, sa franchise et ses qualités humaines. Je lui adresse mes remerciements les plus chaleureux, d'avoir accepté de me prendre sous son aile, pour tout ce qu'elle m'a appris et pour ses encouragements. J'espère dans l'avenir pouvoir transmettre tout ce qu'elle m'a appris, la meilleure façon selon moi de lui rendre hommage.

Je tiens à remercier vivement le Professeur Kamal BOUKHETALA, qui m'a fait l'honneur de présider ce jury, malgré ses nombreuses responsabilités.

Merci infiniment à Madame Karima BELAIDE, et au Professeur Ali LAKSACI, pour l'honneur qu'ils me font d'examiner mon travail, et d'avoir accepté de faire le déplacement malgré leurs multiples responsabilités.

Je suis très honoré, que le Professeur Elias OULD SAÏD, ait accepté d'examiner cette thèse bien qu'il soit extrêmement occupé. J'ai eu l'occasion de discuter avec lui beaucoup de fois. Je tiens à le remercier pour l'intérêt qu'il a porté à mon travail.

Je suis très reconnaissante au Professeur Abdelkader TATACHAK pour toute l'aide qu'il m'a apportée avec autant de gentillesse et de générosité. Je le remercie aussi pour l'honneur qu'il me fait en acceptant de participer au jury de ma soutenance.

Je remercie vivement la Professeure Ourida SADKI, pour l'intérêt qu'elle a manifesté en participant en qualité de membre invité à ce jury.

Je désire en outre remercier tous les membres du Groupe de Travail sur les Données Incomplètes (GTDI) pour leur sympathie, leur amitié. J'ai eu beaucoup de plaisir à travailler avec eux. Le cadre de travail était idéal.

Je dis un grand merci du fond du coeur à tous mes amis pour leur soutien moral.

A titre plus personnel, je remercie chaleureusement mon mari, pour la grande patience, l'encouragement et la confiance qu'il m'a témoignés. Je tiens à le remercier surtout pour son soutien moral ininterrompu et ses nombreux conseils tout le long de ma thèse. Aucun remerciement n'est trop grand pour Toi, mon fils pour lequel je dédie ce travail.

De plus, mes remerciements seraient incomplets, si je ne fais pas mention aux membres de ma famille. Qu'ils trouvent ici toute ma reconnaissance pour le soutien que chacun d'eux a pu apporter pour faire aboutir ce rêve. De même, que tous les membres de ma belle-famille qui m'ont soutenu, de près ou de loin durant mes études doctorales.

Ces remerciements ne peuvent s'achever, sans une pensée pour ma première fan : ma mère. Sa présence et ses encouragements sont pour moi les piliers fondateurs de ce que je suis et de ce que je fais.

Résumé

Les travaux de cette thèse portent sur le problème de l'estimation non paramétrique, par la méthode du noyau, de la fonction de régression lorsque la variable d'intérêt est soumise à une troncature aléatoire à gauche. La motivation essentielle est d'établir des propriétés asymptotiques de l'estimateur à noyau de la fonction de régression introduit par Ould Saïd et Lemdani (2006) tout en considérant un cadre de dépendance des données de type associé.

Cette contribution se compose de trois parties. La première partie est consacrée à l'estimation, par la méthode du noyau, de la densité et de la fonction de régression pour des données complètement observées et associées. Nous rappelons certains résultats existants et étendons ceux-ci au cas multidimensionnel. Dans la deuxième partie, nous nous intéressons à l'estimation à noyau de la fonction de régression sous le modèle aléatoire de troncature à gauche, dans les cas des données indépendantes et fortement mélangeantes (α -mélangeantes). Dans la troisième partie, nous donnons nos résultats où nous établissons la convergence uniforme presque sûre avec vitesse de l'estimateur à noyau de la fonction de régression pour des données tronquées à gauche et associées. En outre, nous montrons que l'estimateur convenablement normalisé est asymptotiquement gaussien. Nous donnons aussi des illustrations de nos résultats sur des données simulées.

Mots clés : Association, Convergence uniforme presque sûre, Données tronquées à gauche, Estimateur à noyau, Normalité asymptotique, Régression non paramétrique, Vitesse de convergence.

Abstract

The work of this thesis focuses on some nonparametric estimation problems. More precisely, considering kernel estimators of the regression function when the interest variable is subject to random left truncation. We aim at establishing some asymptotic properties of the kernel estimator introduced by Ould Saïd et Lemdani (2006) while taking a dependency framework for the data.

The present manuscript includes three parts. The first one focuses on the kernel density and regression estimation when the data are associated and completely observed. we recall some existing results and extend them to the multidimensional case. In the second part, we are interested in kernel regression estimation under the random left truncation model, in both independent and strong mixing condition (α -mixing) cases. In the third part, we give our results. We give the rate of almost sure uniform convergence of the kernel regression estimator when the data are left truncated and associated. Moreover, we show that the suitably standardized estimator is asymptotically normal. We give also illustrations of our results on simulated data.

Key words : Association, Asymptotic normality, Kernel estimator, Left truncated data, Nonparametric regression, Rate of convergence, Strong uniform consistency.

Liste des travaux

Publications dans des revues à comités de lecture

1. Guessoum, Z., and Hamrani, F. (2016), "Convergence rate of the kernel regression estimator for associated and truncated data," *Journal of Nonparametric Statistics*, DOI : 10.1080/10485252.2017.1303059
2. Guessoum, Z., and Hamrani, F. (Submitted for Publication), "Asymptotic normality of the kernel regression estimator for truncated and associated data."

Communications dans des congrès

1. Guessoum, Z., Hamrani, F., "Simulation de l'estimateur à noyau de la fonction de régression dans un modèle de troncature à gauche," *Colloque International de Modélisation Stochastique et Statistique (MSS'2014)*, 23-25 Novembre 2014, USTHB, Alger.
2. Guessoum, Z., Hamrani, F., "Régression non paramétrique pour des données tronquées à gauche et associées," *Colloque International sur la Statistique et ses Applications (CISA'2015)*, 06-07 Mai 2015, Université Dr Tahar MOULAY, Saida.
3. Guessoum, Z., Hamrani, F., "Propriétés asymptotiques d'un estimateur non paramétrique de la fonction de régression pour des données tronquées à gauche et associées," *Congrès des Mathématiciens Algériens (CMA'2016)*, 08-09 Novembre 2016, Université de Batna II, Batna.

Table des matières

1	Introduction générale	1
2	Estimation de la densité et de la fonction de régression pour des données complètes associées	8
2.1	Données associées	8
2.1.1	Propriétés d'association	9
2.1.2	Exemples de vecteurs et processus associés	10
2.1.3	Autres types de dépendance	10
2.1.3.1	Association Positive	10
2.1.3.2	Association Négative	11
2.1.3.3	Quasi-Association	11
2.1.3.4	Dépendance faible	12
2.1.4	Inégalité exponentielle	13
2.2	Estimation de la densité	14
2.2.1	Convergence uniforme presque sûre	15
2.2.2	Normalité asymptotique	24
2.3	Estimation de la fonction de régression	25
2.3.1	Hypothèses et résultat	26
2.3.2	Preuve du résultat	27
3	Estimation de la fonction de régression pour des données tronquées à gauche	29
3.1	Estimation sous le modèle aléatoire de troncature à gauche	29
3.1.1	Modèle aléatoire de troncature à gauche (RLT)	29
3.1.2	Estimation de les fonctions de répartition	30
3.1.3	Estimation de la probabilité de troncature	32
3.1.4	Estimation de la densité de la covariable	33
3.1.5	Estimation de la fonction de régression	35
3.2	Propriétés asymptotiques de l'estimateur à noyau de la fonction de régression sous le modèle RLT	37

3.2.1	Cas de suites de variables aléatoires i.i.d.	37
3.2.1.1	Hypothèses et résultats principaux	38
3.2.1.2	Preuves des résultats	39
3.2.2	Cas de suites de variables aléatoires α –mélangeantes	51
3.2.2.1	Convergence presque sûre	52
3.2.2.2	Normalité asymptotique	53
4	Vitesse de convergence de l'estimateur de la fonction de régression pour des données tronquées à gauche et associées	55
4.1	Introduction	55
4.2	Model and estimators	57
4.3	Assumptions and main results	60
4.4	Simulations study	62
4.4.1	Performances of $\hat{m}_n(\cdot)$ under linear model	63
4.4.2	Some comparisons	65
4.4.3	Nonlinear model	67
4.5	Conclusion	68
5	Normalité asymptotique de l'estimateur de la fonction de régression pour des données tronquées à gauche et associées	77
5.1	Introduction	77
5.2	Présentation du modèle et les estimateurs	79
5.3	Hypothèses et résultats principaux	81
5.4	Étude de simulations	83
5.5	Preuve du résultat principal	86
6	Conclusion et Perspectives	99
6.1	Conclusion	99
6.2	Perspectives	100

Chapitre 1

Introduction générale

L'estimation statistique est un domaine très important de la statistique mathématique qui développe des techniques pour décrire certaines caractéristiques d'ensembles d'observations. Ce domaine est divisé en deux volets principaux, l'estimation paramétrique et l'estimation non paramétrique. Dans le cadre de l'estimation non paramétrique, on s'intéresse à l'estimation à partir des observations, d'une fonction inconnue appartenant à une certaine classe de fonctions. Notons que l'approche non paramétrique est basée sur un procédé indépendant de la loi de l'échantillon d'observations tandis que l'approche paramétrique se restreint à l'estimation d'un nombre fini de paramètres liés à la loi de l'échantillon.

Un des problèmes centraux en statistique est celui de l'estimation de caractéristiques fonctionnelles associées à la loi des observations, comme par exemple, la fonction de densité ou la fonction de régression. Dans le modèle de régression non paramétrique, on suppose l'existence d'une fonction $m(\cdot)$ qui exprime la valeur moyenne d'une variable d'intérêt Y en fonction d'une covariable X . Considérons un N -échantillon $(X_1, Y_1), \dots, (X_N, Y_N)$ du couple aléatoire (X, Y) , la relation de régression est modélisée par :

$$Y_i = m(X_i) + \epsilon_i, \quad \text{pour } i = 1, \dots, N, \quad (1.1)$$

avec $\{\epsilon_i; i = 1, \dots, N\}$ est une suite d'erreurs. Dans notre contexte non paramétrique, plusieurs méthodes existent dans la littérature pour l'estimation de la fonction de régression. On peut citer les estimateurs par la méthode des moindres carrées (Vapnik et Chervonenkis (1971), Devroye et al. (1996)), par fonctions splines (Schumaker (1981), Eubank (1999)), la méthode des plus proches voisins (Stone (1977), Devroye et al. (1994)), les estimateurs par la méthode des ondelettes (Antoniadis (1994), Donoho et al. (1995)), ceux par partition (Walk (1997)) et les estimateurs par la méthode du noyau. Parmi l'ensemble de ces estimateurs, l'un des plus utilisés reste l'estimateur par la méthode du noyau introduit séparément par Nadaraya (1964) et Watson (1964). Cette méthode d'estimation a connu un développement continu. En effet, Devroye (1978) a établi la convergence uniforme presque sûre de cet estimateur. Le taux de convergence optimal pour la régression non

paramétrique a été donné par Stone (1980, 1982). Collomb (1981, 1983, 1984, 1985) apporte une contribution déterminante sur ce modèle. Ces travaux se sont focalisés sur l'utilisation de la régression dans la prévision de séries chronologiques. Nous renvoyons à Bosq et Lecoutre (1987), Schimek (2000), Sarda et Vieu (2000) pour un large éventail de références.

Les applications de modèle de régression ont pris une place importante dans différents domaines tels que l'économie, la médecine, la biologie, l'épidémiologie, l'astronomie, la fiabilité, etc... Remarquons que dans chacun de ces domaines, on s'intéresse à des variables aléatoires généralement positives représentant la durée de temps jusqu'à l'apparition d'un certain événement, appelées durée de survie. La particularité des données de survie réside dans le fait que ces données sont souvent incomplètes à cause du phénomène de censure ou/et de troncature.

Les données censurées à droite sont très fréquentes dans les modèles de survie. On parle de censure droite lorsqu'à la fin de la durée d'observation, un sujet n'a pas connu l'événement d'intérêt pour une raison ou une autre. Ainsi, la durée de survie observée pour ces sujets n'est donc pas la variable d'intérêt Y mais une autre variable de durée C qui correspond à la date à laquelle le sujet sort de l'étude. Parmi les travaux sur l'estimation de la fonction de survie dans ce modèle, en plus de l'article fondateur de Kaplan et Meier (1958) qui ont proposé un estimateur produit-limite pour la fonction de distribution de la variable d'intérêt, citons Peterson (1977) qui établit la consistance de l'estimateur de Kaplan et Meier, Stute et Wang (1993) qui ont établi un résultat du type Loi Forte des Grands Nombres et le livre de Gill (1994). Concernant l'estimation non paramétrique de la fonction de régression dans ce même modèle, Carbonez et al. (1995) introduisent un estimateur à partitions de la fonction de régression et établissent sa consistance forte. Kohler et al. (2002) exploitent l'idée de ce dernier travail pour l'étendre à diverses méthodes (à noyau, plus proches voisins, moindres carrés et spline de lissage). Guessoum et Ould Saïd (2008) établissent la convergence uniforme presque sûre sur des compacts et la normalité asymptotique de l'estimateur à noyau.

Dans cette thèse, nous nous intéressons particulièrement, à un autre cas classique de données incomplètes, celui des données dites tronquées à gauche qui est le cadre dans lequel nous avons apporté de nouveaux résultats.

La troncature à gauche se produit lorsque la durée de survie modélisée par une variable Y doit être assez grande pour être observée. Y doit en fait être plus grande qu'une variable de troncature T . Ainsi les observations ne sont possibles que si $Y \geq T$. Il s'agit d'un modèle qui est tout d'abord apparu en astronomie, où des échantillons sont composés d'objets astraux d'une certaine zone. Les luminosités absolues et apparentes d'un objet astral sont respectivement définies comme étant sa brillance observée à une distance fixe et depuis la terre et l'on n'observe que les objets qui sont suffisamment brillants, c'est-à-dire ceux pour lesquels la luminosité $M \geq m$, m étant la variable de troncature. Un autre exemple de troncature à gauche concerne la durée de vie des retraités. On enregistre l'âge au

décès et l'âge d'entrée au centre de retraite. Un individu donné, doit vivre suffisamment longtemps pour être parmi la communauté du centre de retraite et s'il décède avant, il y a troncature à gauche. Plus d'exemples et références traitant des données tronquées peuvent être trouvés dans Woodroffe (1985), Wang et al. (1986), Tsai et al. (1987), He et Yang (1994) et Chen et al. (1995).

Comme dans le modèle de troncature gauche, nous ne sommes capable d'observer que les durées de vie Y pour lesquelles $Y \geq T$, alors, nous disposons d'un échantillon observé de taille n . Cet échantillon est extrait d'un échantillon de plus grande taille N inconnue. Ainsi, il n'est pas possible d'avoir un échantillon représentant toute la population considérée, car lorsque $Y < T$ rien ne peut être observé. Les résultats asymptotiques devraient être exprimés par rapport à la probabilité \mathbb{P} , relative au N -échantillon qui n'est malheureusement pas observable. Cependant, le fait que $n \leq N$, nous permet de dire que les résultats établis quand $n \rightarrow \infty$ resteront valables quand $N \rightarrow \infty$. Il nous faut donc introduire une probabilité conditionnelle relative au n -échantillon, définie par $\mathbf{P}(\cdot) = \mathbb{P}(\cdot | Y \geq T)$. Les résultats asymptotiques sont donc établis pour $n \rightarrow \infty$ et exprimés relativement à la probabilité \mathbf{P} .

Dans le cadre de ce modèle, Lynden-Bell (1971) introduit les estimateurs produits-limites des fonctions de répartition (f.d.r.) de la variable d'intérêt Y et de la troncature T . Woodroffe (1985) établit les conditions d'identifiabilité du modèle ainsi que la convergence presque sûre des estimateurs de Lynden-Bell avec une vitesse d'ordre $n^{-1/2}$. Stute (1993) établit des représentations fortes de l'estimateur de la fonction de risque cumulatif ainsi que de l'estimateur de Lynden-Bell de la variable d'intérêt Y comme somme de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) avec un terme de reste $o_{\mathbf{P}}(n^{-1}(\log n)^{\delta})$ pour tout $\delta > 3/2$.

Une autre quantité importante inconnue qui intervient dans ce modèle et qu'on doit estimer est la probabilité de (non) troncature qui est la proportion de la population que nous observons. Cette probabilité notée $\alpha := \mathbb{P}(Y \geq T)$ pourrait être estimée par $\frac{n}{N}$ mais cet estimateur ne peut être calculé, car N est inconnue. A ce stade, remarquons que n est une variable aléatoire de loi binomiale $\mathcal{B}(N, \alpha)$. En utilisant les estimateurs de Lynden-Bell, He et Yang (1998) donnent un estimateur calculable de α , ils montrent qu'il ne dépend pas de l'argument et donnent aussi des résultats de convergence asymptotique dont la normalité.

L'estimation non paramétrique de la fonction de régression pour ce même modèle est traité dans Gross et Lai (1996), He et Yang (2003), ... Ould Saïd et Lemdani (2006) construisent un nouvel estimateur à noyau pour la fonction de régression et établissent sa consistance uniforme sur un ensemble compact ainsi que sa normalité asymptotique.

D'autres résultats asymptotiques ont été établis en estimation fonctionnelle par la méthode du noyau pour des données incomplètes (censure ou troncature). Rappelons les travaux de Ould Saïd (2006) et Ould Saïd et Sadki (2008) concernant les quantiles conditionnels dans un modèle de censure

à droite, Ould Saïd et Tatachak (2007) pour le mode conditionnel sous troncature à gauche, Lemdani et al. (2009) ont étudié la fonction des quantiles conditionnels pour des données tronquées à gauche et finalement Khardani et al. (2010) ont traité le problème de l'estimation du mode conditionnel pour des données censurées à droite.

Jusqu'à présent, les travaux cités dans ce cadre supposent l'indépendance des données. Cependant, le fait de supposer que les données étudiées sont toujours indépendantes est peu réaliste, c'est pour cela que plusieurs auteurs ont concentrés leurs études sur un autre type de données, qui sont les données dépendantes. De nombreuses notions ont été introduites pour modéliser la dépendance. Les notions de mélange et d'association sont largement utilisés dans la littérature.

Le concept de α -mélange (ou mélange fort) introduit par Rosenblatt (1956) est le plus faible et donc le moins restrictif parmi toutes les formes de mélange qui existent dans la littérature. Rappelons qu'une suite de variables aléatoires $\{X_i; i = 1, \dots, n\}$ définie sur $(\Omega, \mathcal{F}, \mathbf{P})$ est α -mélangeante (ou fortement mélangeante) si la suite des coefficients

$$\alpha(n) = \sup \{ |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|, A \in \mathcal{F}_1^k \text{ et } B \in \mathcal{F}_{k+n}^\infty, k \geq 1 \}$$

tend vers zéro quand n tend vers l'infini où \mathcal{F}_i^j désigne la tribu engendrée par les variables X_i, \dots, X_j . Pour plus de détails sur les conditions de mélange, nous renvoyons le lecteur aux références suivantes : Doukhan (1994), Bosq (1998), Rio (2000) et Bradley (2007). Un des premiers travaux portant sur l'estimation à noyau de la fonction de régression sous la condition de α -mélange est dû à Györfi et al. (1989) qui établissent la convergence uniforme de l'estimateur de Nadaraya-Watson. Vieu (1991) a donné les termes asymptotiquement exacts de l'erreur quadratique du même estimateur. Dans ce cadre α -mélangeant, la convergence forte uniforme de l'estimateur de Nadaraya-Watson est traitée dans Doukhan (1994), Bosq (1998), Liebscher (2001) ainsi que dans les références qui y sont.

L'étude de données incomplètes sous la condition de mélange fort dans le cadre de l'estimation non paramétrique à noyau de la fonction de régression est très récente. Pour les données censurées à droite, Guessoum et Ould Saïd (2010, 2012) ont établi les propriétés asymptotiques (convergence uniforme presque sûre et normalité asymptotique) de l'estimateur de Nadaraya-Watson. Pour les données tronquées à gauche, Liang et al. (2009) ont établis la convergence uniforme presque sûre de l'estimateur à noyau introduit par Ould Saïd et Lemdani (2006) et Liang (2010) a étudié sa normalité asymptotique.

Dans cette thèse, nous étudions un autre concept de dépendance appelé association. Ce concept est un cas particulier de la notion de faible dépendance introduite par Doukhan et Louhichi (1999). Le concept d'association pour des variables aléatoires a deux origines. L'une est la physique mathématique avec les inégalités FKG (Fortuin, Kasteleyn et Ginibre (1971)) et leurs applications en

théorie de la percolation, au modèle d'Ising, l'autre est la statistique (Esary, Proschan et Walkup (1967)) avec des applications en fiabilité.

L'association se définit de la manière suivante :

Une suite finie de variables aléatoires X_1, \dots, X_N est dite associée si

$$\text{Cov}(f(X_1, \dots, X_N), g(X_1, \dots, X_N)) \geq 0, \quad (1.2)$$

pour toute paire f, g de fonctions de $\mathbb{R}^N \rightarrow \mathbb{R}$ croissantes coordonnée par coordonnée et telles que cette covariance existe. Une suite infinie de variables aléatoires est associée si toute sous suite finie est associée.

Les variables aléatoires indépendantes représentent un exemple des variables aléatoires associées (Esary, Proschan et Walkup (1967)). Il y a plusieurs exemples dans la fiabilité et l'analyse de survie où les variables aléatoires d'intérêt ne sont pas indépendantes mais associées. Par exemple, les variables aléatoires gaussiennes positivement corrélées (Pitt (1982)), les distributions exponentielles multivariées due à Marshall et Olkin (1967), etc. Plusieurs processus utilisés dans l'analyse des séries chronologiques sont aussi associés. Nagaraj et Reddy (1993) donnent des conditions pour que les processus autorégressifs et les processus linéaires soient associés. Lee, Rachev et Samorodnitsky (1990) donnent aussi les conditions nécessaires et suffisantes pour que les processus α -stables soient associés. Pour plus d'exemples et de détails sur le concept d'association, nous renvoyons le lecteur au livre de Bulinski et Shashkin (2007).

Le mélange et l'association sont deux notions complémentaires qui permettent de traiter la dépendance. Cependant, il est plus simple de vérifier si la condition (1.2) d'association est satisfaite ou non, plutôt que de voir si la suite de coefficients de mélange (α_n) tend vers 0. Notons aussi que les classes des processus α -mélangeants et des processus associés sont distinctes mais non disjointes, ceci est illustré dans l'exemple suivant. Soit $\{X_k\}_{k=1}^\infty$ le processus linéaire défini par

$$X_k = \sum_{j=0}^{\infty} a_j v_{k-j}$$

où $\{v_i\}_{i=-\infty}^\infty$ est une suite de variables aléatoires indépendantes identiquement distribuées centrées et de variance σ^2 , et $\{a_i\}_{i=-\infty}^\infty$ une suite de réels telle que $\sum_{i=1}^{\infty} a_i^2 < \infty$. $\{X_k\}_{k=1}^\infty$ est associée si $a_j \geq 0$. D'autre part, Pham and Tran (1985) montrent que $\{X_k\}_{k=1}^\infty$ est α -mélangeant sous des conditions convenables sur a_j . En particulier, Andrew (1984) montre que si $\{v_i\}_{i=-\infty}^\infty$ est une suite de variables aléatoires de Bernoulli et $a_j = \varsigma^j$, $0 < \varsigma \leq \frac{1}{2}$, $\{X_k\}_{k=1}^\infty$ n'est pas α -mélangeant, bien qu'il reste associé. Récemment, Doukhan et Louhichi (1999) ont introduit une nouvelle notion de dépendance faible qui permet de traiter le mélange et l'association dans une approche unifiée.

Les aspects probabilistes et statistiques de variables aléatoires associées ont été largement discutés dans la littérature, citons par exemple, Newman (1980) et Birkel (1988) pour le théorème central

limite, Newman et Wright (1981) et Cox et Grimmett (1984) pour le théorème central limite fonctionnel, Yu (1986) pour la loi du logarithme itéré, Dabrowski (1985) pour la loi du logarithme itéré fonctionnel, Newman (1984), Birkel (1989) et Bagai, Prakasa Rao (1995) et Azevedo (2010) pour la loi forte des grands nombres, Newman et Wright (1981) et Yu (1996) pour le principe d'invariance.

L'estimation non paramétrique à noyau sous des données complètes et associées est largement étudiée dans la littérature. Roussas (1991) a établi sous des conditions de régularité, la convergence ponctuelle et uniforme de l'estimateur à noyau de la densité et ses dérivées et de la fonction de hasard en donnant la vitesse de convergence. Bagai et Prakasa Rao (1995) ont également étudié l'estimateur à noyau de la densité et de la fonction de hasard, ils ont établi la consistance forte ponctuelle et uniforme, sans vitesse de convergence. Roussas (2000) a établi la normalité asymptotique de l'estimateur à noyau de la densité. Masry (2002) a prouvé la convergence uniforme et la normalité asymptotique avec vitesse de convergence de la densité d'une variable vectorielle. Douge (2007) a donné une vitesse de convergence presque sûre uniforme sur les ensembles compacts de l'estimateur à noyau de la densité. Des propriétés asymptotiques des estimateurs à noyau de la densité et de la fonction de régression se trouvent aussi dans le livre de Oliviera (2012).

Dans le cas de données incomplètes et associées, il n'existe pas beaucoup de travaux faits en estimation non paramétrique par la méthode de noyau pour ce type de données. Pour les données censurées à droite, on peut citer Cai et Roussas (1998) qui établissent des propriétés asymptotiques (convergence uniforme et normalité asymptotique) de l'estimateur de Kaplan Meier de la fonction de répartition. Ferrani et al. (2014) ont étudié la convergence uniforme presque sûre de l'estimateur à noyau de la densité et du mode avec vitesse. Menni et Tatachak (2016) ont étudié la convergence uniforme presque sûre de l'estimateur à noyau de la fonction de régression en caractérisant la vitesse de convergence. Pour les données tronquées à gauche, Guessoum et al. (2012) établissent la consistance uniforme forte de l'estimateur de Lynden-Bell de la fonction de répartition. Dans cette thèse, nous donnons une vitesse de convergence presque sûre et uniforme de l'estimateur à noyau de la fonction de régression introduit par Ould Saïd et Lemdani (2006). Nous y montrons aussi sa normalité asymptotique.

Cette thèse est organisée en 6 chapitres : dans le chapitre 2, nous présentons la notion de dépendance "association" en donnant la définition et quelques propriétés et exemples de variables associées. Ensuite, nous nous intéressons à l'estimation de la densité, nous établissons la convergence uniforme presque sûre de l'estimateur à noyau de la fonction de densité pour des variables associées à valeurs dans \mathbb{R}^d et nous étendons aussi les résultats de Roussas (2000) concernant la normalité asymptotique au cas des variables aléatoires associées à valeurs dans \mathbb{R}^d . Nous étudions également la convergence uniforme presque sûre de l'estimateur à noyau de la fonction de régression. Dans le chapitre 3,

nous nous intéressons à l'estimation non paramétrique par la méthode du noyau de la fonction de régression pour des données tronquées à gauche. Nous présentons d'abord le modèle aléatoire de troncature à gauche et les estimateurs de la fonction de répartition, la densité de la covariable et la fonction de régression sous ce modèle. Ensuite, nous étendons les résultats de Ould Saïd et Lemdani (2006) sur la convergence uniforme presque sûre et la normalité asymptotique de ces estimateurs obtenus dans le cas des suites de variables aléatoires indépendantes identiquement distribuées avec une covariable réelle au cas d'une covariable à valeurs dans \mathbb{R}^d ($d \geq 1$). Pour conclure ce chapitre 3, nous énonçons ensuite, sans en donner les preuves, les résultats de Liang, Li et Qi (2009) et Liang (2011) où ils étendent les résultats obtenus dans le cas indépendant au cas des suites de variables aléatoires α -mélangeantes. Les chapitres 4 et 5 constituent l'apport principal de cette thèse et ils sont dédiés à l'étude de l'estimateur à noyau de la fonction de régression sous le modèle aléatoire de troncature à gauche lorsque les données présentent une forme de dépendance qui est l'association. Plus précisément, au chapitre 4, nous donnons une vitesse de convergence uniforme presque sûre de cet estimateur en utilisant l'inégalité exponentielle type Bernstein due à Doukhan et Neumann (2007). Des simulations assez poussées viennent conforter nos résultats à taille d'échantillons finie. Les résultats sont présentés dans la version originale c'est-à-dire celle à paraître dans la revue *Journal of NonParametric statistics*. Le cinquième est une extension de nos résultats à l'étude de la normalité asymptotique de l'estimateur considéré. Nous y incluons aussi des intervalles de confiance pour la fonction de régression. Ici encore, des simulations sont menées pour mettre en évidence le bon comportement gaussien de l'estimateur ainsi que de calcul des probabilités de couverture à 95%. Le dernier chapitre est une conclusion de nos travaux. Nous récapitulons ce que nous avons fait et nous donnons des perspectives de recherches futures.

Chapitre 2

Estimation de la densité et de la fonction de régression pour des données complètes associées

Dans ce chapitre, nous présentons une certaine notion de dépendance appelée "association". Nous en donnons la définition et nous citons quelques propriétés et exemples de variables associées. Ensuite, dans un premier temps, nous nous intéressons à l'estimation de la densité quand les observations forment une suite de variables aléatoires associée. Récemment, plusieurs auteurs se sont intéressés à l'étude des propriétés asymptotiques de l'estimateur à noyau de la densité pour les variables associées. Citons par exemple à cet égard les travaux de Roussas (1991, 2000), Bagai et Prakasa Rao (1995), Masry (2002), Douge (2007) et Oliveira (2012). Nous nous établissons une vitesse de convergence uniforme presque sûre de l'estimateur à noyau de la fonction de densité pour des variables associées à valeurs dans \mathbb{R}^d et nous étendons aussi les résultats de Roussas (2000) concernant la normalité asymptotique au cas des variables aléatoires associées à valeurs dans \mathbb{R}^d . Nous étudions dans un second temps, la convergence uniforme presque sûre de l'estimateur à noyau de la fonction de régression. Pour l'estimation de la fonction de régression pour des données associées, on citera les résultats établis dans Oliveira (2012).

2.1 Données associées

Le concept d'association a été introduit et étudié par Esary, Proschan et Walkup en 1967. Depuis, cette notion de dépendance a suscité une grande attention en mécanique statistique, étant connu comme inégalités de FKG après la contribution par Fortuin, Kasteleyn et Ginibre (FKG, 1971). Ce concept est une généralisation d'une forme de dépendance positive introduite par Lehmann (1966) définie comme suit :

Définition 2.1. *Deux variables aléatoires X et Y sont dites positivement dépendant par quadrant*

(PQD) si pour tous réels x et y

$$\mathbb{P}(X \leq x, Y \leq y) - \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) \geq 0,$$

ou pour toutes fonctions croissantes f et g

$$\text{Cov}(f(X), g(Y)) \geq 0.$$

L'association se définit alors de la manière suivante :

Définition 2.2. Une suite finie de variables aléatoires X_1, \dots, X_N est dite associée si

$$\text{Cov}(f(X_1, \dots, X_N), g(X_1, \dots, X_N)) \geq 0, \quad (2.1)$$

pour toute paire f, g de fonctions de $\mathbb{R}^N \rightarrow \mathbb{R}$ croissantes coordonnée par coordonnée et telles que cette covariance existe.

L'extension de la définition de l'association aux suites infinies de variables aléatoires est donnée par :

Définition 2.3. On dit qu'une suite infinie de variables aléatoires est associée si toute sous suite finie est associée.

Remarque 2.1. (Sur la définition de l'association)

1) On peut remplacer paire de fonctions croissantes par paire de fonctions décroissantes puisque

$$\text{Cov}(f(X), g(Y)) = \text{Cov}(-f(X), -g(Y)).$$

2) Pour vérifier l'association d'un vecteur aléatoire, il peut être commode de disposer de caractérisations où les fonctions de test f, g de (2.1) appartiennent à une classe de fonctions particulières. On peut montrer que pour vérifier (2.1), on peut se contenter de la vérifier pour les fonctions croissantes binaires, pour les fonctions croissantes continues bornées ou pour les fonctions croissantes bornées admettant des dérivées partielles d'ordre 1 bornées .

2.1.1 Propriétés d'association

Nous présentons ici quelques propriétés basiques de l'association données dans Esary et al. (1967).

- a) Un sous-ensemble d'un ensemble fini de variables aléatoires associées est associé.
- b) Si deux ensembles de variables associées sont indépendants l'un de l'autre, leur union est un ensemble associé.
- c) Tout singleton formé d'une variable aléatoire est associé.
- d) Des fonctions décroissantes (ou croissantes) de variables aléatoires associées sont associées.
- e) Si $X^{(k)} = (X_1^{(k)}, \dots, X_N^{(k)})$ est un vecteur associé pour tout $k \geq 1$, et $X^{(k)}$ converge en loi vers $X = (X_1, \dots, X_N)$ quand k tend vers $+\infty$, alors X est associé.

Remarque 2.2. Les propriétés b) et c) entraînent l'association de tout ensemble de variables aléatoires indépendantes.

2.1.2 Exemples de vecteurs et processus associés

Nous donnons maintenant quelques exemples de vecteurs et processus associés.

- 1) Soit $X = (X_1, \dots, X_N)$ un vecteur associé et $Y = (X_{N1}, \dots, X_{NN})$ le vecteur de statistiques d'ordre engendré par X . Alors Y est aussi associé.
- 2) Pour $X = (X_1, \dots, X_N)$ associé et $i = 1, \dots, n$, notons $S_i = \sum_{j=1}^i X_j$. Alors le vecteur (S_1, \dots, S_N) des sommes partielles est associé.
- 3) Tout vecteur gaussien positivement corrélé $X = (X_1, \dots, X_N)$ (i.e. $Cov(X_i, X_j) \geq 0$, $i, j = 1, \dots, N$) est associé. Ce résultat est dû à Pitt (1982).
- 4) Un vecteur $X = (X_1, X_2)$ de variables binaires est associé si et seulement si $Cov(X_1, X_2) \geq 0$.
- 5) *Processus autorégressif d'ordre p* :

Soit $(X_k)_{k \in \mathbb{Z}}$ un processus stationnaire autoregressif d'ordre p défini pour tout $k \in \mathbb{Z}$ par

$$X_k = \phi_1 X_{k-1} + \dots + \phi_p X_{k-p} + \epsilon_k,$$

avec $(\epsilon_k)_{k \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes de moyenne nulle et de variance égale à 1. $(X_k)_{k \in \mathbb{Z}}$ est associé si $\phi_i \geq 0$, $1 \leq i \leq p$. Si $p = 1$ et $\phi_1 < 0$. Alors (X_{2^k}) et (X_{2^k+1}) sont des suites associées (Nagaraj and Reddy (1993)).

- 6) *Processus linéaire* :

Soit $(X_k)_{k \in \mathbb{Z}}$ le processus linéaire défini pour tout $k \in \mathbb{Z}$ par

$$X_k = \sum_{i=0}^{\infty} a_i \epsilon_{k-i},$$

avec $(a_i)_{i \in \mathbb{Z}}$ une suite de réels nonnégatifs telle que $\sum_{i=0}^{\infty} a_i < \infty$ et $(\epsilon_i)_{i \in \mathbb{Z}}$ est une suite de variables aléatoires indépendantes de moyenne nulle et de variance égale à 1. Alors le processus $(X_k)_{k \in \mathbb{Z}}$ est associé (Nagaraj and Reddy (1993)).

2.1.3 Autres types de dépendance

Dans cette sous section nous introduisons brièvement d'autres formes de dépendance appelées association positive, association négative, quasi-association et dépendance faible.

2.1.3.1 Association Positive

Burton, Dabrowski et Dehling (1986) ont introduit la notion d'association positive dite aussi association faible; ils ont formulé ce concept pour des suites de variables aléatoires à valeurs dans \mathbb{R}^d .

Définition 2.4. Une suite finie de variables aléatoire X_1, \dots, X_N est dite positivement associée si pour tous sous ensembles disjoints I et J de $\{1, 2, \dots, N\}$ et toutes fonctions croissantes f et g définies sur $\mathbb{R}^{|I|}$ et $\mathbb{R}^{|J|}$,

$$\text{Cov}(f(Y_i, i \in I), g(Y_j, j \in J)) \geq 0,$$

lorsque cette covariance existe.

Une suite infinie de variables aléatoires est positivement associée si toute sous suite finie est positivement associée.

L'association positive est strictement plus faible que l'association. Ainsi, l'association implique l'association positive.

2.1.3.2 Association Négative

La notion de variables négativement associées a été introduit par Joag-Dev et Proschan (1983) ; ils ont donné de nombreuses propriétés et proposé plusieurs applications en Statistique.

Définition 2.5. Une suite finie de variables aléatoires X_1, \dots, X_N est dite négativement associée si pour tous sous ensembles disjoints I et J de $\{1, 2, \dots, N\}$ et toutes fonctions croissantes f et g définies sur $\mathbb{R}^{|I|}$ et $\mathbb{R}^{|J|}$,

$$\text{Cov}(f(Y_i, i \in I), g(Y_j, j \in J)) \leq 0,$$

lorsque cette covariance existe.

Une suite infinie de variables aléatoires est négativement associée si toute sous suite finie est négativement associée.

2.1.3.3 Quasi-Association

Une propriété fondamentale des variables positivement ou négativement associées est l'équivalence entre la non corrélation et l'indépendance. Bulinski et Suquet (2001) ont introduit un nouveau concept de dépendance appelé quasi-association qui permet d'étudier les variables aléatoires possédant cette propriété.

Définition 2.6. Une suite $\{X_i; i \geq 0\}$ de variables aléatoires réelles est dite quasi-associée si, pour tous sous-ensembles finis disjoints I et J de \mathbb{N} et toutes fonctions lipschitziennes $f : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ et $g : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$, on a

$$|\text{Cov}(f(X_i, i \in I), g(X_j, j \in J))| \leq \sum_{i \in I} \sum_{j \in J} \text{Lip}_i(f) \text{Lip}_j(g) |\text{Cov}(X_i, X_j)| \quad (2.2)$$

où les constantes de lipschitz $\text{Lip}_i(f)$ sont telles que pour $x = (x_i, i \in I)$, $y = (y_i, i \in I)$ dans $\mathbb{R}^{|I|}$

$$|f(x) - f(y)| \leq \sum_{i \in I} \text{Lip}_i(f) |x_i - y_i|$$

avec

$$Lip_i(f) = \sup_{x_i \neq x'_i} \frac{|f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{|I|}) - f(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_{|I|})|}{|x_i - x'_i|}. \quad (2.3)$$

le sup étant pris pour $x_1, x_2, \dots, x_{|I|}, x'_i \in \mathbb{R}$.

Shashkin (2002) a prouvé que toute suite de variables aléatoires gaussiennes est quasi-associée. Bulinski et Shabanovich (1998) ont montré que, toute suite de variables aléatoires positivement ou négativement associée admettant un moment d'ordre 2 fini satisfait (2.2). Le résultat est donné dans le théorème suivant que l'on trouve dans le livre de Bulinski et Shashkin (2007).

Théorème 2.1. (*Bulinski et Shashkin (2007), Theorem 5.3, p.89*) Soit $\{X_t; t \in \mathbb{Z}\}$ une suite de variables aléatoires associée, ou positivement ou négativement associée, tel que $\mathbb{E}(X_t^2) < \infty$ pour tout $t \in \mathbb{Z}$. Soit I et J deux sous-ensembles finis de \mathbb{Z} . Dans le cas de l'association positive ou l'association négative, nous supposons, en plus, que I et J sont disjoints. Alors, pour toutes fonctions lipschitziennes $f : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$ et $g : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$, on a

$$|Cov(f(X_i, i \in I), g(X_j, j \in J))| \leq \sum_{i \in I} \sum_{j \in J} Lip_i(f) Lip_j(g) |Cov(X_i, X_j)|.$$

Par conséquent, l'association positive ou l'association négative impliquent la quasi-association.

2.1.3.4 Dépendance faible

Doukhan et Louhichi (1999) ont défini une notion de dépendance faible qui permet de résumer les propriétés de dépendance de très nombreux modèles via une suite de coefficients.

Définition 2.7. La suite $\{X_N; N \in \mathbb{N}\}$ de variables aléatoires est dite $(\theta, \mathcal{F}, \psi)$ -faiblement dépendante, s'il existe une classe \mathcal{F} de fonctions réelles, une suite $\theta = (\theta_r)_{r \in \mathbb{N}}$ décroissante vers zéro à l'infini, et une fonction ψ d'arguments $(h, k, u, v) \in \mathcal{F}^2 \times \mathbb{N}^2$, telle que pour tous $(u + v)$ -uple $(i_1, \dots, i_u, j_1, \dots, j_v)$, avec

$$i_1 \leq \dots \leq i_u < i_u + r \leq j_1 \leq \dots \leq j_v,$$

on a

$$Cov(h(X_{i_1}, \dots, X_{i_u}), k(X_{j_1}, \dots, X_{j_v})) \leq \psi(h, k, u, v) \theta_r,$$

pour toute fonctions $f, g \in \mathcal{F}$ définies respectivement sur \mathbb{R}^u et \mathbb{R}^v .

Le lemme ci-dessous spécifie la structure de dépendance faible des variables associées. Soit \mathcal{L} l'ensemble des fonctions lipschitziennes bornées de \mathbb{R}^u dans \mathbb{R} , pour $u \in \mathbb{N}^*$ et \mathcal{C}_b^1 est l'ensemble des

fonctions différentiables à dérivées partielles continues et bornées. De plus,

$$Lip(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_1}, \quad (2.4)$$

denote le module de continuité lipschitzien de h avec $\|x\|_1 = |x_1| + \dots + |x_u|$ pour $x = (x_1, \dots, x_u)$.

Lemme 2.1. (Doukhan et Louhichi (1999), Lemma 4, p.321) Si $\{X_N; N \in \mathbb{N}\}$ est une suite de variables aléatoires associées centrées, alors $(X_N)_N$ est $(\theta, \mathcal{L}, \psi)$ -faiblement dépendante avec

$$\theta_r = \sup_i \sum_{j: |i-j| \geq r} Cov(X_i, X_j) \text{ et } \psi(h, k, u, v) = \min(u, v) Lip(h) Lip(k),$$

et elle est $(\theta, \mathcal{C}_b^1, \psi)$ -faiblement dépendante avec

$$\theta_r = \sup_{j: |i-j| \geq r} Cov(X_i, X_j) \text{ et } \psi(h, k, u, v) = \sum_{i=1}^u \sum_{j=1}^v \left\| \frac{\partial h}{\partial x_i} \right\|_{\infty} \times \left\| \frac{\partial k}{\partial x_j} \right\|_{\infty}.$$

Le lemme 2.1 est une conséquence immédiate des deux résultats suivants

Lemme 2.2. (Newman (1984), Proposition 15, p.134) Soit $\{X_N, N \in \mathbb{N}\}$ une suite de variables aléatoires associées et soient A et B deux sous-ensembles de \mathbb{N} . Soient f, g, f_1 et g_1 des fonctions réelles telles que $f_1 - f, f_1 + f, g_1 - g$ et $g_1 + g$ sont des fonctions croissantes. Alors,

$$|Cov(f(X_i, i \in A), g(X_j, j \in B))| \leq |Cov(f_1(X_i, i \in A), g_1(X_j, j \in B))|.$$

Lemme 2.3. (Birkel(1988), Lemma 3.1, p.1692) Soit $\{X_N, N \in \mathbb{N}\}$ une suite de variables aléatoires associées. Soient A et B deux sous-ensembles de \mathbb{N} et f et g deux fonctions réelles définies sur $\mathbb{R}^{|A|}$ et $\mathbb{R}^{|B|}$ admettant des dérivées partielles bornées. Alors,

$$|Cov(f(X_i, i \in A), g(X_j, j \in B))| \leq \sum_{i \in A} \sum_{j \in B} \left\| \frac{\partial f}{\partial x_i} \right\|_{\infty} \times \left\| \frac{\partial g}{\partial x_j} \right\|_{\infty} Cov(X_i, X_j).$$

2.1.4 Inégalité exponentielle

L'outil principal que nous allons utiliser pour établir la convergence presque sûre et caractériser la vitesse de convergence est l'inégalité exponentielle de type Bernstein. Il existe plusieurs versions de ce type d'inégalité dans la littérature pour des variables dépendantes. Pour les variables associées, une première inégalité exponentielle a été établie par Prakasa Rao (1993). Ensuite, une autre inégalité qui était une extension au cadre de variables associées de l'inégalité de Bernstein dans le cas indépendant a été établie par Ioannides et Roussas (1999) dans le cas des variables bornées, et par Oliveira (2005) pour le cas des variables non bornées. Une nouvelle inégalité exponentielle a été établie récemment par Douge (2007). Cependant, pour établir nos résultats, nous utiliserons une inégalité exponentielle établie par Doukhan et Neumann (2007) pour des variables aléatoires réelles vérifiant une condition sur les covariances.

Théorème 2.2. (Doukhan et Neumann (2007), Theorem 1, p.880) Soient X_1, \dots, X_N des variables aléatoires réelles centrées définies sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et $\Psi : \mathbb{N}^2 \rightarrow \mathbb{N}$ une des fonctions suivantes :

- (i) $\Psi(u, v) = 2v$,
- (ii) $\Psi(u, v) = u + v$,
- (iii) $\Psi(u, v) = uv$,
- (iv) $\Psi(u, v) = a(u + v) + (1 - a)uv$, pour $a \in (0, 1)$.

Supposons qu'il existe des constantes $K, M, L_1, L_2 < \infty, \mu, \nu \geq 0$ et une suite décroissante de coefficients réels $(\delta(N))_{N \geq 0}$ tels que, pour tous u -uplet (s_1, \dots, s_u) et tous v -uplet (t_1, \dots, t_v) avec $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq N$, les inégalités suivantes sont vérifiées :

$$|Cov(X_{s_1} \cdots X_{s_u}, X_{t_1} \cdots X_{t_v})| \leq K^2 M^{u+v-2} ((u+v)!)^\nu \Psi(u, v) \delta(t_1 - s_u),$$

où

$$\sum_{s=0}^{\infty} (s+1)^k (\delta(s)) \leq L_1 L_2^k (k!)^\mu, \forall k \geq 0,$$

et

$$\mathbb{E}(|X_t|^k) \leq (k!)^\nu M^k, \forall k \geq 0.$$

Alors, pour tout $t \geq 0$,

$$\mathbb{P}\left(\sum_{i=1}^N X_i \geq t\right) \leq \exp\left(-\frac{t^2/2}{A_N + B_N^{1/(\mu+\nu+2)} t^{(2\mu+2\nu+3)/(\mu+\nu+2)}}\right) \quad (2.5)$$

où A_N peut être choisi comme tout nombre supérieur ou égal à $\sigma_N^2 := \text{Var}\left(\sum_{i=1}^N X_i\right)$ et

$$B_N = 2(K \vee M) L_2 \left(\frac{2^{4+\mu+\nu} N K^2 L_1}{A_N} \vee 1 \right).$$

2.2 Estimation de la densité

Les techniques utilisées pour étudier les estimateurs de la densité ou de la régression sont similaires. Cependant, le problème de l'estimation de la régression est quelque peu plus compliqué car nous avons besoin de traiter la dépendance entre la variable d'intérêt et la covariable. Nous commençons donc par établir des propriétés asymptotiques de l'estimateur de la densité, en introduisant les techniques dans un cadre simple et ensuite, on les étend à l'étude de la régression.

Dans cette section, nous établissons des propriétés asymptotiques de l'estimateur à noyau de la densité, en supposant, qu'elle existe. Nous nous intéressons au cas des données complètement observées et associées.

Soit $\{X_i; 1 \leq i \leq N\}$ une suite strictement stationnaire de variables aléatoires associées définie sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{R}^d ($d \geq 1$), ayant la même distribution que X qui admet une densité de probabilité inconnue v . On considère l'estimateur à noyau de Parzen-Rosenblatt de v défini par

$$\hat{v}_N(x) = \frac{1}{Nh_N^d} \sum_{i=1}^N K_d \left(\frac{x - X_i}{h_N} \right), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d, \quad (2.6)$$

où $K_d : \mathbb{R}^d \rightarrow \mathbb{R}$ est la fonction noyau et $(h_N)_{N \geq 1}$ une suite de nombres réels positifs appelée fenêtre tendant vers zéro à l'infini.

L'estimation à noyau de la densité pour ce type de données a été récemment traité dans la littérature par plusieurs auteurs. Dans le cas univarié, Roussas (1991) a étudié la consistance uniforme forte des estimateurs à noyau des dérivées d'ordre r ($r \geq 0$) de la densité pour des suites strictement stationnaires. Il a donné également une caractérisation de la vitesse de convergence. Bagai et Prakasa Rao (1995) ont établi la consistance ponctuelle et uniforme de l'estimateur à noyau, sans vitesse de convergence. Roussas (2000) a établi sa normalité asymptotique. Masry (2002) a étendu les résultats de Roussas (1991) au cas multivarié. Douge (2007) a établi une nouvelle inégalité exponentielle, qui lui a permis d'obtenir une vitesse de convergence presque sûre uniforme sur les ensembles compacts de l'estimateur à noyau de la densité dans le cas univarié d'ordre de $O \left(\frac{\log \log N}{N} \right)^{\rho/(2\rho+1)}$, $\rho \in]0, \infty[$. Oliveira (2012) (Chapitre 3, Théorème 3.31) a aussi établi la convergence presque sûre de l'estimateur à noyau de la densité dans le cas univarié avec une vitesse de convergence d'ordre de $O \left(\frac{\log \log N}{Nh_N} \right)^{\gamma/2}$, $\gamma \in (0, 1)$. Nous obtenons dans le Théorème 2.3 une vitesse de convergence presque sûre uniforme dans le cas multivarié. Pour cela, nous utilisons l'inégalité exponentielle établie par Doukhan et Neumann (2007) pour des variables aléatoires centrées donnée dans le Théorème 2.2.

2.2.1 Convergence uniforme presque sûre

Nous établissons la convergence uniforme presque sûre de l'estimateur \hat{v}_N sur un compact D de \mathbb{R}^d . Notons $\Delta_{i,j}$ le coefficient de covariance défini par

$$\Delta_{i,j} = \sum_{k=1}^d \sum_{l=1}^d \text{Cov}(X_{i,k}, X_{j,l}),$$

où $X_{i,k}$ est la k -ème composante de X_i .

Les hypothèses que nous devons imposer pour établir ce résultat sont les suivantes :

(H2.1) La fenêtre h_N satisfie :

$$h_N \rightarrow 0, \quad Nh_N^d \rightarrow \infty \quad \text{et} \quad \frac{\log^5 N}{Nh_N^d} \rightarrow 0 \quad \text{quand} \quad N \rightarrow +\infty;$$

(H2.2) La fonction noyau satisfait

(i) K_d est une densité de probabilité bornée à support compact et hölderienne d'exposant $\beta > 0$;

(ii) Pour tout $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $K_d(x) = \prod_{i=1}^d K(x_i)$, avec K est une fonction noyau lipschitzienne définie sur \mathbb{R} ;

(iii) $\int_{\mathbb{R}^d} z_i K_d(z) dz = 0$, $i = 1, \dots, d$, $\int_{\mathbb{R}^d} |z_1^{i_1} \dots z_d^{i_d}| K_d(z) dz < +\infty$, $i_1 + \dots + i_d = 2$,
 $\int_{\mathbb{R}^d} K_d^2(z) dz < +\infty$ et $\int_{\mathbb{R}^d} |z_1 + \dots + z_d| K_d^2(z) dz < +\infty$;

(H2.3) Le terme de covariance définit par $\rho(s) := \sup_{|i-j| \geq s} \Delta_{i,j}$, $s > 0$, vérifie $\rho(s) \leq \gamma_0 e^{-\gamma s}$, $\gamma_0 > 0$, $\gamma > 0$;

(H2.4) La densité $v(\cdot)$ est bornée et deux fois différentiable avec $\sup_{x \in D} \left| \frac{\partial^k v}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$, $i, j = 1, \dots, d$, $k = 1, 2$;

(H2.5) la densité jointe $v_{i,j}$ du couple (X_i, X_j) pour $i \neq j$ existe et satisfait $\sup_{x_1, x_2 \in D} |v_{i,j}(x_1, x_2)| < +\infty$.

Les différentes hypothèses sont classiques en estimation fonctionnelle. Les hypothèses (H2.1), (H2.2) et (H2.4) portent respectivement sur la fenêtre h_N , le noyau K_d et la densité v . L'hypothèse (H2.5) est indispensable pour avoir une bonne majoration des covariances. L'hypothèse (H2.3) quantifie la décroissance exponentielle de terme de covariance qui nous permet d'établir la convergence presque sûre uniforme.

Nous allons maintenant donner le résultat dans le théorème suivant

Théorème 2.3. *Sous les hypothèses (H2.1)-(H2.5), on a*

$$\sup_{x \in D} |\hat{v}_N(x) - v(x)| = O \left(\sqrt{\frac{\log N}{N h_N^d}} \vee h_N^2 \right) \text{ p.s, lorsque } N \rightarrow +\infty.$$

En particulier si nous choisissons $h_N = O \left(\left(\frac{\log N}{N} \right)^{1/(d+4)} \right)$, nous obtenons la vitesse de convergence optimale

$$\sup_{x \in D} |\hat{v}_N(x) - v(x)| = O \left(\left(\frac{\log N}{N} \right)^{2/(d+4)} \right), \text{ p.s, lorsque } N \rightarrow +\infty,$$

optimalité au sens du critère de convergence uniforme de l'estimation de la densité (Stute (1982)).

Preuve du Théorème 2.3. Dans toute la suite, C et c désigneront des constantes dont les valeurs peuvent varier d'une expression à une autre.

Afin de pouvoir utiliser le Théorème 2.2, nous devons montrer d'abord que les variables $U_i(x)$ définies par

$$U_i(x) := U(X_i; x) = K_d \left(\frac{x - X_i}{h_N} \right) - \mathbb{E} \left(K_d \left(\frac{x - X_i}{h_N} \right) \right) \quad (2.7)$$

vérifient les conditions de ce théorème. Cela fait l'objet du lemme suivant. Dans ce Lemme, pour alléger les notations, on note $U_i(x)$ par U_i .

Lemme 2.4. *Sous les hypothèses (H2.2)-(H2.5), il existe des constantes $K, M, L_1, L_2 < \infty, \mu, \lambda \geq 0$ et une suite décroissante de coefficients réels $(\delta(n))_{n \geq 0}$, telles que pour tout u -uplet $(s_1, \dots, s_u) \in \mathbb{N}^u$ et tout v -uplet $(t_1, \dots, t_v) \in \mathbb{N}^v$ avec $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$, on a*

- a) $Cov(U_{s_1} \dots U_{s_u}, U_{t_1} \dots U_{t_v}) \leq c^{u+v} h_N^{-2} uv \delta(t_1 - s_u),$
- b) $\sum_{s=0}^{\infty} (s+1)^k \delta(s) \leq L_1 L_2^k (k!)^\mu, \forall k \geq 0,$
- c) $\mathbb{E}(|U_i|^k) \leq (k!)^\lambda M^k, \forall k \geq 0.$

Preuve. Pour montrer a), nous allons utiliser le Théorème 2.1. Pour cela, nous avons à déterminer les constantes de lipschitz de la fonction $\phi_m, m \in \mathbb{N}$ définie pour

$$w = (w_1, \dots, w_m) = (w_{1,1}, \dots, w_{1,d}, \dots, w_{m,1}, \dots, w_{m,d}) \in \mathbb{R}^{md}$$

par

$$\phi_m(w) = \prod_{i=1}^m \left[K_d \left(\frac{x_1 - w_{i,1}}{h_N}, \dots, \frac{x_d - w_{i,d}}{h_N} \right) - \mathbb{E} \left(K_d \left(\frac{x - X_1}{h_N} \right) \right) \right].$$

En utilisant l'hypothèse (H2.2)(ii), on peut écrire

$$\phi_m(w) = \prod_{i=1}^m \left[\left(\prod_{j=1}^d K \left(\frac{x_j - w_{i,j}}{h_N} \right) \right) - \mathbb{E} \left(K_d \left(\frac{x - X_1}{h_N} \right) \right) \right].$$

D'après (2.3), les constantes de lipschitz de la fonction ϕ_m , notées $Lip_s(\phi_m)$ pour $s = (i-1)d + k, 1 \leq i \leq m, 1 \leq k \leq d$ sont données par

$$\begin{aligned} Lip_s(\phi_m) &= \sup_{\substack{w_{i,k} \neq w'_{i,k} \\ l=1 \\ l \neq i}} \prod_{l=1}^m |U(w_l; x)| \frac{\left| \prod_{\substack{j=1 \\ j \neq k}}^d K \left(\frac{x_j - w_{i,j}}{h_N} \right) \left[K \left(\frac{x_k - w_{i,k}}{h_N} \right) - K \left(\frac{x_k - w'_{i,k}}{h_N} \right) \right] \right|}{|w_{i,k} - w'_{i,k}|} \\ &\leq (2\|K_d\|_\infty)^{m-1} (\|K\|_\infty)^{d-1} \sup_{w_{i,k} \neq w'_{i,k}} \frac{\left| K \left(\frac{x_j - w_{i,j}}{h_N} \right) - K \left(\frac{x_k - w'_{i,k}}{h_N} \right) \right|}{|w_{i,k} - w'_{i,k}|} \\ &\leq (2\|K_d\|_\infty)^{m-1} \|K\|_\infty^{d-1} h_N^{-1} Lip(K), \end{aligned} \tag{2.8}$$

où $Lip(K)$ désigne le module de lipschitz de K donné par (2.4).

En utilisant maintenant le Théorème 2.1 sous l'hypothèse (H2.2)(ii), on a

$$|Cov(U_{s_1} \dots U_{s_u}, U_{t_1} \dots U_{t_v})| \leq \sum_{i=s_1}^{s_u} \sum_{j=t_1}^{t_v} Lip_i(U_{s_1} \dots U_{s_u}) Lip_j(U_{t_1} \dots U_{t_v}) \Delta_{i,j}. \tag{2.9}$$

Remarquons que

$$\prod_{i=1}^u U_{s_i} = \prod_{i=1}^u U(X_{s_i}; x) = \phi_u(X_{s_1}, \dots, X_{s_u})$$

et

$$\prod_{i=1}^v U_{t_i} = \prod_{i=1}^v U(X_{t_i}; x) = \phi_v(X_{t_1}, \dots, X_{t_v}).$$

Ainsi, en utilisant (2.8), on obtient pour $s = (i-1)d + k, 1 \leq i \leq u, 1 \leq k \leq d$,

$$Lip_s(\phi_u) \leq (2\|K_d\|_\infty)^{u-1} \|K\|_\infty^{d-1} h_N^{-1} Lip(K) \quad (2.10)$$

et pour $s = (i-1)d + k, 1 \leq i \leq v, 1 \leq k \leq d$,

$$Lip_s(\phi_v) \leq (2\|K_d\|_\infty)^{v-1} \|K\|_\infty^{d-1} h_N^{-1} Lip(K). \quad (2.11)$$

En combinant (2.10) et (2.11), l'inégalité (2.9) devient

$$|Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})| \leq (2\|K_d\|_\infty)^{u+v-2} (\|K\|_\infty^{d-1} Lip(K))^2 h_N^{-2} \sum_{i=s_1}^{s_u} \sum_{j=w_1}^{w_v} \Delta_{i,j}. \quad (2.12)$$

L'hypothèse **(H2.3)** et la stationnarité entraînent que

$$|Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})| \leq C^{u+v} h_N^{-2} uv \rho(t_1 - s_u). \quad (2.13)$$

D'autre part, sous les hypothèses **(H2.2)(i)**, **(H2.4)** et **(H2.5)**, nous avons

$$\mathbb{E} \left(K_d \left(\frac{x - X_i}{h_n} \right) K_d \left(\frac{x - X_j}{h_n} \right) \right) \leq C h_N^{2d} \quad (2.14)$$

et

$$\mathbb{E} \left(K_d \left(\frac{x - X_i}{h_n} \right) \right) \leq C h_N^d. \quad (2.15)$$

Nous déduisons de (2.14) et (2.15) que

$$|Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})| \leq C^{u+v} h_N^{2d}. \quad (2.16)$$

En écrivant

$$Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v}) = Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})^{\frac{d}{2d+2}} \cdot Cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})^{\frac{d+1}{2d+2}}$$

et en utilisant (2.13) et (2.16), nous obtenons

$$|cov(U_{s_1} \cdots U_{s_u}, U_{t_1} \cdots U_{t_v})| \leq C^{u+v} h_N^d uv (\rho(t_1 - s_u))^{\frac{d}{2d+2}}. \quad (2.17)$$

Posons $\delta(\cdot) = \rho^{\frac{d}{2d+2}}(\cdot)$, on obtient a). Pour tout $k \geq 0$, sous l'hypothèse **(H2.3)**, nous avons

$$\begin{aligned} \sum_{s=0}^{\infty} (s+1)^k (\rho(s))^{d/(2d+2)} &\leq \sum_{s=0}^{\infty} (s+1) \cdots (s+k) e^{-\gamma s d/(2d+2)} \\ &= \frac{d^k}{dt^k} \left(\frac{1}{1-t} \right) \Big|_{t=e^{-\gamma d/(2d+2)}} \\ &= (k)! \left(\frac{1}{1-e^{-\gamma d/(2d+2)}} \right)^{k+1}. \end{aligned}$$

En choisissant $\mu = 1$ et $L_1 = L_2 = \frac{1}{1-e^{-\gamma d/(2d+2)}}$, nous obtenons le résultat b).

Pour la preuve de c), nous avons pour tout $k \geq 0$,

$$\begin{aligned}\mathbb{E}(|U_i|^k) &\leq (2\|K_d\|_\infty)^k \\ &= (k!)^\lambda M^k,\end{aligned}$$

où $\lambda = 0$ et $M = 2\|K_d\|_\infty$. Ceci termine la preuve de Lemme 2.4. \square

Revenons maintenant à la preuve du Théorème 2.3. On a

$$\begin{aligned}\sup_{x \in D} |\hat{v}_N(x) - v(x)| &= \sup_{x \in D} |\hat{v}_N(x) - \mathbb{E}(\hat{v}_N(x)) + \mathbb{E}(\hat{v}_N(x)) - v(x)| \\ &\leq \sup_{x \in D} |\hat{v}_N(x) - \mathbb{E}(\hat{v}_N(x))| + \sup_{x \in D} |\mathbb{E}(\hat{v}_N(x)) - v(x)| \\ &=: J_1 + J_2.\end{aligned}\tag{2.18}$$

Remarquons que $\hat{v}_N(x) - \mathbb{E}(\hat{v}_N(x)) = \frac{1}{Nh_N^d} \sum_{i=1}^N U_i(x)$. Ainsi, $J_1 = \sup_{x \in D} \left| \frac{1}{Nh_N^d} \sum_{i=1}^N U_i(x) \right|$. Pour étudier le terme J_1 , nous employons une technique classique qui consiste à utiliser le recouvrement d'un compact. Ainsi, l'ensemble compact D peut être recouvert par un nombre fini l_N de boules $B_k(x_k, a_N^d)$ de centres $x_k = (x_{k,1}, \dots, x_{k,d})$, $k \in \{1, \dots, l_N\}$ et de même rayon $a_N^d = N^{-\frac{1}{2\beta}} h_N^{1+\frac{d}{2\beta}}$ où β est donné dans l'hypothèse **(H2.2)(i)**.

Comme D est borné, alors il existe une constante C tel que $l_N a_N^d \leq C$. Pour tout x dans D , il existe une boule B_k qui contient x tel que $\|x - x_k\| \leq a_N^d$, $\|\cdot\|$ est une norme définie sur \mathbb{R}^d . Ensuite, on utilise la décomposition suivante

$$\begin{aligned}J_1 &= \sup_{x \in D} \left| \frac{1}{Nh_N^d} \sum_{i=1}^N (U_i(x) - U_i(x_k)) - \frac{1}{nh_N^d} \sum_{i=1}^N U_i(x_k) \right| \\ &\leq \max_{1 \leq k \leq l_N} \sup_{x \in B_k} \frac{1}{Nh_N^d} \left| \sum_{i=1}^N (U_i(x) - U_i(x_k)) \right| + \max_{1 \leq k \leq l_N} \frac{1}{Nh_N^d} \left| \sum_{i=1}^N U_i(x_k) \right| \\ &=: J_{11} + J_{12}.\end{aligned}\tag{2.19}$$

Commençons par le terme J_{11} . Nous avons

$$\begin{aligned}\sup_{x \in B_k} \left\{ \frac{1}{Nh_N^d} \left| \sum_{i=1}^N (U_i(x) - U_i(x_k)) \right| \right\} &\leq \sup_{x \in B_k} \left\{ \frac{1}{Nh_N^d} \sum_{i=1}^N \left| \left[K_d \left(\frac{x - X_i}{h_N} \right) - K_d \left(\frac{x_k - X_i}{h_N} \right) \right] \right. \right. \\ &\quad \left. \left. - \mathbb{E} \left[K_d \left(\frac{x - X_i}{h_N} \right) - K_d \left(\frac{x_k - X_i}{h_N} \right) \right] \right| \right\} \\ &\leq \sup_{x \in B_k} \left\{ \frac{1}{Nh_N^d} \sum_{i=1}^N \left| K_d \left(\frac{x - X_i}{h_N} \right) - K_d \left(\frac{x_k - X_i}{h_N} \right) \right| \right. \\ &\quad \left. + \frac{1}{h_N^d} \mathbb{E} \left[\left| K_d \left(\frac{x - X_1}{h_N} \right) - K_d \left(\frac{x_k - X_1}{h_N} \right) \right| \right] \right\}.\end{aligned}$$

Ainsi, l'hypothèse **(H2.2)(i)** entraîne que

$$\begin{aligned} \sup_{x \in B_k} \left\{ \frac{1}{Nh_N^d} \left| \sum_{i=1}^N (U_i(x) - U_i(x_k)) \right| \right\} &\leq \frac{2 \sup_{x \in B_k} \|x - x_k\|^\beta}{h_N^{d+\beta}} \\ &\leq C(a_n^d)^\beta h_N^{-d-\beta} = O\left(\frac{1}{\sqrt{Nh_N^d}}\right). \end{aligned}$$

D'où,

$$J_{11} = \max_{1 \leq k \leq l_N} \sup_{x \in B_k} \frac{1}{Nh_N^d} \left| \sum_{i=1}^N (U_i(x) - U_i(x_k)) \right| = O\left(\frac{1}{\sqrt{Nh_N^d}}\right). \quad (2.20)$$

Etudions maintenant le terme J_{12} dans (2.19). On utilisant le Lemme 2.4 et le Théorème 2.2, nous avons, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\sum_{i=1}^N U_i(x) \geq \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2/2}{A_N + B_N^{1/(\mu+\lambda+2)} \varepsilon^{(2\mu+2\lambda+3)/(\mu+\lambda+2)}}\right), \quad (2.21)$$

où A_N peut être choisi tel que $A_N \leq \sigma_N^2$ avec

$$\sigma_N^2 := \text{Var}\left(\sum_{i=1}^N U_i(x)\right),$$

et

$$B_N = 2cL_2 \left(\frac{2^{4+\mu+\lambda} N c h_N^d L_1}{A_N} \vee 1 \right).$$

Pour déterminer σ_N^2 , nous devons calculons $\text{Var}\left(\sum_{i=1}^N U_i(x)\right)$. Nous avons

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^N U_i(x)\right) &= (Nh_N^d)^2 \text{Var}(\hat{v}_N(x)) \\ &= \text{Var}\left(\sum_{i=1}^N K_d\left(\frac{x - X_i}{h_N}\right)\right) \\ &= N \text{Var}\left(K_d\left(\frac{x - X_1}{h_N}\right)\right) + \sum_{\substack{1 \leq i, j \leq N \\ i \neq j}} \text{Cov}\left(K_d\left(\frac{x - X_i}{h_N}\right), K_d\left(\frac{x - X_j}{h_N}\right)\right) \\ &=: V + S. \end{aligned}$$

Calculons maintenant le terme de variance V .

$$\begin{aligned} V &= N \mathbb{E}\left(K_d^2\left(\frac{x - X_1}{h_N}\right)\right) - N \mathbb{E}^2\left(K_d\left(\frac{x - X_1}{h_N}\right)\right) \\ &= N \int_{\mathbb{R}^d} K_d^2\left(\frac{x - u}{h_N}\right) v(u) du - N \left(\int_{\mathbb{R}^d} K_d\left(\frac{x - u}{h_N}\right) v(u) du \right)^2 \\ &= Nh_N^d \int_{\mathbb{R}^d} K_d^2(z) v(x - h_N z) dz - Nh_N^{2d} \left(\int_{\mathbb{R}^d} K_d(z) v(x - h_N z) dz \right)^2 \quad (\text{en posant } z = \frac{x-u}{h_N}) \\ &:= V_1 - V_2 \end{aligned}$$

Sous les hypothèses **(H2.2)(iii)** et **(H2.4)**, en faisant un développement de Taylor de v au voisinage de x , on obtient

$$\begin{aligned} V_1 &= Nh_N^d \left\{ \int_{\mathbb{R}^d} K_d^2(z) \left[v(x) - h_N \left(z_1 \frac{\partial v}{\partial z_1}(x^*) + \dots + z_d \frac{\partial v}{\partial z_d}(x^*) \right) \right] dz \right\} \quad (\text{tel que } x - h_N z \leq x^* \leq x) \\ &\leq Nh_N^d \left\{ v(x) \int_{\mathbb{R}^d} K_d^2(z) dz - h_N \left[\sup_{x \in D} \left| \frac{\partial v}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d^2(z) dz + \dots + \sup_{x \in D} \left| \frac{\partial v}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d^2(z) dz \right] \right\} \\ &= O(Nh_N^d). \end{aligned}$$

De la même manière, sous les hypothèses **(H2.2)(i)**, **(H2.2)(iii)** et **(H2.4)**, on obtient

$$\begin{aligned} V_2 &= Nh_N^{2d} \left\{ \int_{\mathbb{R}^d} K_d(z) \left[v(x) - h_N \left(z_1 \frac{\partial v}{\partial z_1}(x^*) + \dots + z_d \frac{\partial v}{\partial z_d}(x^*) \right) \right] dz \right\}^2 \quad (\text{tel que } x - h_N z \leq x^* \leq x) \\ &\leq Nh_N^{2d} \left\{ v(x) - h_N \left[\sup_{x \in D} \left| \frac{\partial v}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d(z) dz + \dots + \sup_{x \in D} \left| \frac{\partial v}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d(z) dz \right] \right\}^2 \\ &= O(Nh_N^{2d}). \end{aligned}$$

Par conséquent,

$$V = V_1 - V_2 = O(Nh_N^d).$$

Quant au terme de covariance, en faisant un changement de variable, les hypothèses **(H2.2)(i)**, **(H2.4)** et **(H2.5)** entraînent

$$\begin{aligned} \left| Cov \left(K_d \left(\frac{x - X_i}{h_N} \right), K_d \left(\frac{x - X_j}{h_N} \right) \right) \right| &= \left| \mathbb{E} \left[K_d \left(\frac{x - X_i}{h_N} \right) K_d \left(\frac{x - X_j}{h_N} \right) \right] - \mathbb{E} \left[K_d \left(\frac{x - X_i}{h_N} \right) \right] \right. \\ &\quad \left. \mathbb{E} \left[K_d \left(\frac{x - X_j}{h_N} \right) \right] \right| \\ &= \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - r}{h_N} \right) K_d \left(\frac{x - s}{h_N} \right) v_{i,j}(r, s) dr ds \right. \\ &\quad \left. - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - r}{h_N} \right) K_d \left(\frac{x - s}{h_N} \right) v(r) v(s) dr ds \right| \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - r}{h_N} \right) K_d \left(\frac{x - s}{h_N} \right) |v_{i,j}(r, s) - v(r) v(s)| dr ds \\ &= h_N^{2d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_d(u) K_d(w) |v_{i,j}(x - h_N u, x - h_N w) - \\ &\quad v(x - h_N u) v(x - h_N w)| du dw \\ &\leq Ch_N^{2d}. \end{aligned} \tag{2.22}$$

Notons aussi que ces covariances peuvent être contrôlées en utilisant le Théorème 2.1.

$$\begin{aligned} \left| Cov \left(K_d \left(\frac{x - X_i}{h_N} \right), K_d \left(\frac{x - X_j}{h_N} \right) \right) \right| &\leq \frac{(Lip(K_d))^2}{h_N^2} \Delta_{ij} \\ &\leq Ch_N^{-2} \rho(|i - j|). \end{aligned} \tag{2.23}$$

En écrivant $Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right) = Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)^{\frac{d}{2d+2}}$
 $\cdot Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)^{\frac{d+1}{2d+2}}$ et en utilisant (2.22) et (2.23), nous obtenons

$$\left|Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)\right| \leq Ch_N^d \rho(|i-j|)^{\frac{d}{2d+2}}. \quad (2.24)$$

Pour évaluer S , nous utilisons une technique développée dans Masry (1986). Pour cela, on considère une suite ϑ_N qui tend vers l'infini, dont on précisera ultérieurement sa valeur exacte. On a

$$S \leq \sum_{1 \leq |i-j| \leq \vartheta_N} \left|Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)\right| + \sum_{\vartheta_N < |i-j| \leq N-1} \left|Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), \left(\frac{x-X_j}{h_N}\right)\right)\right|. \quad (2.25)$$

En utilisant (2.22), le premier terme de covariance de (2.25) devient

$$\sum_{1 \leq |i-j| \leq \vartheta_N} \left|Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)\right| \leq CNh_N^{2d} \vartheta_N. \quad (2.26)$$

Pour le second terme, en utilisant (2.24) conjointement avec l'hypothèse **(H2.3)**, nous obtenons

$$\begin{aligned} \sum_{\vartheta_N < |i-j| \leq N-1} \left|Cov\left(K_d\left(\frac{x-X_i}{h_N}\right), K_d\left(\frac{x-X_j}{h_N}\right)\right)\right| &\leq \sum_{\vartheta_N < |i-j| \leq N-1} CNh_N^d e^{-\frac{\gamma^d}{2d+2}|i-j|} \\ &\leq CNh_N^d \int_{\vartheta_N}^N e^{-\frac{\gamma^d}{2d+2}u} du \\ &\leq CNh_N^d e^{-\frac{\gamma^d}{2d+2}\vartheta_N}. \end{aligned} \quad (2.27)$$

En choisissant $\vartheta_N = O(h_N^{-\nu})$ avec $0 < \nu < d$, dans (2.26) et (2.27), nous obtenons $S = o(Nh_N^d)$ et donc $\sigma_N^2 = O(Nh_N^d)$. Revenons maintenant à l'inégalité (2.21) et posons $A_N = \sigma_N^2$. Selon la preuve du Lemme 2.4, nous avons $\mu = 1$, $\lambda = 0$ et $L_1 = L_2 = \frac{1}{1-e^{-\gamma^d/(2d+2)}}$, ce qui entraîne que $B_n = O(1)$. Alors, l'inégalité (2.21) devient

$$\mathbb{P}\left(\sum_{i=1}^N U_i(x) \geq \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2/2}{C.Nh_N^d + \varepsilon^{5/3}}\right). \quad (2.28)$$

Ensuite, si on choisit $x = x_k$ et $\varepsilon = \varepsilon_0 \sqrt{\frac{\log N}{Nh_N^d}}$, pour tout $\varepsilon_0 > 0$, nous obtenons sous l'hypothèse **(H2.1)**

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leq k \leq l_N} \left|\sum_{i=1}^N U_i(x_k)\right| > Nh_N^d \varepsilon_0 \sqrt{\frac{\log N}{Nh_N^d}}\right) &\leq \sum_{k=1}^{l_N} \mathbb{P}\left(\left|\sum_{i=1}^N U_i(x_k)\right| > Nh_N^d \varepsilon_0 \sqrt{\frac{\log N}{Nh_N^d}}\right) \\ &\leq 2l_N \exp\left(-\frac{\frac{\varepsilon_0^2}{2} \log N}{C + \varepsilon_0^{\frac{5}{3}} \left(\frac{\log^5 N}{Nh_N^d}\right)^{\frac{1}{6}}}\right) \\ &\leq C(a_N^d)^{-1} e^{-C\varepsilon_0^2 \log N} \\ &= CN^{\frac{1}{2\beta}} h_N^{-1-\frac{d}{2\beta}} N^{-C\varepsilon_0^2} \\ &= C \frac{1}{\sqrt{(Nh_N)^{\frac{2\beta+d}{\beta}}}} N^{-C\varepsilon_0^2+1+\frac{d+1}{2\beta}}. \end{aligned}$$

Pour un choix approprié de ε_0 (par exemple $\varepsilon_0^2 > \frac{1}{C} \left(2 + \frac{d+1}{\beta}\right)$), cette dernière quantité peut être considérée comme le terme général d'une série convergente. Ainsi

$$\sum_{N=1}^{\infty} \mathbb{P} \left(\sum_{i=1}^N U_i(x_k) \geq \varepsilon \right) < \infty.$$

Par le lemme de Borel Cantelli

$$J_{12} = \max_{1 \leq k \leq l_N} \frac{1}{N h_N^d} \left| \sum_{i=1}^N U_i(x_k) \right| = O \left(\sqrt{\frac{\log N}{N h_N^d}} \right) \text{ p.s., lorsque } N \rightarrow \infty. \quad (2.29)$$

En combinant (2.19), (2.20) et (2.29), on obtient

$$J_1 = \sup_{x \in D} |\hat{v}_N(x) - v(x)| = O \left(\sqrt{\frac{\log N}{N h_N^d}} \right) \text{ p.s., lorsque } N \rightarrow \infty.$$

Etudions maintenant le terme de biais J_2 . D'après **(H2.2)(iii)**, **(H2.4)** et un développement de Taylor de v au voisinage de x , on a

$$\begin{aligned} \mathbb{E}(\hat{v}_N(x)) - v(x) &= \mathbb{E} \left[\frac{1}{N h_N^d} \sum_{i=1}^N K_d \left(\frac{x - X_i}{h_N} \right) \right] - v(x) \\ &= \frac{1}{h_N^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_N} \right) v(u) du - v(x) \\ &= \int_{\mathbb{R}^d} K_d(z) [v(x - h_N z) - v(x)] dz \quad (\text{en posant } z = \frac{x-u}{h_N}) \\ &= \int_{\mathbb{R}^d} K_d(z) \left[-h_N \left(z_1 \frac{\partial v}{\partial z_1}(x) + \dots + z_d \frac{\partial v}{\partial z_d}(x) \right) + \frac{h_N^2}{2} \left(z_1^2 \frac{\partial^2 v}{\partial z_1^2}(x^*) + \dots + z_d^2 \frac{\partial^2 v}{\partial z_d^2}(x^*) \right. \right. \\ &\quad \left. \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} z_i z_j \frac{\partial^2 v}{\partial z_i \partial z_j}(x^*) \right) \right] dz \quad (\text{tel que } x - h_N z \leq x^* \leq x) \\ &\leq -h_N \left[\sup_{x \in D} \left| \frac{\partial v}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d(z) dz + \dots + \sup_{x \in D} \left| \frac{\partial v}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d(z) dz \right] \\ &\quad + \frac{h_N^2}{2} \left[\sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_1^2}(x) \right| \int_{\mathbb{R}^d} z_1^2 K_d(z) dz + \dots + \sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_d^2}(x) \right| \int_{\mathbb{R}^d} z_d^2 K_d(z) dz \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_i \partial z_j}(x) \right| \int_{\mathbb{R}^d} z_i z_j K_d(z) dz \right] \\ &= \frac{h_N^2}{2} \left[\sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_1^2}(x) \right| \int_{\mathbb{R}^d} z_1^2 K_d(z) dz + \dots + \sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_d^2}(x) \right| \int_{\mathbb{R}^d} z_d^2 K_d(z) dz \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \sup_{x \in D} \left| \frac{\partial^2 v}{\partial z_i \partial z_j}(x) \right| \int_{\mathbb{R}^d} z_i z_j K_d(z) dz \right] \\ &\leq C h_N^2. \end{aligned}$$

On en déduit que

$$J_2 = \sup_{x \in D} |\mathbb{E}(\hat{v}_N(x)) - v(x)| = O(h_N^2).$$

Remplaçant J_1 et J_2 dans (2.18), nous obtenons le résultat donné dans le Théorème 2.3. \square

2.2.2 Normalité asymptotique

Roussas (2000) a établi la normalité asymptotique de l'estimateur à noyau \hat{v}_N de la densité v pour des variables strictement stationnaires associées ou négativement associées à valeurs dans \mathbb{R} . Nous étendons ce résultat au cas des variables aléatoires associées strictement stationnaires à valeurs dans \mathbb{R}^d . Pour cela, nous avons besoin des hypothèses **(H2.3)**, **(H2.4)** et **(H2.5)** précédentes et des hypothèses suivantes :

(H'2.1) La fenêtre h_N satisfie : $Nh_N^{d+4} \rightarrow 0$ quand $N \rightarrow +\infty$;

(H'2.2) La fonction noyau satisfait

(i) K_d est une densité de probabilité bornée à support compact ;

(ii) Pour tout $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $K_d(x) = \prod_{i=1}^d K(x_i)$, avec K est une fonction noyau lipschitzienne définie sur \mathbb{R} ;

(iii) $\int_{\mathbb{R}^d} z_i K_d(z) dz = 0, i = 1, \dots, d$ et $\int_{\mathbb{R}^d} |z_1^{i_1} \cdots z_d^{i_d}| K_d(z) dz < +\infty, i_1 + \dots + i_d = 2$;

(iv) Les dérivées partielles $\frac{\partial K_d}{\partial u_i}(u)$ existent pour tout $u \in \mathbb{R}^d$ et $i = 1, \dots, d$ et elles sont bornées ;

(H2.6) Il existe des suites de nombres entiers $(p_N)_N$, $(q_N)_N$ et $(k_N)_N$ définie par $k_N := \left\lfloor \frac{N}{p_N + q_N} \right\rfloor$, (où $[x]$ désigne la partie entière de x) tendant vers ∞ quand N tend vers ∞ , telles que $k_N(p_N + q_N) \leq N$ et $\frac{k_N(p_N + q_N)}{N} \xrightarrow{N \rightarrow +\infty} 1$ satisfaisant :

(i) $\frac{p_N k_N}{N} \xrightarrow{N \rightarrow +\infty} 1$,

(ii) $p_N h_N^d \xrightarrow{N \rightarrow +\infty} 0$ et $\frac{p_N^2}{N h_N^d} \xrightarrow{N \rightarrow +\infty} 0$,

(iii) $\frac{e^{-\gamma q_N}}{h_N^{d+2}} \xrightarrow{N \rightarrow +\infty} 0$.

Théorème 2.4. *Sous les hypothèses **(H'2.2)** et **(H2.3)**-**(H2.6)**, on a*

$$\sqrt{N h_N^d} (\hat{v}_N(x) - \mathbb{E}(\hat{v}_N(x))) \xrightarrow[N \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)),$$

avec

$$\sigma^2(x) = v(x) \int_{\mathbb{R}^d} K_d^2(z) dz,$$

et $\xrightarrow{\mathcal{L}}$ désigne la convergence en loi.

Corollaire 2.1. *Sous les hypothèses du Théorème 2.4 et si de plus, l'hypothèse **(H'2.1)** est vérifiée, on a :*

$$\sqrt{Nh_N^d} (\hat{v}_N(x) - v(x)) \xrightarrow[N \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)).$$

Les preuves de Théorème 2.4 et Corollaire 2.1 sont similaires dans leurs grandes lignes à celles du Théorème 2.1 et Corollaire 2.1 dans Roussas (2000).

2.3 Estimation de la fonction de régression

Soit $\{(X_i, Y_i); 1 \leq i \leq N\}$ une suite strictement stationnaire de vecteurs aléatoires associés définis sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d \times \mathbb{R}$ ($d \geq 1$), ayant la même distribution que (X, Y) qui admet une densité jointe inconnue $f(., .)$. Nous définissons alors la fonction de régression ou espérance conditionnelle de Y sachant $X = x$, par

$$m(x) := \mathbb{E}[Y|X = x] = \frac{\int_{\mathbb{R}} y f(x, y) dy}{\int_{\mathbb{R}} f(x, y) dy} =: \frac{\psi(x)}{v(x)},$$

pour tout $x \in \mathbb{R}^d$ lorsque $v(x)$ est différente de zéro. Pour estimer la fonction de régression m , nous considérons l'estimateur \hat{m}_N de Nadaraya-Watson (1964) défini, pour tout $x \in \mathbb{R}^d$, par

$$\hat{m}_N(x) = \sum_{i=1}^N W_{i,N}(x) Y_i, \quad (2.30)$$

où $W_{i,N}(x)$ sont des poids dépendants de x et de X_1, \dots, X_N définis par

$$\begin{aligned} W_{i,N}(x) &= \frac{K_d\left(\frac{x-X_i}{h_N}\right)}{\sum_{j=1}^N K_d\left(\frac{x-X_j}{h_N}\right)} \\ &= \frac{\frac{1}{Nh_N^d} K_d\left(\frac{x-X_i}{h_N}\right)}{\hat{v}_N(x)}, \end{aligned}$$

où $\hat{v}_N(x)$ est l'estimateur à noyau de Parzen-Rosenblatt donné dans (2.6). Si on note

$$\hat{\psi}_N(x) := \frac{1}{Nh_N^d} \sum_{i=1}^N Y_i K_d\left(\frac{x-X_i}{h_N}\right),$$

alors

$$\hat{m}_N(x) = \frac{\hat{\psi}_N(x)}{\hat{v}_N(x)}.$$

A notre connaissance, le seul résultat qui existe dans la littérature sur la convergence presque sûre de l'estimateur \hat{m}_N pour des données associées est établi dans le livre de Oliveira (2012) (Chapitre

3, Théorème 3.38) dans le cas univarié, sans vitesse de convergence. Cependant, il existe un résultat concernant la convergence presque sûre uniforme sur un compact de l'estimateur \hat{m}_N qui est établi par Ango Nze, Bühlmann et Doukhan (2002) pour des variables faiblement dépendantes qui incluent les variables associées dans le cas univarié. Ils ont ainsi obtenu, sous des conditions de régularité d'ordre $\rho \in]0, \infty[$ de ψ et v , une vitesse de convergence de l'ordre de $O\left(\frac{\log^4 N}{N}\right)^{\rho/(2\rho+1)}$. Nous obtenons dans le Théorème 2.3 une vitesse de convergence presque sûre uniforme dans le cas multivarié. Pour cela nous utilisons l'inégalité exponentielle établie par Doukhan et Neumann (2007) pour des variables aléatoires centrées donnée dans le Théorème 2.2.

2.3.1 Hypothèses et résultat

Soit \mathcal{U} un sous ensemble compact de $\dot{\mathcal{U}} = \{x \in \mathbb{R}^d | v(x) > \delta > 0\}$ pour un réel $\delta > 0$. Notons $\theta_{i,j}$ le coefficient de covariance défini par

$$\theta_{i,j} := \sum_{k=1}^d \sum_{l=1}^d \text{Cov}(X_{i,k}, X_{j,l}) + 2 \sum_{k=1}^d \text{Cov}(X_{i,k}, Y_j) + \text{Cov}(Y_i, Y_j).$$

Nous avons besoin des hypothèses **(H2.1)**, **(H2.2)**, **(H2.4)** et **(H2.5)** précédentes et des hypothèses suivantes :

(H'2.3) Le terme de covariance défini par $\rho(s) := \sup_{|i-j| \geq s} \theta_{i,j}$, $s > 0$, vérifie $\rho(s) \leq \gamma_0 e^{-\gamma s}$, $\gamma_0 > 0$, $\gamma > 0$;

(H2.7) La fonction $\psi(\cdot)$ est bornée et deux fois différentiable avec $\sup_{x \in \mathcal{U}} \left| \frac{\partial^k \psi}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ pour $i, j = 1, \dots, d$ et $k = 1, 2$;

(H2.8) La fonction $\psi_1(x) := \int_{\mathbb{R}} y^2 f(x, y) dy$ est bornée et différentiable avec $\sup_{x \in \mathcal{U}} \left| \frac{\partial \psi_1}{\partial x_i}(x) \right| < +\infty$ pour $i = 1, \dots, d$.

Le résultat est donné dans le théorème suivant

Théorème 2.5. *Supposons que les hypothèses **(H2.1)**-**(H2.2)**, **(H'2.3)** et **(H2.4)**-**(H2.7)** soient vérifiées, nous avons*

- i) $\sup_{x \in \mathcal{U}} \left| \hat{\psi}_N(x) - \psi(x) \right| = O\left(\sqrt{\frac{\log N}{N h_N^d}} \vee h_N^2\right)$ p.s, lorsque $N \rightarrow +\infty$,
- ii) $\sup_{x \in \mathcal{U}} |\hat{m}_N(x) - m(x)| = O\left(\sqrt{\frac{\log N}{N h_N^d}} \vee h_N^2\right)$ p.s, lorsque $N \rightarrow +\infty$.

En particulier, si $h_N = O\left(\left(\frac{\log N}{N}\right)^{1/(4+d)}\right)$, on obtient la la vitesse de convergence optimale $O\left(\left(\frac{\log N}{N}\right)^{2/(4+d)}\right)$, optimalité au sens du critère de convergence uniforme de l'estimation de la densité (Stute (1982)).

2.3.2 Preuve du résultat

Preuve du Théorème 2.5. On a

$$\begin{aligned} \sup_{x \in \mathcal{U}} \left| \hat{\psi}_N(x) - \psi(x) \right| &= \sup_{x \in \mathcal{U}} \left| \hat{\psi}_N(x) - \mathbb{E}(\hat{\psi}_N(x)) + \mathbb{E}(\hat{\psi}_N(x)) - \psi(x) \right| \\ &\leq \sup_{x \in \mathcal{U}} \left| \hat{\psi}_N(x) - \mathbb{E}(\hat{\psi}_N(x)) \right| + \sup_{x \in \mathcal{U}} \left| \mathbb{E}(\hat{\psi}_N(x)) - \psi(x) \right|. \end{aligned} \quad (2.31)$$

Sous les hypothèses **(H2.2)(iii)** et **(H2.7)**, nous avons

$$\begin{aligned} \mathbb{E} \left(\hat{\psi}_N(x) \right) - \psi(x) &= \mathbb{E} \left[\frac{1}{N h_N^d} \sum_{i=1}^N Y_i K_d \left(\frac{x - X_i}{h_N} \right) \right] - \psi(x) \\ &= \frac{1}{h_N^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_N} \right) f(u, y) du dy - \psi(x) \\ &= \frac{1}{h_N^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_N} \right) \psi(u) du - \psi(x) \\ &= \int_{\mathbb{R}^d} K_d(z) [\psi(x - h_N z) - \psi(x)] dz \quad (\text{en posant } z = \frac{x-u}{h_N}) \\ &= \int_{\mathbb{R}^d} K_d(z) \left[-h_N \left(z_1 \frac{\partial \psi}{\partial z_1}(x) + \dots + z_d \frac{\partial \psi}{\partial z_d}(x) \right) + \frac{h_N^2}{2} \left(z_1^2 \frac{\partial^2 \psi}{\partial z_1^2}(x^*) + \dots + z_d^2 \frac{\partial^2 \psi}{\partial z_d^2}(x^*) \right) \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} z_i z_j \frac{\partial^2 \psi}{\partial z_i \partial z_j}(x^*) \right] dz \quad (\text{tel que } x - h_N z \leq x^* \leq x) \\ &\leq \frac{h_N^2}{2} \left[\sup_{x \in \mathcal{U}} \left| \frac{\partial^2 \psi}{\partial z_1^2}(x) \right| \int_{\mathbb{R}^d} z_1^2 K_d(z) dz + \dots + \sup_{x \in \mathcal{U}} \left| \frac{\partial^2 \psi}{\partial z_d^2}(x) \right| \int_{\mathbb{R}^d} z_d^2 K_d(z) dz \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \sup_{x \in \mathcal{U}} \left| \frac{\partial^2 \psi}{\partial z_i \partial z_j}(x) \right| \int_{\mathbb{R}^d} z_i z_j K_d(z) dz \right] \\ &\leq C h_N^2. \end{aligned}$$

Ce qui donne

$$\sup_{x \in \mathcal{U}} \left| \mathbb{E}(\hat{\psi}_N(x)) - \psi(x) \right| = O(h_N^2). \quad (2.32)$$

Pour le premier terme de (2.31), nous avons

$$\hat{\psi}_N(x) - \mathbb{E}(\hat{\psi}_N(x)) = \frac{1}{N h_N^d} \sum_{i=1}^N V_i(x),$$

où

$$V_i(x) := V(X_i, Y_i; x) = Y_i K_d \left(\frac{x - X_i}{h_N} \right) - \mathbb{E} \left(Y_i K_d \left(\frac{x - X_i}{h_N} \right) \right).$$

En suivant pas à pas la preuve du Théorème 2.3 et remplaçant $U_i(x)$ par $V_i(x)$, on montre que

$$\sup_{x \in \mathcal{U}} \left| \hat{\psi}_N(x) - \mathbb{E}(\hat{\psi}_N(x)) \right| = O \left(\sqrt{\frac{\log N}{N h_N^d}} \right). \quad (2.33)$$

La preuve de i) s'achève en combinant (2.32) et (2.33). La preuve de ii) repose sur la décomposition suivante

$$\begin{aligned} \hat{m}_N(x) - m(x) &= \frac{\hat{\psi}_N(x)}{\hat{v}_N(x)} - \frac{\psi(x)}{v(x)} \\ &= \frac{1}{\hat{v}_N(x)} \left\{ \hat{\psi}_N(x) - m(x) \hat{v}_N(x) \right\} \\ &= \frac{1}{\hat{v}_N(x)} \left\{ (\hat{\psi}_N(x) - \psi(x)) + \frac{\psi(x)}{v(x)} (v(x) - \hat{v}_N(x)) \right\}. \end{aligned}$$

Ainsi, nous avons

$$\begin{aligned} \sup_{x \in \mathcal{U}} |\hat{m}_N(x) - m(x)| &\leq \frac{1}{\inf_{x \in \mathcal{U}} (\hat{v}_N(x))} \left\{ \sup_{x \in \mathcal{U}} |\hat{\psi}_N(x) - \psi(x)| + \frac{\sup_{x \in \mathcal{U}} |\psi(x)|}{\inf_{x \in \mathcal{U}} v(x)} \sup_{x \in \mathcal{U}} |\hat{v}_N(x) - v(x)| \right\} \\ &\leq \frac{1}{\delta - \sup_{x \in \mathcal{U}} |\hat{v}_N(x) - v(x)|} \left\{ \sup_{x \in \mathcal{U}} |\hat{\psi}_N(x) - \psi(x)| + \delta^{-1} \sup_{x \in \mathcal{U}} |\psi(x)| \sup_{x \in \mathcal{U}} |\hat{v}_N(x) - v(x)| \right\}. \end{aligned}$$

En utilisant le résultat i) et le Théorème 2.3, on obtient ii). □

Chapitre 3

Estimation de la fonction de régression pour des données tronquées à gauche

Dans ce chapitre, nous présentons premièrement le modèle aléatoire de troncature à gauche (RLT) et les estimateurs à noyau de la densité de la covariable et de la fonction de régression introduits par Ould-Saïd et Lemdani (2006) dans lesquels est pris en compte l'effet de troncature des données. Ensuite, nous étendons les résultats de Ould Saïd et Lemdani (2006) sur la convergence uniforme presque sûre et la normalité asymptotique de ces estimateurs obtenus dans le cas des suites de variables aléatoires indépendantes identiquement distribuées (i.i.d.) avec une covariable réelle au cas d'une covariable à valeurs dans $\mathbb{R}^d (d \geq 1)$. Nous concluons ce chapitre par les résultats de Liang, Li et Qi (2009) et Liang (2011) où ils étendent les résultats obtenus dans le cas indépendant au cas des suites de variables aléatoires α -mélangeantes.

3.1 Estimation sous le modèle aléatoire de troncature à gauche

3.1.1 Modèle aléatoire de troncature à gauche (RLT)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

Soit Y une variable aléatoire réelle (durée de vie) de fonction de répartition (f.d.r.) F inconnue et T une variable aléatoire réelle indépendante de Y de fonction de répartition G inconnue. Considérons une suite de variables aléatoires i.i.d. $\{Y_i; i = 1, \dots, N\}$ de la variable Y . Soit $\{T_i; i = 1, \dots, N\}$ une suite de variables aléatoires i.i.d. de la variable T . On suppose aussi que ces variables sont indépendantes des Y_i .

Ainsi la f.d.r. jointe de Y et T est

$$\begin{aligned} H(y, t) &:= \mathbb{P}(Y \leq y, T \leq t) \\ &= F(y)G(t). \end{aligned}$$

Dans le modèle RLT, la variable aléatoire d'intérêt Y est interférée par une variable de troncature T , tel que les quantités Y et T sont observables seulement si $Y \geq T$ tandis que rien n'est observé si $Y < T$. Sans confusion possible, on note $\{(Y_i, T_i); i = 1, \dots, n\}$, ($n \leq N$) l'échantillon observé (i.e. $Y_i \geq T_i$) parmi le N -échantillon d'origine. La taille N est donc déterministe mais inconnue. Une conséquence de la troncature est que la taille de l'échantillon vraiment observé n est une variable aléatoire distribuée selon la loi Binomiale de paramètre N et α où α est la probabilité de (non) troncature définie par

$$\alpha := \mathbb{P}(Y \geq T).$$

Il est clair que si $\alpha = 0$, aucune donnée peut être observée. Pour cela, nous supposons, dorénavant, que $\alpha \neq 0$. Par la loi forte des grands nombres on a, lorsque N tend vers ∞

$$\check{\alpha}_n := \frac{n}{N} \rightarrow \alpha, \quad \mathbb{P} - p.s. \quad (3.1)$$

Conditionnellement à la valeur de n , Lemdani et Ould-Said (2007) ont prouvé que la propriété i.i.d de l'échantillon observé de taille n est déduite de celle de l'échantillon de taille N . Comme N est inconnue et n est connu (bien qu'il soit aléatoire), sous ce modèle, les résultats ne seront pas établis par rapport à la probabilité \mathbb{P} (relative au N -échantillon) mais par rapport à une nouvelle probabilité \mathbf{P} (relative au n -échantillon) définit par

$$\mathbf{P}(\cdot) = \mathbb{P}(\cdot | Y \geq T).$$

Nous noterons respectivement \mathbb{E} et \mathbf{E} les espérances relatives aux probabilités \mathbb{P} and \mathbf{P} . Dans la suite, nous dénotons par un exposant $(*)$, toute caractéristique liée aux données observées (i.e. conditionnellement à la valeur de n).

3.1.2 Estimation de les fonctions de répartitions

Sous le modèle RLT, la f.d.r. jointe (Stute (1993) et Zhou (1996)) d'un (Y, T) observé devient

$$\begin{aligned} H^*(y, t) &:= \mathbf{P}(Y \leq y, T \leq t) \\ &= \mathbb{P}(Y \leq y, T \leq t | Y \geq T) \\ &= \frac{\mathbb{P}(Y \leq y, T \leq t, Y \geq T)}{\mathbb{P}(Y \geq T)} \\ &= \frac{\mathbb{P}(Y \leq y, T \leq t, T \leq y)}{\mathbb{P}(Y \geq T)} \\ &= \frac{1}{\alpha} \int_{-\infty}^y G(t \wedge u) dF(u). \end{aligned}$$

où $t \wedge u := \min(t, u)$. Les f.d.r.'s marginales respectives de Y et T sont donc définies par

$$\begin{aligned} F^*(y) &:= H^*(y, \infty) \\ &= \frac{1}{\alpha} \int_{-\infty}^y G(u) dF(u) \end{aligned}$$

et

$$\begin{aligned}
G^*(t) &:= H^*(\infty, t) \\
&= \frac{1}{\alpha} \int_{-\infty}^{\infty} G(t \wedge u) dF(u) \\
&= \frac{1}{\alpha} \int_{-\infty}^{\infty} \int_{-\infty}^{t \wedge u} dG(v) dF(u) \\
&= \frac{1}{\alpha} \int_{-\infty}^t dG(v) \int_v^{\infty} dF(u) \\
&= \frac{1}{\alpha} \int_{-\infty}^t (1 - F(v)) dG(v),
\end{aligned}$$

qui peuvent être estimés respectivement par

$$F_n^*(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq y\}}$$

et

$$G_n^*(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{T_i \leq t\}}.$$

F_n^* et G_n^* sont les estimateurs empiriques respectifs de F^* et G^* , mais ne sont en aucun cas des estimateurs de F et G . Pour celles-ci, Lynden-Bell (1971) a obtenu des estimateurs non paramétriques que nous introduirons par la suite. Pour toute f.d.r. W , nous noterons par a_W et b_W , respectivement, les bornes inférieure et supérieure du support de W définies respectivement par

$$a_W := \inf\{u : W(u) > 0\} \text{ et } b_W := \sup\{u : W(u) < 1\}. \quad (3.2)$$

Afin de nous assurer de l'identifiabilité du modèle, nous supposons, en suivant Woodroffe (1985),

$$a_G \leq a_F, \quad b_G \leq b_F \quad \text{et} \quad \int_{a_F}^{\infty} \frac{dF}{G} < \infty. \quad (3.3)$$

Il est à noter que la dernière condition est requise pour le cas $a_G = a_F$.

Sous les conditions (3.3),

$$\alpha = \mathbb{P}(Y \geq T) = \int_0^{\infty} G(u) dF(u). \quad (3.4)$$

Maintenant, pour définir les estimateurs de F et G , nous avons besoin d'introduire la fonction $C(\cdot)$ définie pour un $y \in [a_F; \infty)$ par

$$\begin{aligned}
C(y) &:= \mathbf{P}\{T \leq y \leq Y\} \\
&= G^*(y) - F^*(y^-) \\
&= \mathbb{P}(T \leq y \leq Y | Y \geq T) \\
&= \frac{\mathbb{P}(T \leq y \leq Y, Y \geq T)}{\mathbb{P}(Y \geq T)} \\
&= \frac{\mathbb{P}(T \leq y \leq Y)}{\mathbb{P}(Y \geq T)} \\
&= \frac{1}{\alpha} \mathbb{P}(T \leq y) \mathbb{P}(Y \geq y) \\
&= \frac{1}{\alpha} G(y) (1 - F(y^-)),
\end{aligned} \tag{3.5}$$

qui peut être estimée par

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

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

$$F_n(y) = 1 - \prod_{i: Y_i \leq y} \left[\frac{n C_n(Y_i) - 1}{n C_n(Y_i)} \right], \tag{3.6}$$

et

$$G_n(t) = \prod_{i: T_i > t} \left[\frac{n C_n(T_i) - 1}{n C_n(T_i)} \right]. \tag{3.7}$$

Les propriétés asymptotiques de ces estimateurs sont étudiées par Woodroffe (1985) et il a établi sous les conditions (3.3) et la continuité de F et G , les résultats de convergence presque sûre suivants

$$\sup_{y \geq a_F} |F_n(y) - F(y)| \rightarrow 0, \quad \sup_{t \geq a_G} |G_n(t) - G(t)| \rightarrow 0, \quad \mathbf{P} - p.s., \quad \text{quand } n \rightarrow \infty.$$

3.1.3 Estimation de la probabilité de troncature

Une autre quantité inconnue importante qu'on doit aussi estimer est la probabilité de troncature α car l'estimateur défini dans (3.1) ne peut pas être calculé puisque N est inconnu.

En utilisant (3.4), α peut être estimée par

$$\tilde{\alpha}_n = \int_0^\infty \tilde{G}_n(u) d\tilde{F}_n(u). \tag{3.8}$$

où \tilde{F}_n et \tilde{G}_n sont, respectivement, des estimateurs de F et G . Sous les conditions (3.3) et la continuité de F et G , Woodroffe (1985) a montré que si on remplace \tilde{F}_n et \tilde{G}_n par les estimateurs de Lynden-Bell définis, respectivement, en (3.6) et (3.7), alors $\tilde{\alpha}_n$ converge en probabilité vers α quand $n \rightarrow \infty$. Sous des conditions similaires, la normalité asymptotique de $\tilde{\alpha}_n$ a été étudiée notamment par Chao (1987) et Keiding et Gill (1990).

En utilisant (3.5), nous avons

$$\alpha = \frac{G(y)(1 - F(y^-))}{C(y)}.$$

Ainsi, He et Yang (1998) proposent d'utiliser cette relation pour définir un nouvel estimateur pour α donné, pour tout y tel que $C_n(y) > 0$, par

$$\alpha_n = \frac{G_n(y)(1 - F_n(y^-))}{C_n(y)} \quad (3.9)$$

où F_n et G_n sont les estimateurs de Lynden-Bell définis, respectivement, en (3.6) et (3.7) et C_n est l'estimateur empirique de la fonction C .

He et Yang (1998) montrent que α_n est indépendant de y pourvu que $C_n(y) > 0$ et ils montrent que

$$\alpha_n \rightarrow \alpha, \quad \mathbf{P} - p.s., \quad \text{quand } n \rightarrow \infty.$$

Ils montrent aussi que si dans (3.8) \tilde{F}_n et \tilde{G}_n sont les estimateurs de Lynden-Bell, alors α_n et $\tilde{\alpha}_n$ sont égaux.

3.1.4 Estimation de la densité de la covariable

Maintenant, en plus des deux variables considérées auparavant Y et T , nous introduisons un vecteur aléatoire $X \in \mathbb{R}^d$ ($d \geq 1$) de covariables de fonction de répartition V et admettant une densité continue v .

Soit $\{(X_i, Y_i, T_i); i = 1, \dots, N\}$ une suite i.i.d de vecteurs aléatoires du triplet (X, Y, T) . On note par $\{(X_i, Y_i, T_i); i = 1, \dots, n\}$ l'échantillon observé (i.e. $Y_i \geq T_i$). Dorénavant, on suppose que T est indépendante de (X, Y) et la notation $X \leq x$ sera utilisée pour les vecteurs $X = (X_1, \dots, X_d)$ et $x = (x_1, \dots, x_d)$ pour dénoter l'événement $\{X_1 \leq x_1, \dots, X_d \leq x_d\}$. Nous rappelons la construction des estimateurs de V et v établie par Ould Saïd et Lemdani (2006). Dans l'espace $(\Omega, \mathcal{F}, \mathbb{P})$, l'estimateur à noyau naturel de la densité de la covariable v est donné par

$$v_N(x) = \frac{1}{Nh_N^d} \sum_{i=1}^N K_d \left(\frac{x - X_i}{h_N} \right),$$

où $K_d : \mathbb{R}^d \rightarrow \mathbb{R}$ est la fonction noyau et $(h_N)_{N \geq 1}$ une suite de nombres réels positifs appelée fenêtre tendant vers zéro à l'infini. Comme N est inconnu et comme on dispose seulement de n observations de X (parmi N), on ne peut pas alors utiliser ce dernier estimateur. D'autre part, si nous considérons

l'estimateur

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

qui est calculé à partir de l'échantillon observé, la densité estimée sera la densité conditionnelle v^* . Pour surmonter cette difficulté, nous considérons d'abord la f.d.r. jointe L^* d'un triplet (X, Y, T) observé suivante

$$\begin{aligned} L^*(x, y, t) &= \mathbf{P}(X \leq x, Y \leq y, T \leq t) \\ &= \mathbb{P}(X \leq x, Y \leq y, T \leq t | Y \geq T) \\ &= \frac{\mathbb{P}(X \leq x, Y \leq y, T \leq t, Y \geq T)}{\mathbb{P}(Y \geq T)} \\ &= \frac{1}{\alpha} \int_{-\infty}^x \int_{a_G}^y G(w \wedge t) dF(u, w), \end{aligned}$$

où $F(.,.)$ est la f.d.r. jointe de couple (X, Y) . Prenant $t = +\infty$, un couple (X, Y) observé a donc la f.d.r. jointe $F^*(.,.)$ suivante

$$\begin{aligned} F^*(x, y) &= L^*(x, y, +\infty) \\ &= \alpha^{-1} \int_{-\infty}^x \int_{a_G}^y G(w) dF(u, w). \end{aligned} \quad (3.10)$$

Par différentiation de (3.10), on obtient

$$F(dx, dy) = \frac{\alpha}{G(y)} F^*(dx, dy), \text{ pour } y \geq a_G. \quad (3.11)$$

D'où

$$f(x, y) = \frac{\alpha}{G(y)} f^*(x, y) \quad (3.12)$$

avec $f(.,.)$ est la densité jointe de couple (X, Y) .

En intégrant (3.11) par rapport à y , on obtient la f.d.r. de X

$$\begin{aligned} V(x) &= \alpha \int_{-\infty}^x \int_{a_G}^{+\infty} \frac{1}{G(y)} F^*(du, dy) \\ &= \mathbf{E} \left(\frac{\alpha}{G(Y)} \mathbb{I}_{\{X \leq x\}} \right). \end{aligned}$$

Un estimateur naturel de V est donné par

$$\widehat{V}_n(x) := \frac{\alpha_n}{n} \sum_{i=1}^n \frac{1}{G_n(Y_i)} \mathbb{I}_{\{X_i \leq x\}}. \quad (3.13)$$

Notons que dans (3.13) la somme est sur i telle que $G_n(Y_i) \neq 0$. Finalement (3.13) donne un estimateur de la densité de X par le produit de convolution du noyau K_d et de l'estimateur \widehat{V}_n .

$$\begin{aligned} \hat{v}_n(x) &:= \frac{1}{h_n^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_n} \right) d\widehat{V}_n(u) \\ &= \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right). \end{aligned} \quad (3.14)$$

3.1.5 Estimation de la fonction de régression

Nous souhaitons estimer la fonction de régression définie, pour tout $x \in \mathbb{R}^d$, par

$$m(x) = \mathbb{E}[Y|X = x] \quad (3.15)$$

qui peut s'écrire sous la forme $m(x) = \frac{\psi(x)}{v(x)}$ avec $\psi(x) = \int_{\mathbb{R}} y f(x, y) dy$. En se basant sur un échantillon $\{(X_i, Y_i, T_i); i = 1, \dots, N\}$ du triplet (X, Y, T) , la fonction de régression qui peut être estimée est donnée par

$$m^*(x) = \mathbb{E}[Y|X = x, Y \geq T]. \quad (3.16)$$

Puisque les fonctions de distribution de Y et de T dans le n -échantillon observé sont différentes de leurs fonctions de distribution dans le N -échantillon original, alors $m^*(x) \neq m(x)$. Ainsi, un estimateur de $m^*(.)$ ne peut pas être utilisé comme un estimateur de $m(.)$.

Lemdani et Ould Saïd (2006) ont construit un nouvel estimateur à noyau de la fonction de régression sous un modèle RLT de la manière suivante.

Soit $\xi : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction mesurable et considérons l'estimation de l'espérance $\mathbb{E}[\xi(X, Y)]$ à partir de l'échantillon $\{(X_i, Y_i, T_i); i = 1, \dots, N\}$. Puisque seulement n triplets sont observés (parmi N), nous essayons de construire un estimateur de $\mathbb{E}[\xi(X, Y)]$ qui est calculable seulement si $Y \geq T$. Suivant l'idée introduite par Carbonez, Györfi et van der Meulin (1995), un estimateur sans biais sous le modèle RLT est donné par

$$\frac{1}{N} \sum_{i=1}^N \frac{\xi(X_i, Y_i)}{G(Y_i)} \mathbb{I}_{\{Y_i \geq T_i\}}. \quad (3.17)$$

En effet, en utilisant les propriétés de l'espérance conditionnelle et le fait que T et (X, Y) sont indépendants, nous avons

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \frac{\xi(X_i, Y_i)}{G(Y_i)} \mathbb{I}_{\{Y_i \geq T_i\}} \right] &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\mathbb{E} \left[\frac{\xi(X_i, Y_i)}{G(Y_i)} \mathbb{I}_{\{Y_i \geq T_i\}} | X_i, Y_i \right] \right] \\ &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\frac{\xi(X_i, Y_i)}{G(Y_i)} \mathbb{E} [\mathbb{I}_{\{Y_i \geq T_i\}} | X_i, Y_i] \right] \\ &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} [\xi(X_i, Y_i)] \\ &= \mathbb{E} [\xi(X, Y)]. \end{aligned}$$

Pour le cas particulier où $\xi(X, Y) = (Y|X)$, nous obtenons un estimateur de la fonction de régression $m(x) = \mathbb{E}[Y|X = x]$. Cependant, N et G sont inconnus, l'estimateur donné dans (3.17) ne peut pas être utilisé en pratique et il doit être adapté. Premièrement, G peut être estimée par l'estimateur consistant G_n donné par (3.7) et N par $\frac{n}{\alpha_n}$ où α_n est donné par (3.9). Il n'est non plus possible

d'utiliser les poids de Nadaraya-Watson définis par

$$\begin{aligned} W_{i,N}(x) &= \frac{K_d\left(\frac{x-X_i}{h_N}\right)}{\sum_{j=1}^N K_d\left(\frac{x-X_j}{h_N}\right)} \\ &= \frac{\frac{1}{Nh_N^d} K_d\left(\frac{x-X_i}{h_N}\right)}{\hat{v}_N(x)}, \end{aligned} \quad (3.18)$$

pour estimer $\mathbb{E}[Y|X=x]$.

En effet, si nous considérons les $W_{i,\frac{n}{\alpha_n}}(x)$, il n'est pas possible de calculer le dénominateur car on dispose seulement de n observations de X (parmi N). D'autre part, si nous considérons les poids

$$W_{i,n}(x) = \frac{K_d\left(\frac{x-X_i}{h_n}\right)}{\sum_{j=1}^n K_d\left(\frac{x-X_j}{h_n}\right)}$$

qui sont calculés à partir de l'échantillon observé, la fonction de régression estimée sera $m^*(x)$ ce qui n'est pas notre but. En utilisant (5.7), nous adaptons (3.18) pour obtenir des nouveaux poids

$$\bar{W}_{i,n}(x) = \frac{\frac{1}{G_n(Y_i)} K_d\left(\frac{x-X_i}{h_n}\right)}{\sum_{j=1}^n \frac{1}{G_n(Y_j)} K_d\left(\frac{x-X_j}{h_n}\right)}.$$

Ces poids mènent à un estimateur compliqué pour ψ et donc pour m , ainsi ils sont modifiés en

$$\widetilde{W}_{i,n}(x) = \frac{\frac{1}{\alpha_n} K_d\left(\frac{x-X_i}{h_n}\right)}{\sum_{j=1}^n \frac{1}{G_n(Y_j)} K_d\left(\frac{x-X_j}{h_n}\right)}$$

afin d'obtenir un estimateur plus simple et plus naturel $\hat{\psi}_n$ de ψ . Ceci est motivé par le fait que chacun de $\frac{1}{\alpha_n}$ et $\frac{1}{n} \sum_{i=1}^n \frac{1}{G_n(Y_i)}$ sont des estimateurs consistants de $\frac{1}{\alpha}$. Maintenant, l'estimateur donné dans (3.17) peut être écrit sous la forme

$$\frac{1}{n} \times \frac{n}{N} \sum_{i=1}^N \frac{\xi(X_i, Y_i)}{G(Y_i)}, \quad (3.19)$$

ce qui donne un nouvel estimateur de $m(x)$ en remplaçant les poids uniformes $\frac{1}{n}$ par $\widetilde{W}_{i,n}(x)$. Nous obtenons

$$\begin{aligned}\hat{m}_n(x) &:= \alpha_n \sum_{i=1}^n \widetilde{W}_{i,n}(x) \frac{Y_i}{G_n(Y_i)} \\ &= \frac{\sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)} \\ &= \frac{\hat{\psi}_n(x)}{\hat{v}_n(x)}\end{aligned}\tag{3.20}$$

avec

$$\hat{\psi}_n(x) = \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)\tag{3.21}$$

et $\hat{v}_n(x)$ est donné par (5.7). L'estimateur $\hat{m}_n(\cdot)$ peut être utilisé dans la pratique. Notons que $\hat{\psi}_n(\cdot)$ est un estimateur de la fonction $\psi(\cdot)$ qui peut être écrit comme suit

$$\hat{\psi}_n(x) = \frac{1}{h_n^d} \int_{\mathbb{R}} K_d\left(\frac{x - u}{h_n}\right) d\widehat{\Psi}_n(u)$$

avec

$$\widehat{\Psi}_n(x) = \frac{\alpha_n}{n} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} \mathbb{I}_{\{X_i \geq x\}}\tag{3.22}$$

qui est un estimateur de

$$\Psi(x) = \int_{-\infty}^x \psi(u) du.$$

3.2 Propriétés asymptotiques de l'estimateur à noyau de la fonction de régression sous le modèle RLT

Dans toute la suite nous supposons la continuité de les f.d.r.'s F et G ,

$$a_G < a_F \quad \text{et} \quad b_G \leq b_F.$$

3.2.1 Cas de suites de variables aléatoires i.i.d.

Lemdani et Ould Saïd (2006) ont établi la consistance uniforme forte de l'estimateur à noyau de la fonction de régression pour des variables aléatoires indépendantes tronquée à gauche dans le cas d'une covariable réelle. Ils ont établi également la normalité asymptotique de l'estimateur. Nous étendons ces résultats au cas où la covariable est à valeurs dans \mathbb{R}^d .

3.2.1.1 Hypothèses et résultats principaux

Soit $\Omega_0 = \{x \in \mathbb{R}^d | v(x) > 0\}$ et soit $\mathcal{C} \subset \Omega_0$ un sous ensemble compact et $\eta = \inf_{x \in \Omega} v(x) > 0$. Les hypothèses pour établir les résultats sont les suivantes :

(H3.1) La fenêtre h_n satisfie :

$$h_n \rightarrow 0 \text{ et } \frac{\log n}{nh_n^d} \rightarrow 0 \text{ quand } n \rightarrow +\infty;$$

(H3.2) Le noyau K_d est une densité de probabilité bornée, holderienne d'exposant β , satisfaisant

$$\int_{\mathbb{R}^d} z_i K_d(z) dz = 0, i = 1, \dots, d, \int_{\mathbb{R}^d} |z_1^{i_1} \dots z_d^{i_d}| K_d(z) dz < +\infty, i_1 + \dots + i_d = 2, \\ \int_{\mathbb{R}^d} K_d^2(z) dz < +\infty \text{ et } \int_{\mathbb{R}^d} |z_1 + \dots + z_d| K_d^2(z) dz < +\infty;$$

(H3.3) La fonction $\psi(\cdot)$ est bornée et différentiable avec $\sup_{x \in \mathcal{C}} \left| \frac{\partial^k \psi}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ pour $i, j = 1, \dots, d$ et $k = 1, 2$;

(H3.4) La fonction $\psi_1(x) := \int_{\mathbb{R}} \frac{y^2}{G(y)} f(x, y) dy$ est bornée et deux fois différentiable avec $\sup_{x \in \mathcal{U}} \left| \frac{\partial \psi_1}{\partial x_i}(x) \right| < +\infty$ pour $i = 1, \dots, d$;

(H3.5) La densité $v(\cdot)$ est bornée et deux fois différentiable avec $\sup_{x \in \mathcal{C}} \left| \frac{\partial^k v}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ pour $i, j = 1 \dots d$ et $k = 1, 2$.

Les hypothèses suivantes interviennent dans le résultat de la normalité asymptotique.

(H3.6) La fenêtre h_n satisfie

$$nh_n^{d+4} \rightarrow 0 \text{ quand } n \rightarrow +\infty;$$

(H3.7) Il existe $\nu > 2$ tel que

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}} y^\nu f(x, y) dx dy < +\infty.$$

Le premier résultat concernant la convergence uniforme presque sûre de l'estimateur \hat{m}_n défini en (4.9) sur un compact est donné dans le théorème suivant

Théorème 3.1. *Supposons que les hypothèses (H3.1)-(H3.5) sont vérifiées, nous avons*

$$\sup_{x \in \mathcal{U}} |\hat{m}_n(x) - m(x)| = O \left(\sqrt{\frac{\log n}{nh_n^d}} \vee h_n^2 \right) \text{ P-p.s, lorsque } n \rightarrow +\infty.$$

En particulier si nous choisissons $h_n = O \left(\left(\frac{\log n}{n} \right)^{1/(d+4)} \right)$, nous obtenons la vitesse de convergence optimale

$$\sup_{x \in \mathcal{U}} |\hat{m}_n(x) - m(x)| = O \left(\left(\frac{\log n}{n} \right)^{2/(d+4)} \right), \text{ p.s, lorsque } n \rightarrow +\infty,$$

optimalité au sens du critère de convergence uniforme de l'estimation de la densité (Stute 1982).

Le résultat suivant concerne la normalité asymptotique de l'estimateur \hat{m}_n . Soit

$$\Sigma(x) = \begin{pmatrix} \Sigma_0(x) & \Sigma_1(x) \\ \Sigma_1(x) & \Sigma_2(x) \end{pmatrix} \quad (3.23)$$

avec

$$\Sigma_j(x) = \int_{\mathbb{R}} \frac{y^{2-j}}{G(y)} f(x, y) dy, \quad j = 0, 1, 2. \quad (3.24)$$

Ainsi, nous avons le théorème suivant

Théorème 3.2. *Sous les hypothèses (H3.1)-(H3.3) et (H3.5)-(H3.7) et pour tout x tel que $v(x) > 0$, nous avons*

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)),$$

avec $\xrightarrow{\mathcal{L}}$ désigne la convergence en loi et

$$\sigma^2(x) := \frac{\alpha [\Sigma_0(x)v^2(x) - 2\Sigma_1(x)\psi(x)v(x) + \Sigma_2(x)\psi^2(x)]}{v^4(x)} \kappa, \quad (3.25)$$

avec

$$\kappa = \int_{\mathbb{R}^d} K_d^2(z) dz < +\infty.$$

Remarque 3.1. *Notons d'une part que $\Sigma_2(x) \geq v(x)$ et d'autre part, d'après l'inégalité de Cauchy-Schwartz, nous avons $\Sigma_1^2(x) < \Sigma_0(x)\Sigma_2(x)$. Ainsi, $\Sigma(x)$ est définie positive pourvu que $v(x) > 0$.*

3.2.1.2 Preuves des résultats

Dans la preuve du Théorème 3.1, nous utilisons une inégalité de type exponentiel, plus précisément l'inégalité de Bernstein et nous utilisons les techniques de recouvrement d'un ensemble compact par un nombre fini de boules pour obtenir l'uniformité. Ould Saïd et Lemdani (2006) utilisent une autre idée qui est basé sur l'utilisation des classes de Vapnik-Červonenkis (VC). Cependant, si nous suivons la même méthodologie, nous obtenons une vitesse de convergence de l'ordre de $O\left(\sqrt{\frac{\log n}{nh_n^{2d}}} \vee h_n^2\right)$ (vitesse obtenue pour $d = 1$ dans Ould Saïd et Lemdani (2006)), alors que celle obtenue en utilisant l'inégalité de Bernstein est de l'ordre de $O\left(\sqrt{\frac{\log n}{nh_n^d}} \vee h_n^2\right)$ qui est donc meilleure. Dans cette partie, C désigne une constante générique.

Preuve du Théorème 3.1. Afin de montrer ce théorème, nous considérons d'abord les estimateurs suivants

$$\tilde{\psi}_n(x) = \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)$$

et

$$\tilde{v}_n(x) = \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{1}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)$$

de $v(x)$ et $\psi(x)$, respectivement. Ces estimateurs ne peuvent pas être calculé car α et G sont inconnus mais seront utiles pour contrôler les termes de fluctuations et de biais . Nous considérons aussi les deux lemmes suivants.

Lemme 3.1. *Sous les hypothèses (H3.1)-(H3.5), nous avons*

$$\sup_{x \in \mathcal{C}} |\hat{\psi}_n(x) - \psi(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee h_n^2 \right\} \mathbf{P} - p.s, \text{ lorsque } n \rightarrow \infty.$$

Preuve. On utilise la décomposition suivante

$$\hat{\psi}_n(x) - \psi(x) = \left(\hat{\psi}_n(x) - \tilde{\psi}_n(x) \right) + \left(\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right) + \left(\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right),$$

ce qui donne

$$\begin{aligned} \sup_{x \in \mathcal{C}} |\hat{\psi}_n(x) - \psi(x)| &\leq \sup_{x \in \mathcal{C}} \left| \hat{\psi}_n(x) - \tilde{\psi}_n(x) \right| + \sup_{x \in \mathcal{C}} \left| \tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right| + \sup_{x \in \mathcal{C}} \left| \mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right| \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

Pour le terme I_1 , nous avons

$$\left| \hat{\psi}_n(x) - \tilde{\psi}_n(x) \right| \leq \left\{ \frac{|\alpha_n - \alpha|}{G_n(a_F)} + \frac{\alpha}{G_n(a_F)G(a_F)} \sup_{y \geq a_F} |G_n(y) - G(y)| \right\} \frac{1}{nh_n^d} \sum_{i=1}^n Y_i K_d \left(\frac{x - X_i}{h_n} \right).$$

D'après le Théorème 3.2 de He et Yang (1998),

$$|\alpha_n - \alpha| = O_{\mathbf{P}}(n^{-1/2}). \quad (3.26)$$

Et d'après la Remarque 6 de Woodroffe (1985),

$$\sup_{y \geq a_F} |G_n(y) - G(y)| = O_{\mathbf{P}}(n^{-1/2}). \quad (3.27)$$

Par la Loi Forte des Grands Nombres, lorsque $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K_d \left(\frac{x - X_i}{h_n} \right) \longrightarrow \mathbf{E} \left(\frac{1}{h_n^d} K_d \left(\frac{x - X_1}{h_n} \right) \right).$$

Ainsi, en utilisant ce dernier résultat et le fait que Y est bornée, on obtient sous les hypothèses (H3.2) et (H3.5)

$$\frac{1}{nh_n^d} \sum_{i=1}^n Y_i K_d \left(\frac{x - X_i}{h_n} \right) = O_{\mathbf{P}}(1). \quad (3.28)$$

Finalement, en utilisant (3.26), (3.27), (3.28) et le fait que $G_n(a_F) \xrightarrow{\mathbf{P}-p.s.} G(a_F)$, nous obtenons

$$I_1 = O_{\mathbf{P}}(n^{-1/2}). \quad (3.29)$$

Etudions maintenant le terme I_2 . La compacité de \mathcal{C} , nous permet d'utiliser une technique classique qui consiste à recouvrir le compact \mathcal{C} par un nombre fini l_n de boules $B_k(x_k, a_n^d)$ de centres $x_k = (x_{k,1}, \dots, x_{k,d})$, $k \in \{1, \dots, l_n\}$ et de même rayon a_n^d où l_n et a_n^d vérifient

$$l_n \leq C a_n^{-d} \quad \text{et} \quad a_n^d = n^{-\frac{1}{2\beta}} h_n^{1+\frac{d}{2\beta}}.$$

Pour tout x dans \mathcal{C} , il existe une boule B_k qui contient x tel que $\|x - x_k\| \leq a_n^d$, $\|\cdot\|$ est une norme définie sur \mathbb{R}^d . Nous décomposons alors le terme I_2 comme suit

$$\begin{aligned} I_2 &= \sup_{x \in \mathcal{C}} \left| \left(\tilde{\psi}_n(x) - \tilde{\psi}_n(x_k) \right) + \left(\mathbf{E}(\tilde{\psi}_n(x_k)) - \mathbf{E}(\tilde{\psi}_n(x)) \right) + \left(\tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right) \right| \\ &\leq \max_{1 \leq k \leq l_n} \sup_{x \in B_k} \left| \tilde{\psi}_n(x) - \tilde{\psi}_n(x_k) \right| + \max_{1 \leq k \leq l_n} \sup_{x \in B_k} \left| \mathbf{E}(\tilde{\psi}_n(x)) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| + \max_{1 \leq k \leq l_n} \left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| \\ &=: I_{21} + I_{22} + I_{23}. \end{aligned} \tag{3.30}$$

Sous l'hypothèse **(H3.2)**, on a

$$\begin{aligned} \left| \tilde{\psi}_n(x) - \tilde{\psi}_n(x_k) \right| &= \left| \frac{\alpha}{n h_n^d} \sum_{i=1}^n \frac{Y_i}{G(Y_i)} \left[K_d \left(\frac{x - X_i}{h_n} \right) - K_d \left(\frac{x_k - X_i}{h_n} \right) \right] \right| \\ &\leq \frac{C}{n h_n^d G(a_F)} \sum_{i=1}^n \left| K_d \left(\frac{x - X_i}{h_n} \right) - K_d \left(\frac{x_k - X_i}{h_n} \right) \right| \\ &\leq \frac{C}{h_n^d G(a_F)} \left\| \frac{x - X_1}{h_n} - \frac{x_k - X_1}{h_n} \right\|^\beta \\ &\leq \frac{\|x - x_k\|^\beta}{G(a_F) h_n^{\beta+d}}. \end{aligned}$$

Pour tout $x \in B_k$, nous avons $\|x - x_k\| \leq a_n^d$, ainsi

$$\sup_{x \in B_k} \left| \tilde{\psi}_n(x) - \tilde{\psi}_n(x_k) \right| \leq \frac{(a_n^d)^\beta}{G(a_F) h_n^{\beta+d}} = O \left(\frac{1}{\sqrt{n h_n^d}} \right).$$

D'où,

$$I_{21} = \max_{1 \leq k \leq l_n} \sup_{x \in B_k} \left| \tilde{\psi}_n(x) - \tilde{\psi}_n(x_k) \right| = O \left(\frac{1}{\sqrt{n h_n^d}} \right). \tag{3.31}$$

En raisonnant d'une manière similaire, on obtient aussi

$$I_{22} = \max_{1 \leq k \leq l_n} \sup_{x \in B_k} \left| \mathbf{E}(\tilde{\psi}_n(x)) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| = O \left(\frac{1}{\sqrt{n h_n^d}} \right). \tag{3.32}$$

Concernant I_{23} , on utilise l'inégalité de Bernstein qui spécifie que, si $(Z_i)_i$ est une suite de variables aléatoires indépendantes et centrées et s'il existe une constante positive $M < \infty$, telle que $|Z_1| \leq M$, alors on a

$$\forall \varepsilon > 0, \mathbf{P} \left(\left| \sum_{i=1}^n Z_i \right| > \varepsilon n \right) \leq 2 \exp \left\{ \frac{-\varepsilon^2 n}{2\sigma^2 \left(1 + \frac{M\varepsilon}{\sigma^2} \right)} \right\}, \tag{3.33}$$

où $\sigma^2 = \mathbf{E}(Z_1^2)$.

Pour cela, posons

$$Z_i(x) := \frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right).$$

Il est clair que

$$\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) = \frac{1}{nh_n^d} \sum_{i=1}^n Z_i(x),$$

$\mathbf{E}(Z_i(x)) = 0$ et $|Z_i(x)| \leq \frac{2C\|K_d\|_\infty}{G(a_F)} =: M$.

Nous devons maintenant calculer $\mathbf{E}(Z_1^2)$.

$$\begin{aligned} \mathbf{E}(Z_1^2) &= \mathbf{E} \left[\frac{\alpha^2 Y_1^2}{G^2(Y_1)} K_d^2 \left(\frac{x - X_1}{h_n} \right) \right] - \mathbf{E}^2 \left[\frac{\alpha Y_1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right] \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha^2 y^2}{G^2(y)} K_d^2 \left(\frac{x - u}{h_n} \right) f^*(u, y) du dy - \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha y}{G(y)} K_d \left(\frac{x - u}{h_n} \right) f^*(u, y) du dy \right)^2 \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha y^2}{G(y)} K_d^2 \left(\frac{x - u}{h_n} \right) f(u, y) du dy - \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy \right)^2 \quad (\text{en utilisant (3.12)}) \\ &= \int_{\mathbb{R}^d} \alpha K_d^2 \left(\frac{x - u}{h_n} \right) \psi_1(u) du - \left(\int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_n} \right) \psi(u) du \right)^2 \\ &= h_n^d \int_{\mathbb{R}^d} \alpha K_d^2(z) \psi_1(x - h_n z) dz - h_n^{2d} \left(\int_{\mathbb{R}^d} K_d(z) \psi(x - h_n z) dz \right)^2 \quad (\text{en posant } z = \frac{x-u}{h_n}) \\ &=: E_1 - E_2. \end{aligned}$$

Sous les hypothèses **(H3.2)** et **(H3.4)**, en faisant un développement de Taylor de ψ_1 au voisinage de x , on obtient

$$\begin{aligned} E_1 &= h_n^d \left\{ \int_{\mathbb{R}^d} K_d^2(z) \left[\psi_1(x) - h_n \left(z_1 \frac{\partial \psi_1}{\partial z_1}(x^*) + \dots + z_d \frac{\partial \psi_1}{\partial z_d}(x^*) \right) \right] dz \right\} \quad (\text{tel que } x - h_n z \leq x^* \leq x) \\ &\leq h_n^d \left\{ \psi_1(x) \int_{\mathbb{R}^d} K_d^2(z) dz - h_n \left[\sup_{x \in \mathcal{C}} \left| \frac{\partial \psi_1}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d^2(z) dz + \dots + \sup_{x \in \mathcal{C}} \left| \frac{\partial \psi_1}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d^2(z) dz \right] \right\} \\ &= O(h_n^d). \end{aligned}$$

De la même manière, sous les hypothèses **(H3.2)** et **(H3.3)**, en faisant un développement de Taylor de ψ au voisinage de x , on obtient

$$\begin{aligned} E_2 &= h_n^{2d} \left\{ \int_{\mathbb{R}^d} K_d(z) \left[\psi(x) - h_n \left(z_1 \frac{\partial \psi}{\partial z_1}(x^*) + \dots + z_d \frac{\partial \psi}{\partial z_d}(x^*) \right) \right] dz \right\}^2 \quad (\text{tel que } x - h_n z \leq x^* \leq x) \\ &\leq h_n^{2d} \left\{ \psi(x) - h_n \left[\sup_{x \in \mathcal{C}} \left| \frac{\partial \psi}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d(z) dz + \dots + \sup_{x \in \mathcal{C}} \left| \frac{\partial \psi}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d(z) dz \right] \right\}^2 \\ &= O(h_n^{2d}). \end{aligned}$$

Par conséquent,

$$\mathbf{E}(Z_1^2) = O(h_n^d) - O(h_n^{2d}) = O(h_n^d).$$

En appliquant maintenant l'inégalité de Bernstein donné en (3.33), on aura

$$\begin{aligned}
\mathbf{P} \left(\max_{1 \leq k \leq l_n} \left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| > \varepsilon \right) &\leq \sum_{j=1}^{l_n} \mathbf{P} \left(\left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| > \varepsilon \right) \\
&= l_n \sup_{x \in \mathcal{C}} \mathbf{P} \left(\left| \frac{1}{nh_n^d} \sum_{i=1}^n Z_i(x_k) \right| > \varepsilon \right) \\
&= l_n \sup_{x \in \mathcal{C}} \mathbf{P} \left(\left| \sum_{i=1}^n Z_i(x_k) \right| > nh_n^d \varepsilon \right) \\
&\leq 2l_n \exp \left\{ \frac{-(h_n^d \varepsilon)^2 n}{2\sigma^2 \left(1 + \frac{Mh_n^d \varepsilon}{\sigma^2} \right)} \right\} \\
&\leq 2l_n \exp \left\{ \frac{-h_n^d \varepsilon^2 n}{2(C + M\varepsilon)} \right\}.
\end{aligned}$$

En choisissant $\varepsilon = \varepsilon_0 \sqrt{\frac{\log n}{nh_n^d}}$, pour tout $\varepsilon_0 > 0$, on aura

$$\begin{aligned}
\mathbf{P} \left(\max_{1 \leq k \leq l_n} \left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| > \varepsilon_0 \sqrt{\frac{\log n}{nh_n^d}} \right) &\leq Ca_N^{-d} \exp \left\{ \frac{-\varepsilon_0^2 \log n}{2 \left(C + M\varepsilon_0 \sqrt{\frac{\log n}{nh_n^d}} \right)} \right\} \\
&= Cn^{\frac{1}{2\beta}} h_n^{-1 - \frac{d}{2\beta}} n^{-\frac{\varepsilon_0^2}{2 \left(C + M\varepsilon_0 \sqrt{\frac{\log n}{nh_n^d}} \right)}} \\
&\approx \frac{C}{h_n^{-1 - \frac{d}{2\beta}}} n^{\frac{1}{2} \left(\frac{1}{\beta} - \frac{\varepsilon_0^2}{C} \right)} \quad (\text{sous l'hypothèse (H3.1)}).
\end{aligned}$$

Pour un choix approprié de ε_0 , le terme de droite est le terme général d'une série convergente. Ainsi,

$$\sum_{n=1}^{\infty} \mathbf{P} \left(\max_{1 \leq k \leq l_n} \left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| > \varepsilon_0 \sqrt{\frac{\log n}{nh_n^d}} \right) < \infty$$

et en utilisant le lemme de Borel Cantelli, nous avons

$$I_{23} = \max_{1 \leq k \leq l_n} \left| \tilde{\psi}_n(x_k) - \mathbf{E}(\tilde{\psi}_n(x_k)) \right| = O \left(\sqrt{\frac{\log n}{nh_n^d}} \right) \quad \text{p.s., lorsque } n \rightarrow \infty. \quad (3.34)$$

En combinant (3.31), (3.32), (3.34) avec (3.30), on obtient

$$I_2 = \sup_{x \in D} \left| \hat{\psi}_n(x) - \psi(x) \right| = O \left(\sqrt{\frac{\log n}{nh_n^d}} \right) \quad \text{p.s., lorsque } n \rightarrow \infty. \quad (3.35)$$

Pour le terme de biais I_3 , les hypothèses **(H3.2)** et **(H3.3)** et un développement de Taylor au

voisinage de x permettent d'écrire

$$\begin{aligned}
\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) &= \mathbf{E} \left(\frac{1}{nh_n^d} \sum_{i=1}^n \frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) - \psi(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_R \frac{\alpha y}{G(y)} K_d \left(\frac{x - u}{h_n} \right) f^*(u, y) du dy - \psi(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy - \psi(x) \\
&= \int_{\mathbb{R}^d} K_d(z) (\psi(x - zh_n) - \psi(x)) dz \\
&= \int_{\mathbb{R}^d} K_d(z) \left[-h_n \left(z_1 \frac{\partial \psi}{\partial z_1}(x) + \dots + z_d \frac{\partial \psi}{\partial z_d}(x) \right) + \frac{h_n^2}{2} \left(z_1^2 \frac{\partial^2 \psi}{\partial z_1^2}(x^*) + \dots + z_d^2 \frac{\partial^2 \psi}{\partial z_d^2}(x^*) \right. \right. \\
&\quad \left. \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} z_i z_j \frac{\partial^2 \psi}{\partial z_i \partial z_j}(x^*) \right) \right] dz \quad (\text{tel que } x - h_n z \leq x^* \leq x) \\
&\leq \frac{h_n^2}{2} \left[\sup_{x \in \mathcal{C}} \left| \frac{\partial^2 \psi}{\partial z_1^2}(x) \right| \int_{\mathbb{R}^d} z_1^2 K_d(z) dz + \dots + \sup_{x \in \mathcal{C}} \left| \frac{\partial^2 \psi}{\partial z_d^2}(x) \right| \int_{\mathbb{R}^d} z_d^2 K_d(z) dz \right. \\
&\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \sup_{x \in \mathcal{C}} \left| \frac{\partial^2 \psi}{\partial z_i \partial z_j}(x) \right| \int_{\mathbb{R}^d} z_i z_j K_d(z) dz \right] \\
&= O(h_n^2). \tag{3.36}
\end{aligned}$$

Ce qui donne

$$I_3 = \sup_{x \in \mathcal{C}} |\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x)| = O(h_n^2). \tag{3.37}$$

Finalement, l'utilisation de (3.29), (3.35) et (3.37) permet d'obtenir le résultat donné dans le lemme.

□

Lemme 3.2. *Sous les hypothèses (H3.1), (H3.2) et (H3.5), nous avons*

$$\sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee h_n^2 \right\} \mathbf{P} - p.s, \text{ lorsque } n \rightarrow \infty.$$

Preuve. Nous avons

$$\begin{aligned}
\sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)| &\leq \sup_{x \in \mathcal{C}} |\hat{v}_n(x) - \tilde{v}_n(x)| + \sup_{x \in \mathcal{C}} |\tilde{v}_n(x) - \mathbf{E}(\tilde{v}_n(x))| + \sup_{x \in \mathcal{C}} |\mathbf{E}(\tilde{v}_n(x)) - v(x)| \\
&=: II_1 + II_2 + II_3.
\end{aligned}$$

Commençons par le terme II_1 . On a

$$|\hat{v}_n(x) - \tilde{v}_n(x)| \leq \left\{ \frac{|\alpha_n - \alpha|}{G_n(a_F)} + \frac{\alpha}{G_n(a_F)G(a_F)} \sup_{y \geq a_F} |G_n(y) - G(y)| \right\} \frac{1}{nh_n^d} \sum_{i=1}^n K_d \left(\frac{x - X_i}{h_n} \right).$$

Ainsi, en utilisant les mêmes arguments que pour le terme I_1 dans la preuve du lemme (3.1), nous obtenons

$$II_1 = O_{\mathbf{P}}(n^{-1/2}). \quad (3.38)$$

Pour le terme II_2 , on reprend les mêmes étapes de la preuve du terme I_2 dans la preuve du lemme (3.1) en posant

$$Z_i(x) := \frac{\alpha}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right) - \mathbf{E}\left(\frac{\alpha}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)\right),$$

on montre que

$$II_2 = \sup_{x \in D} |\hat{v}_n(x) - v(x)| = O\left(\sqrt{\frac{\log n}{nh_n^d}}\right) \text{ p.s., lorsque } n \rightarrow \infty. \quad (3.39)$$

Finalement, par un changement de variable et un développement de Taylor, sous les hypothèses **(H3.2)** et **(H3.5)**, nous avons

$$\begin{aligned} \mathbf{E}(\tilde{v}_n(x)) - v(x) &= \mathbf{E}\left(\frac{1}{nh_n^d} \sum_{i=1}^n \frac{\alpha}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)\right) - v(x) \\ &= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_R \frac{\alpha}{G(y)} K_d\left(\frac{x - u}{h_n}\right) f^*(u, y) du dy - v(x) \\ &= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} K_d\left(\frac{x - u}{h_n}\right) f(u, y) du dy - v(x) \\ &= \int_{\mathbb{R}^d} K_d(z) (v(x - zh_n) - v(x)) dz \\ &= \int_{\mathbb{R}^d} K_d(z) \left[-h_n \left(z_1 \frac{\partial v}{\partial z_1}(x) + \dots + z_d \frac{\partial v}{\partial z_d}(x) \right) + \frac{h_n^2}{2} \left(z_1^2 \frac{\partial^2 v}{\partial z_1^2}(x^*) + \dots + z_d^2 \frac{\partial^2 v}{\partial z_d^2}(x^*) \right) \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} z_i z_j \frac{\partial^2 v}{\partial z_i \partial z_j}(x^*) \right] dz \quad (\text{tel que } x - h_n z \leq x^* \leq x) \\ &\leq \frac{h_n^2}{2} \left[\sup_{x \in \mathcal{C}} \left| \frac{\partial^2 v}{\partial z_1^2}(x) \right| \int_{\mathbb{R}^d} z_1^2 K_d(z) dz + \dots + \sup_{x \in \mathcal{C}} \left| \frac{\partial^2 v}{\partial z_d^2}(x) \right| \int_{\mathbb{R}^d} z_d^2 K_d(z) dz \right. \\ &\quad \left. + 2 \sum_{\substack{1 \leq i, j \leq d \\ i \neq j}} \sup_{x \in \mathcal{C}} \left| \frac{\partial^2 v}{\partial z_i \partial z_j}(x) \right| \int_{\mathbb{R}^d} z_i z_j K_d(z) dz \right] \\ &= O(h_n^2). \end{aligned} \quad (3.40)$$

Alors, on obtient

$$II_3 = O(h_n^2). \quad (3.41)$$

En regroupant les résultats (3.38), (3.39) et (3.41), on conclut le résultat donné dans le lemme. \square

Revenons maintenant à la preuve du Théorème (3.1). Comme

$$\hat{m}_n(x) - m(x) = \frac{\hat{\psi}_n(x) - \psi(x)}{\hat{v}_n(x)} + m(x) \frac{v(x) - \hat{v}_n(x)}{\hat{v}_n(x)}.$$

Alors,

$$\sup_{x \in \mathcal{C}} |\hat{m}_n(x) - m(x)| \leq \frac{1}{\inf_{x \in \mathcal{C}} |v_n(x)|} \left\{ \sup_{x \in \mathcal{C}} |\hat{\psi}_n(x) - \psi(x)| + \frac{\sup_{x \in \mathcal{C}} |\psi(x)|}{\inf_{x \in \mathcal{C}} |v(x)|} \sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)| \right\}. \quad (3.42)$$

D'autre part, nous avons $|v_n(x)| = |v(x) + (v_n(x) - v(x))| \geq ||v(x)| - |v_n(x) - v(x)|| \geq \eta - \sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)|$. Ainsi, (3.42) devient

$$\sup_{x \in \mathcal{C}} |\hat{m}_n(x) - m(x)| \leq \frac{1}{\eta - \sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)|} \left\{ \sup_{x \in \mathcal{C}} |\hat{\psi}_n(x) - \psi(x)| + \eta^{-1} \sup_{x \in \mathcal{C}} |\psi(x)| \sup_{x \in \mathcal{C}} |\hat{v}_n(x) - v(x)| \right\}.$$

L'utilisation de Lemme 3.1 et Lemme 3.2 complète la preuve. \square

Preuve du Théorème 3.2. Pour montrer le résultat donné dans ce théorème, nous utilisons la delta méthode donnée dans le lemme suivant.

Lemme 3.3. (*Shao (2003), Corollary 1.1, p. 61*) Soient X_1, X_2, \dots et Y des vecteurs aléatoires de \mathbb{R}^d vérifiant

$$a_n(X_n - b) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \Sigma),$$

où $b \in \mathbb{R}^d$ et $(a_n)_n$ est une suite de nombres positifs avec $a_n \xrightarrow[n \rightarrow +\infty]{} \infty$. Soit φ une fonction de \mathbb{R}^d dans \mathbb{R} . Si φ est différentiable au point b , alors

$$a_n[\varphi(X_n) - \varphi(b)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, [\nabla \varphi(b)]^T \Sigma \nabla \varphi(b)),$$

où $\nabla \varphi(x)$ dénote le gradient de φ au point x .

Observons maintenant que

$$\hat{m}_n(x) = \frac{\alpha^{-1} \hat{\psi}_n(x)}{\alpha^{-1} \hat{v}_n(x)}, \quad (3.43)$$

avec

$$\frac{\hat{\psi}_n(x)}{\alpha} = \frac{1}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right)$$

et

$$\frac{\hat{v}_n(x)}{\alpha} = \frac{1}{nh_n^d} \sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right).$$

De (3.43), nous pouvons écrire

$$\hat{m}_n(x) - m(x) = \varphi(\alpha^{-1} \hat{\psi}_n(x), \alpha^{-1} \hat{v}_n(x)) - \varphi(\alpha^{-1} \psi(x), \alpha^{-1} v(x)),$$

où φ est une fonction définie par :

$$\begin{aligned} \varphi &: \mathbb{R} \times \mathbb{R}^* \rightarrow \mathbb{R} \\ (x, y) &\mapsto \frac{x}{y} \end{aligned}$$

Pour utiliser le Lemme 3.3, nous devons déterminer la loi de probabilité de

$$\sqrt{nh_n^d} \left[\alpha^{-1}(\hat{\psi}_n(x), \hat{v}_n(x)) - \alpha^{-1}(\psi(x), v(x)) \right].$$

Pour cela, nous utilisons les décompositions suivantes

$$\begin{aligned} \frac{\hat{\psi}_n(x)}{\alpha} - \frac{\psi(x)}{\alpha} &= \frac{1}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \frac{\psi(x)}{\alpha} \\ &= \frac{1}{nh_n^d} \sum_{i=1}^n \left(\frac{1}{G_n(Y_i)} - \frac{1}{G(Y_i)} \right) Y_i K_d \left(\frac{x - X_i}{h_n} \right) \\ &\quad + \frac{1}{nh_n^d} \sum_{i=1}^n \left[\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right] \\ &\quad + \left[\frac{1}{nh_n^d} \sum_{i=1}^n \mathbf{E} \left(\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) - \frac{\psi(x)}{\alpha} \right] \\ &=: \Lambda_{n1}(x) + \Lambda_{n2}(x) + \Lambda_{n3}(x). \end{aligned} \tag{3.44}$$

$$\begin{aligned} \frac{\hat{v}_n(x)}{\alpha} - \frac{v(x)}{\alpha} &= \frac{1}{nh_n^d} \sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \frac{v(x)}{\alpha} \\ &= \frac{1}{nh_n^d} \sum_{i=1}^n \left(\frac{1}{G_n(Y_i)} - \frac{1}{G(Y_i)} \right) K_d \left(\frac{x - X_i}{h_n} \right) \\ &\quad + \frac{1}{nh_n^d} \sum_{i=1}^n \left[\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right] \\ &\quad + \left[\frac{1}{nh_n^d} \sum_{i=1}^n \mathbf{E} \left(\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) - \frac{v(x)}{\alpha} \right] \\ &=: \Gamma_{n1}(x) + \Gamma_{n2}(x) + \Gamma_{n3}(x). \end{aligned} \tag{3.45}$$

Nous commençons par montrer que les termes $\sqrt{nh_n^d}\Lambda_{n1}(x)$, $\sqrt{nh_n^d}\Lambda_{n3}(x)$, $\sqrt{nh_n^d}\Gamma_{n1}(x)$ et $\sqrt{nh_n^d}\Gamma_{n3}(x)$ dans (3.44) et (3.45) sont négligeables. Ceci est fait dans les Lemmes 3.4 et 3.5. Ensuite, nous montrons dans le Lemme 3.6 que le vecteur formé des deux termes dominants $\sqrt{nh_n^d}\Lambda_{n2}(x)$ et $\sqrt{nh_n^d}\Gamma_{n2}(x)$ est asymptotiquement gaussien en utilisant le théorème central limite de Lyapounov (voir Pollard (1984), p. 51).

Lemme 3.4. *Sous les hypothèses (H3.1), (H3.2), (H3.5) et pour tout x , les deux termes $\sqrt{nh_n^d}\Lambda_{n1}(x)$ et $\sqrt{nh_n^d}\Gamma_{n1}(x)$ sont $o_{\mathbf{P}}(1)$ lorsque $n \rightarrow \infty$.*

Preuve. En utilisant (3.27), (3.28) et les hypothèses (H3.1), (H3.2) et (H3.5), nous avons

$$\begin{aligned} \sqrt{nh_n^d}\Lambda_{n1}(x) &\leq \sqrt{nh_n^d} \frac{\sup_{y \geq a_F} |G_n(y) - G(y)|}{G_n(a_F)G(a_F)} \times \frac{1}{nh_n^d} \sum_{i=1}^n Y_i K_d \left(\frac{x - X_i}{h_n} \right) \\ &= O_{\mathbf{P}}(\sqrt{h_n^d}) = o_{\mathbf{P}}(1). \end{aligned}$$

D'une manière similaire, nous avons

$$\sqrt{nh_n^d}\Gamma_{n1}(x) = O_{\mathbf{P}}(\sqrt{h_n^d}) = o_{\mathbf{P}}(1).$$

□

Lemme 3.5. *Sous les hypothèses (H3.2), (H3.3), (H3.5), (H3.6) et pour tout x , les deux termes $\sqrt{nh_n^d}\Lambda_{n3}(x)$ et $\sqrt{nh_n^d}\Gamma_{n3}(x)$ sont $o_{\mathbf{P}}(1)$ lorsque $n \rightarrow \infty$.*

Preuve. Nous avons

$$\begin{aligned}\sqrt{nh_n^d}\Lambda_{n3}(x) &= \sqrt{nh_n^d} \left[\frac{1}{nh_n^d} \sum_{i=1}^n \mathbf{E} \left(\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) - \frac{\psi(x)}{\alpha} \right] \\ &= \frac{\sqrt{nh_n^d}}{\alpha} \left[\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right]\end{aligned}$$

Ainsi, en utilisant (3.36) et l'hypothèse (H3.6), nous obtenons

$$\sqrt{nh_n^d}\Lambda_{n3}(x) = O_{\mathbf{P}}(\sqrt{nh_n^{d+4}}) = o_{\mathbf{P}}(1).$$

D'une manière similaire, en utilisant (3.40) et l'hypothèse (H3.6), nous obtenons

$$\sqrt{nh_n^d}\Gamma_{n3}(x) = O_{\mathbf{P}}(\sqrt{nh_n^{d+4}}) = o_{\mathbf{P}}(1).$$

□

Lemme 3.6. *Sous les hypothèses (H3.1)-(H3.3), (H3.5) et (H3.7), nous avons*

$$\sqrt{nh_n^d}(\Lambda_{n2}(x), \Gamma_{n2}(x))^T \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \alpha^{-1} \kappa \Sigma(x))$$

où $\Sigma(x)$ est donné par (3.23).

Preuve. Commençons par calculer les variances et les covariances nécessaires pour montrer ce lemme.

1. Calcul de la variance de $\sqrt{nh_n^d}\Lambda_{n2}(x)$:

$$\begin{aligned}\text{Var} \left(\sqrt{nh_n^d}\Lambda_{n2}(x) \right) &= \frac{1}{nh_n^d} \times n \text{Var} \left[\frac{Y_1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right] \\ &= \frac{1}{\alpha h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{y^2}{G(y)} K_d^2 \left(\frac{x - u}{h_n} \right) f(u, y) du dy \\ &\quad - \frac{1}{\alpha^2 h_n^d} \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy \right\}^2 \\ &= \frac{1}{\alpha h_n^d} \int_{\mathbb{R}^d} K_d^2 \left(\frac{x - u}{h_n} \right) \Sigma_0(u) du - \frac{1}{\alpha^2 h_n^d} \left\{ \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_n} \right) \psi(u) du \right\}^2 \\ &= \frac{1}{\alpha} \int_{\mathbb{R}^d} K_d^2(z) \Sigma_0(x - zh_n) dz - \frac{h_n^d}{\alpha^2} \left\{ \int_{\mathbb{R}^d} K_d(z) \psi(x - zh_n) du \right\}^2.\end{aligned}$$

D'après le théorème de convergence dominée,

$$\frac{1}{\alpha} \int_{\mathbb{R}^d} K_d^2(z) \Sigma_0(x - h_n z) dz \xrightarrow{n \rightarrow +\infty} \frac{1}{\alpha} \Sigma_0(x) \int_{\mathbb{R}^d} K_d^2(z) dz. \quad (3.46)$$

Sous les hypothèses **(H3.1)**, **(H3.2)** et **(H3.3)**, en faisant un développement de Taylor de ψ au voisinage de x , on obtient

$$\begin{aligned} \frac{h_n^d}{\alpha^2} \int_{\mathbb{R}^d} K_d(z) \psi(x - zh_n) du &\leq \frac{h_n^d}{\alpha^2} \left\{ \psi(x) - h_n \left[\sup_{x \in \mathcal{C}} \left| \frac{\partial \psi}{\partial z_1}(x) \right| \int_{\mathbb{R}^d} z_1 K_d(z) dz + \dots \right. \right. \\ &\quad \left. \left. + \sup_{x \in \mathcal{C}} \left| \frac{\partial \psi}{\partial z_d}(x) \right| \int_{\mathbb{R}^d} z_d K_d(z) dz \right] \right\} \\ &= O(h_n^d) = o(1), \end{aligned} \quad (3.47)$$

En combinant (3.46) et (3.47), nous obtenons

$$\text{Var} \left(\sqrt{nh_n^d} \Lambda_{n2}(x) \right) = \frac{\kappa}{\alpha} \Sigma_0(x) + o(1). \quad (3.48)$$

2. Calcul de la variance de $\sqrt{nh_n^d} \Gamma_{n2}(x)$:

$$\begin{aligned} \text{Var} \left(\sqrt{nh_n^d} \Gamma_{n2}(x) \right) &= \frac{1}{nh_n^d} \times n \text{Var} \left[\frac{1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right] \\ &= \frac{1}{\alpha h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{1}{G(y)} K_d^2 \left(\frac{x - u}{h_n} \right) f(u, y) du dy \\ &\quad - \frac{1}{\alpha^2 h_n^d} \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}} K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy \right\}^2 \\ &= \frac{1}{\alpha h_n^d} \int_{\mathbb{R}^d} K_d^2 \left(\frac{x - u}{h_n} \right) \Sigma_2(u) du - \frac{1}{\alpha^2 h_n^d} \left\{ \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_n} \right) v(u) du \right\}^2 \\ &= \frac{1}{\alpha} \int_{\mathbb{R}^d} K_d^2(z) \Sigma_2(x - zh_n) dz - \frac{h_n^d}{\alpha^2} \left\{ \int_{\mathbb{R}^d} K_d(z) v(x - zh_n) du \right\}^2. \end{aligned}$$

En procédant comme en 1), sous les hypothèses **(H3.1)**, **(H3.2)** et **(H3.5)**, nous obtenons

$$\text{Var} \left(\sqrt{nh_n^d} \Gamma_{n2}(x) \right) = \frac{\kappa}{\alpha} \Sigma_2(x) + o(1). \quad (3.49)$$

3. Calcul de la covariance entre $\sqrt{nh_n^d} \Lambda_{n2}(x)$ et $\sqrt{nh_n^d} \Gamma_{n2}(x)$:

Notons $COV := Cov \left(\sqrt{nh_n^d} \Lambda_{n2}(x), \sqrt{nh_n^d} \Gamma_{n2}(x) \right)$, ainsi

$$\begin{aligned} COV &= \frac{1}{h_n^d} \left\{ \mathbf{E} \left(\frac{Y_1}{G^2(Y_1)} K_d^2 \left(\frac{x - X_1}{h_n} \right) \right) - \mathbf{E} \left(\frac{Y_1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right) \mathbf{E} \left(\frac{1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right) \right\} \\ &= \frac{1}{\alpha h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{y}{G(y)} K_d^2 \left(\frac{x - u}{h_n} \right) f(u, y) du dy - \frac{1}{\alpha^2 h_n^d} \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy \right\} \\ &\quad \left\{ \int_{\mathbb{R}^d} \int_{\mathbb{R}} K_d \left(\frac{x - u}{h_n} \right) f(u, y) du dy \right\} \\ &= \frac{1}{\alpha} \int_{\mathbb{R}^d} K_d^2(z) \Sigma_1(x - zh_n) dz - \frac{h_n^d}{\alpha^2} \left\{ \int_{\mathbb{R}^d} K_d(z) \psi(x - zh_n) du \right\} \left\{ \int_{\mathbb{R}^d} K_d(z) v(x - zh_n) du \right\}. \end{aligned}$$

En procédant aussi comme en 1), sous les hypothèses **(H3.1)**-**(H3.3)** et **(H3.5)**, nous obtenons

$$\text{Cov} \left(\sqrt{nh_n^d} \Lambda_{n2}(x), \sqrt{nh_n^d} \Gamma_{n2}(x) \right) = \frac{\kappa}{\alpha} \Sigma_1(x) + o(1). \quad (3.50)$$

Nous allons maintenant démontrer que toute combinaison linéaire de $\sqrt{nh_n^d} \Lambda_{n2}(x)$ et $\sqrt{nh_n^d} \Gamma_{n2}(x)$ est asymptotiquement gaussienne. Pour un vecteur $c = (c_1, c_2)^T$ de nombres réels, posons

$$\begin{aligned} \Delta_n(x) &= \sqrt{nh_n^d} [c_1 \Lambda_{n2}(x) + c_2 \Gamma_{n2}(x)] \\ &= \sum_{i=1}^n \left\{ c_1 \frac{1}{\sqrt{nh_n^d}} \left[\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right] \right. \\ &\quad \left. + c_2 \frac{1}{\sqrt{nh_n^d}} \left[\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right] \right\} \\ &=: \sum_{i=1}^n \Delta_{ni}(x) \end{aligned}$$

Il est clair que les $\Delta_{ni}(x)$ sont i.i.d. Nous allons appliquer le théorème central limite de Lyapounov à la suite $\Delta_{ni}(x)$. Soit $\rho_{ni}^\nu(x) = \mathbf{E} [|\Delta_{ni}(x)|^\nu]$, pour $\nu > 2$. Par la C_r -inégalité (voir Loève (1965), p.156), nous obtenons

$$\begin{aligned} \rho_{ni}^\nu(x) &= \mathbf{E} \left| c_1 \frac{1}{\sqrt{nh_n^d}} \left\{ \frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right\} \right. \\ &\quad \left. + c_2 \frac{1}{\sqrt{nh_n^d}} \left\{ \frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \right\} \right|^\nu \\ &\leq 2^{\nu-1} \left(\frac{h_n^d}{n} \right)^{\frac{\nu}{2}} \left\{ c_1^\nu \mathbf{E} \left[\left| \frac{1}{h_n^d} \frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right|^\nu \right] + \right. \\ &\quad \left. + c_2^\nu \mathbf{E} \left[\left| \frac{1}{h_n^d} \frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right|^\nu \right] \right\} \end{aligned}$$

qui implique que, sous l'hypothèse **(H3.7)**,

$$\rho_n^\nu(x) := \sum_{i=1}^n \rho_{ni}^\nu(x) = O \left(n^{1-\frac{\nu}{2}} h_n^{\frac{d\nu}{2}} \right) = o(1) \quad (\text{comme } 1 - \frac{\nu}{2} < 0). \quad (3.51)$$

D'autre part, en utilisant (3.48), (3.49) et (3.50), nous obtenons

$$\begin{aligned} s_n^2(x) := \text{Var}(\Delta_n(x)) &= c_1^2 \text{Var} \left(\sqrt{nh_n^d} \Lambda_{n2}(x) \right) + c_2^2 \text{Var} \left(\sqrt{nh_n^d} \Gamma_{n2}(x) \right) \\ &\quad + 2c_1 c_2 \text{Cov} \left(\sqrt{nh_n^d} \Lambda_{n2}(x), \sqrt{nh_n^d} \Gamma_{n2}(x) \right) \\ &= c_1^2 \frac{\kappa}{\alpha} \Sigma_0(x) + c_2^2 \frac{\kappa}{\alpha} \Sigma_2(x) + 2c_1 c_2 \frac{\kappa}{\alpha} \Sigma_1(x) + o(1). \end{aligned}$$

Ainsi,

$$s_n^2(x) \xrightarrow{n \rightarrow +\infty} \alpha^{-1} \kappa c^T \Sigma(x) c > 0, \quad (3.52)$$

pour tout $c \neq 0$ pourvu que $v(x) > 0$. Alors, (3.51) et (3.52) donne

$$\frac{1}{(Var(\Delta_n(x)))^{\frac{\nu}{2}}} \sum_{i=1}^n \mathbf{E} [|\Delta_{ni}(x)|^\nu] = \frac{\rho_n^\nu(x)}{s_n^\nu(x)} \xrightarrow{n \rightarrow +\infty} 0.$$

La condition de Lyapounov étant satisfaite, le lemme est donc démontré par le théorème central limite de Lyapounov. \square

Revenons maintenant à la preuve de Théorème 3.2. En combinant les résultats des lemmes précédents, nous obtenons par le théorème de Mann-Wald (voir Rao (1965) p. 321)

$$\sqrt{nh_n^d} [\alpha^{-1}(\hat{\psi}_n(x), \hat{v}_n(x)) - \alpha^{-1}(\psi(x), v(x))] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \alpha^{-1} \kappa \Sigma(x)).$$

Puis, le Lemme 3.3 nous permet d'avoir

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \alpha^{-1} \kappa \nabla \varphi^T \Sigma(x) \nabla \varphi),$$

où le gradient $\nabla \varphi = \left(\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y} \right)$ est évalué au point $\alpha^{-1}(\psi(x), v(x))$.

La preuve se termine en remarquant que

$$\begin{aligned} \alpha^{-1} \kappa \nabla \varphi_{(\alpha^{-1}(\psi(x), v(x)))}^T \Sigma(x) \nabla \varphi_{(\alpha^{-1}(\psi(x), v(x)))} &= \alpha^{-1} \kappa \begin{pmatrix} \frac{\alpha}{v(x)} & \frac{-\alpha \psi(x)}{v^2(x)} \end{pmatrix} \begin{pmatrix} \Sigma_0(x) & \Sigma_1(x) \\ \Sigma_1(x) & \Sigma_2(x) \end{pmatrix} \begin{pmatrix} \frac{\alpha}{v(x)} \\ \frac{-\alpha \psi(x)}{v^2(x)} \end{pmatrix} \\ &= \alpha^{-1} \kappa \begin{pmatrix} \frac{\alpha \Sigma_0(x)}{v(x)} - \frac{\alpha \psi(x) \Sigma_1(x)}{v^2(x)} & \frac{\alpha \Sigma_1(x)}{v(x)} - \frac{\alpha \psi(x) \Sigma_2(x)}{v^2(x)} \end{pmatrix} \begin{pmatrix} \frac{\alpha}{v(x)} \\ \frac{-\alpha \psi(x)}{v^2(x)} \end{pmatrix} \\ &= \frac{\alpha \kappa [\Sigma_0(x) v^2(x) - 2 \Sigma_1(x) \psi(x) v(x) + \Sigma_2(x) \psi^2(x)]}{v^4(x)} \\ &= \sigma^2(x). \end{aligned}$$

\square

3.2.2 Cas de suites de variables aléatoires α -mélangeantes

Cette partie porte sur l'étude des propriétés asymptotiques de les estimateurs de la fonction de régression et de la densité des covariées introduits par Ould Saïd et Lemdani (2006) pour des données tronquées à gauche et α -mélangeantes. Plus précisément, nous reprenons les résultats obtenus par Liang, Li et Qi (2009) où ils montrent la convergence presque sûre de ces estimateurs et les résultats obtenus par Liang (2011) où il établit leurs normalité asymptotique.

Dans ce qui suit, la suite observée $\{(X_i, Y_i, T_i); 1 \leq i \leq n\}$ est supposée stationnaire et α -mélangeante de coefficient de mélange $\alpha(k)$. Rappelons qu'une suite de variables aléatoires $\{\xi_i; i \geq 1\}$ est α -mélangeante si le coefficient de α -mélange

$$\alpha(n) = \sup \{ |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|, A \in \mathcal{F}_1^i \text{ et } B \in \mathcal{F}_{i+n}^\infty, i \geq 1 \}$$

tend vers zéro quand n tend vers l'infini où \mathcal{F}_a^b désigne la tribu engendrée par $\{\xi_i; a \leq i \leq b\}$.

Remarque 3.2. Nous faisons remarquer ici que, c'est la suite des variables réellement observées qui est supposée α -mélangeante et non pas la suite des données d'origine. En effet, si la suite des données d'origine est supposée α -mélangeante, nous ne savons pas s'il est de même pour la suite des données observées.

Dans la suite de ce chapitre, $\|\cdot\|$ désignera la norme euclidienne de \mathbb{R}^d .

3.2.2.1 Convergence presque sûre

Soit $D \subset \mathbb{R}^d$ un ensemble compact. Posons

$$\hat{D} = \{x : \text{il existe } y \in D, \|x - y\| < \zeta\} \text{ pour un } \zeta > 0.$$

Soit $L : \mathbb{N} \rightarrow \mathbb{R}$ une fonction décroissante telle que $\sum_{n=3}^{\infty} n^{-1}(\log(n))^{-1}(L(n))^{-1} < \infty$. Cette condition est vérifiée si on choisit $L(n) = (\log \log(n))^{1+\delta}$ avec $\delta > 0$.

Nous commençons par présenter la convergence presque sûre de l'estimateur \hat{v}_n . Nous avons besoin des hypothèses suivantes

(A3.0) $h_n \rightarrow 0$ et $\frac{\log n}{nh_n^d} \rightarrow 0$ quand $n \rightarrow +\infty$;

(A3.1) Soit $p \geq 1$ un entier. K_d est une fonction lipschitzienne avec $K_d(t) = 0$ pour $\|t\| > 1$, $\int_{\mathbb{R}^d} K_d(t) dt = 1$ et

$$\int_{\mathbb{R}^d} \prod_{j=1}^l z_{i_j} K_d(z_1, \dots, z_d) = 0 \quad (i_1, \dots, i_l \in \{1, \dots, d\}, l = 1, \dots, p-1).$$

(A3.2) Pour tout entier assez grand j , la densité jointe v_j de X_1 et X_{1+j} existe et pour un $\delta' > 0$, $v_j(x, y) \leq C_1, \forall x, y \in \hat{D} : \|x - y\| < \delta'$, $v(x) \leq C_2, \forall x \in \hat{D}$, pour certaines constantes positives C_1, C_2 .

Théorème 3.3. (Liang et al. (2009)) Supposons que $\alpha(k) = O(k^{-r})$ pour $r > 3$, et que les hypothèses **(A3.0)**-**(A3.3)** soient vérifiées. Supposons de plus que v admet des dérivées partielles d'ordre p bornées dans \hat{D} . Alors,

$$\sup_{x \in D} |\hat{v}_n(x) - v(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee w_n \vee h_n^p \right\} \mathbf{P} - p.s, \text{ lorsque } n \rightarrow \infty, \quad (3.53)$$

où $w_n = \left(n^{1-r} h_n^{-d(r+2)} (\log n)^{r+1} L(n) \right)^{1/(r+d+1)} \rightarrow 0$ quand $n \rightarrow \infty$. De plus, si on choisit $h_n = O \left(\left(\frac{\log n}{n} \right)^{1/(2p+d)} \right)$, nous obtenons la vitesse de convergence optimale

$$\sup_{x \in D} |\hat{v}_n(x) - v(x)| = O \left(\left(\frac{\log n}{n} \right)^{p/(2p+d)} \right) \mathbf{P} - p.s, \text{ lorsque } n \rightarrow \infty, \quad (3.54)$$

pourvu que $r > 3 + d + 3d/p$.

Maintenant, nous donnons le résultat concernant la convergence presque sûre de l'estimateur \hat{m}_n . Nous avons besoin des hypothèses **(A3.0)**, **(A3.1)** et des hypothèses **(A'3.2)** et **(A3.3)** suivantes (**(A'3.2)** est une version légèrement modifiée de **(A3.2)**) :

(A'3.2) Pour tout entier assez grand j , la densité jointe v_j de X_1 et X_{1+j} existe et pour un $\delta' > 0$,

$$v_j(x, y) \leq C_1, \quad \forall x, y \in \hat{D} : \|x - y\| < \delta',$$

$$0 < C_3 \leq v(x) \leq C_2, \quad \forall x \in \hat{D}, \quad \text{pour certaines constantes positives } C_1, C_2, C_3.$$

(A3.3) Pour $\gamma > 0$, $\mathbf{E}(Y_i^\gamma | X_i = x) \leq C_4, \forall x \in \mathbb{R}^d$,

$$|\mathbf{E}(Y_i Y_j | X_i = x, X_j = y)| \leq C_5, \quad \forall x, y \in \hat{D} : \|x - y\| < \delta''$$

pour certaines constantes positives δ'', C_4, C_5 .

Théorème 3.4. (*Liang et al. (2009)*) Supposons que $\alpha(k) = O(k^{-r})$ pour $r > 3$, et que les hypothèses **(A3.0)**, **(A3.1)**, **(A'3.2)** et **(A3.3)** sont vérifiées. Supposons aussi que v et m admettent des dérivées partielles d'ordre p bornées dans \hat{D} .

(i) Si $\gamma \geq 2(r-1)/(r-2)$, alors

$$\sup_{x \in \hat{D}} |\hat{m}_n(x) - m(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee w_n \vee h_n^p \right\} \mathbf{P} - p.s. \text{ lorsque } n \rightarrow \infty, \quad (3.55)$$

$$\text{où } w_n = \left(n^{\gamma(1-r)+r+1} h_n^{-d\gamma(r+2)} (\log n)^{(\gamma+1)(r+1)} (L(n))^{\gamma+r+1} \right)^{1/(\gamma(r+d+1))} \rightarrow 0 \text{ quand } n \rightarrow \infty.$$

(ii) Si, en plus, $|Y_i| \leq C_6 < \infty$ p.s. par une certaine constante C_6 , alors $\sup_{x \in \hat{D}} |\hat{m}_n(x) - m(x)|$ a la même vitesse de convergence que celle donnée dans (3.53) et (3.54) du Théorème 3.3.

Remarque 3.3. les deuxièmes termes donnés dans la vitesse de convergence de (3.53) et (3.55) sont additionnels si on compare par rapport aux résultats correspondants dans le cas des suites de variables aléatoires indépendantes donnés dans le Théorème 3.1 et Lemme 3.2. Ils décrivent l'effet de la dépendance de la suite observée $\{(X_i, Y_i, T_i); 1 \leq i \leq n\}$.

Pour les preuves des résultats, nous renvoyons à l'article de Liang et al. (2009).

3.2.2.2 Normalité asymptotique

Soit $\mathcal{U}(x)$ un voisinage de x et $\text{supp}(v) = \{x \in \mathbb{R}^d | v(x) > 0\}$. La normalité asymptotique de l'estimateur \hat{v}_n repose sur les hypothèses suivantes :

(B3.0) $h_n \rightarrow 0$, $nh_n^d \rightarrow +\infty$ et $nh_n^{d+4} \rightarrow 0$ quand $n \rightarrow +\infty$;

(B3.1) Le noyau K_d est une fonction bornée de \mathbb{R}^d à support compact et

$$\int_{\mathbb{R}^d} K_d(x) dx = 1, \quad \int_{\mathbb{R}^d} (x_1 + \dots + x_d) K_d(x) dx = 0.$$

(B3.2) La suite $\alpha(k)$ satisfait

(i) Pour tout $q = q_n$ tel que $q = o(\sqrt{nh_n^d})$, $\sqrt{nh_n^{-d}} \alpha(q) \xrightarrow[n \rightarrow +\infty]{} 0$;

(ii) Il existe $\gamma > 2$ et $\delta > 1 - 2/\gamma$ tel que $\sum_{l=1}^{\infty} l^{\delta} [\alpha(l)]^{1-2/\gamma} < \infty$.

(B3.3) Pour tout entier $j \geq 1$, la densité conditionnelle jointe v_j^* de X_1 et X_{1+j} existe et pour un $\delta' > 0$,

$$v_j^*(x, y) \leq C_1, \quad \forall x, y \in \mathbb{R}^d : \|x - y\| < \delta',$$

$$v(x) \leq C_2, \quad \forall x \in \mathbb{R}^d, \quad \text{pour certaines constantes positives } C_1, C_2.$$

Théorème 3.5. (Liang (2011)) Soit $\alpha(k) = O(k^{-r})$ pour $r > 3$ et $x \in \text{supp}(v)$. Supposons que les hypothèses **(B3.0)**-**(B3.3)** sont vérifiées. Si $\sup_{u \in \mathcal{U}(x)} \left| \frac{\partial^2 v}{\partial u_i \partial u_j}(u) \right| < \infty$ pour $i, j = 1, \dots, d$, alors

$$\sqrt{nh_n^d} [\hat{v}_n(x) - v(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \alpha \kappa \Sigma_2(x)),$$

avec $\Sigma_2(x)$ est définie en (3.24) et $\kappa = \int_{\mathbb{R}^d} K_d^2(z) dz < +\infty$.

Nous présentons maintenant le résultat sur la normalité asymptotique de l'estimateur \hat{m}_n . Nous avons besoin des hypothèses supplémentaires suivantes :

(B3.4) $\mathbf{E}(Y_i^\gamma | X_i = x, X_j = y) \leq C_3, \forall i, j$ et $x, y \in \mathbb{R}^d : \|x - y\| < \delta''$ pour certaines constantes positives δ'', C_3 et γ est le même que dans **(B3.2)**.

(B3.5) $f(., .)$ est deux fois différentiable par rapport à la première composante et

$$\int_{\mathbb{R}} \sup_{u \in \mathcal{U}(x)} \left| y \frac{\partial^{i_1 + \dots + i_d}}{\partial u_1^{i_1} \dots \partial u_d^{i_d}} f(u, y) \right| dy < \infty \quad \text{pour } i_1 + \dots + i_d \leq 2, \quad \text{où } x \in \text{supp}(v).$$

(B3.6) $n^{-1} h_n^{-d(1+4/\gamma)} = O(1)$, où γ est le même que dans **(B3.2)**.

Théorème 3.6. (Liang (2011)) Soit $\alpha(k) = O(k^{-r})$ pour $r > \max(3, \gamma^2/(\gamma - 2))$ et $x \in \text{supp}(v)$. Supposons que les hypothèses **(B3.0)**-**(B3.6)** sont vérifiées. Si $\sup_{u \in \mathcal{U}(x)} \left| \frac{\partial^2 v}{\partial u_i \partial u_j}(u) \right| < \infty$ pour $i, j = 1, \dots, d$, alors

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma_2(x)),$$

avec

$$\sigma^2(x) := \frac{\alpha \kappa [\Sigma_0(x) v^2(x) - 2 \Sigma_1(x) \psi(x) v(x) + \Sigma_2(x) \psi^2(x)]}{v^4(x)}.$$

Nous renvoyons aussi ici à l'article de Liang (2011) pour les preuves de ces résultats.

Chapitre 4

Vitesse de convergence de l'estimateur de la fonction de régression pour des données tronquées à gauche et associées

Dans ce chapitre, nous établissons la convergence uniforme presque sûre sur un compact avec vitesse de l'estimateur à noyau de la fonction de régression dans le cas des données tronquées à gauche et présentant une forme de dépendance de type associé. Nous illustrons nos résultats théoriques par des simulations. Nous présentons nos résultats dans leur version originale tels qu'ils sont parus dans la revue *Journal of NonParametric Statistics*.

4.1 Introduction

Let Y be a real random variable (rv) of interest with a continuous distribution function (df) F , and X an \mathbb{R}^d -valued random vector of covariates with a joint df V and a joint density v . We wish to estimate Y given X using a regression-based approach. This means looking for a function which realizes the minimum of the mean squared error (MSE). The regression function that achieves this minimum is defined on \mathbb{R}^d by the conditional expectation of Y given $X = x$ that is

$$m(x) := \mathbb{E}(Y | X = x).$$

In the case of complete data there is a vast literature devoted to the study of the nonparametric kernel estimator of $m(\cdot)$. Far from being exhaustive, we can quote Walk (2002) and the references therein.

Nevertheless, in many survival practical applications, it happens that one is not able to observe a subject's entire lifetime. The subject may leave the study, may survive to the closing date, or may enter the study at some time after its lifetime has started. The most current forms of such incomplete data are censorship and truncation. The model studied here is based on the random left truncated

(RLT) data, where the observation (X, Y) is interfered by another independent rv T such that the random quantities Y, X and T are observable only if $Y \geq T$.

The RLT model is originally appeared in astronomy and economics (Woodroffe (1985), Chen et al. (1995)), then extended to several domains as epidemiology, demographics, reliability testing and actuarial (Wang et al. (1986) and Tsai et al. (1987)). For example, in an AIDS study (Kalbeisch and Lawless (1989)), let X be the infection time where 1 represents January 1978 and let T be the incubation time in months for people who were infected by contaminated blood transfusions and developed AIDS by 1 July 1986. Since the total study period is 102 months only individuals with $X + T < 102$ were included in the sample. Then, letting $Y = 102 - X$ yields the model described : (Y, T) is observed only if $T < Y$. Another example is that of a retirement centre (Klein and Moeschberger (2003)), where subjects are observed only if they live long enough to enter the centre. The lifetime Y is then left truncated by the retirement house entry age, T . People who enter the centre earlier may get better medical attention and therefore live longer. On the other hand, people with poor health and shorter expected lifetime may retire earlier.

In the i.i.d. case, Ould Saïd and Lemdani (2006) constructed a nonparametric kernel estimator of the regression function $m(\cdot)$ under RLT model. They established its strong uniform consistency as well as its asymptotic properties. The aim of this paper is to extend some of their results to the case of dependent data. Two types of dependency are widely used in literature : mixing and association.

The α -mixing condition, also called strong mixing, is the weakest among mixing conditions known in the literature. Many stochastic processes satisfy the α -mixing condition, see for example, Doukhan (1994) and Carrasco et al (2007).

For truncated data, under α -mixing condition, the strong convergence of the estimator of the regression function defined by Ould Saïd and Lemdani (2006) is treated in Liang et al. (2009) and its asymptotic normality is established later in Liang (2011).

In this paper, we focus on the concept of association which has been introduced and defined by Esary et al. (1967). A set of finite family of rvs $Y = (Y_1, \dots, Y_N)$ is said associated if

$$cov(f(Y), g(Y)) \geq 0$$

for all nondecreasing functions f and g from \mathbb{R}^N to \mathbb{R} for which this covariance exists. An infinite family is said to be associated if every finite subfamily is associated.

It is of interest to note that association and mixing define two distinct but not disjoint classes of processes (see Doukhan and Louhichi (1999)). In the linear time series framework, the sequence X_k defined by $X_k = \sum_{j=0}^{\infty} a_j v_{k-j}$ (where $(v_i)_{-\infty < i < +\infty}$ is a sequence of i.i.d. rvs with mean zero and variance σ^2), is associated if $a_j \geq 0$. On the other hand Pham and Tran (1985) showed that $(X_k)_{k \geq 0}$ is α -mixing under suitable conditions on a_j . In particular, Andrew (1984) showed that when $(v_i)_{-\infty < i < +\infty}$ is a sequence of i.i.d. Bernoulli rvs and $a_j = \varsigma$ with $0 < \varsigma < \frac{1}{2}$, the sequence $(X_k)_{k \geq 0}$ is not α -mixing, whereas it is still associated.

For more details on the concept of association, we refer the reader to Bulinski and Shashkin (2007). In that book, the reader can find some results and examples related to associated random sequences and random fields.

In the complete associated data case, there is a vast literature devoted to the study of the nonparametric kernel estimation and many papers deal with density estimation. We cite only a few of them as Bagai et al. (1995) who have obtained the strong uniform consistency of the kernel density estimator. They also proposed an estimator of the survival function and established its consistency. Roussas (2000) has established the asymptotic normality of the usual kernel estimate of the marginal probability density function. Douge (2007) has stated a new exponential inequality and has derived a uniform almost sure rate of convergence over compact sets for the kernel density estimator.

In the incomplete data case, for associated rvs there are no much works done for this kind of model. Under random right censoring one, we can cite Cai and Roussas (1998) who have established uniform strong consistency and asymptotic normality of the Kaplan-Meier estimator. Ferrani et al. (2016) established the strong uniform consistency of the kernel estimator of the underlying density function and the almost sure convergence of a smooth kernel mode estimator under right censored model. Under RLT model, Guessoum et al. (2012) established the strong uniform convergence with a rate of the Lynden-Bell estimator.

To the best of our knowledge, the problem of estimating the regression function under association and truncation has not been addressed in the literature. The goal of this paper is to establish the strong uniform convergence with a rate for the kernel regression estimate, under RLT model when the variable of interest Y and the multivariate covariates X are associated.

This paper is organized as follows. In Section 2, we recall some results stated in RLT model with the estimators studied in the current work. In Section 3, we list the assumptions and give our main results. In Section 4, some simulations are given. The proofs of the main results are detailed in the appendix with some preliminary lemmas.

4.2 Model and estimators

Let Y be a bounded real rv defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with a continuous df F and T a real rv independent from Y , defined on the same probability space with a Lipschitz df G .

In the following, $\{Y_i; i = 1, \dots, N\}$ and $\{T_i; i = 1, \dots, N\}$ denote respectively, a strictly stationary associated sequence and an i.i.d. sequence of N copies of Y and T , where the sample size N is fixed but unknown.

In the RLT model, as mentioned above, the rv of interest Y and the truncated rv T are observable only when $Y \geq T$, whereas nothing is observed if $Y < T$. Then without possible confusion, we still denote $\{(Y_i, T_i); i = 1, \dots, n\}$ ($n \leq N$) the actually observed sample. Note that as the original sequence of interest is associated, the observed one is associated (by property (P1) in Esary et al.

(1967)) and the observed sequence of truncation is also i.i.d. (by proposition 2.1 in Lemdani and Ould Saïd (2007)).

Let $\alpha := \mathbb{P}(Y \geq T)$ be the probability to observe at least one pair from (Y, T) . As a consequence of truncation, the size n of the actually observed sample is a $\text{Bin}(N, \alpha)$ rv. We suppose hereafter that $\alpha > 0$ otherwise no data can be observed. Since N is unknown and n known (although random), our results will not be stated with respect to the probability measure \mathbb{P} (related to the N -sample) but will involve the probability \mathbf{P} (related to the n -sample) defined as $\mathbf{P}(\cdot) = \mathbb{P}(\cdot | Y \geq T)$. In the same way \mathbb{E} and \mathbf{E} will denote the expectation operators related to \mathbb{P} and \mathbf{P} , respectively.

Following Stute (1993), the joint \mathbf{P} -distribution of an observed (Y, T) is given by

$$\begin{aligned} H^*(y, t) &:= \mathbf{P}\{Y \leq y, T \leq t\} \\ &= \mathbb{P}\{Y \leq y, T \leq t | Y \geq T\} \\ &= \frac{1}{\alpha} \int_{-\infty}^y G(t \wedge u) dF(u) \end{aligned}$$

where $t \wedge u := \min(t, u)$. The marginal distributions of Y and T , respectively, are defined by

$$F^*(y) := H^*(y, \infty) = \frac{1}{\alpha} \int_{-\infty}^y G(u) dF(u), \quad \text{and} \quad G^*(t) := H^*(\infty, t) = \frac{1}{\alpha} \int_{-\infty}^{\infty} G(t \wedge u) dF(u)$$

which can be estimated by

$$F_n^*(y) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq y\}} \quad \text{and} \quad G_n^*(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{T_i \leq t\}}$$

respectively, where \mathbb{I}_A denotes the indicator function of the set A . Let $C(\cdot)$ be defined by

$$\begin{aligned} C(y) &:= \mathbf{P}\{T \leq y \leq Y\} = \mathbb{P}\{T \leq y \leq Y | Y \geq T\} \\ &= \frac{1}{\alpha} G(y)(1 - F(y)) = G^*(y) - F^*(y) \end{aligned}$$

with empirical estimator

$$C_n(y) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{T_i \leq y \leq Y_i\}} = G_n^*(y) - F_n^*(y).$$

The well-known nonparametric estimators of F and G in RLT model, proposed by Lynden-Bell (1971) are

$$\begin{aligned} F_n(y) &:= 1 - \prod_{i: Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)} \right] \\ G_n(t) &:= \prod_{i: T_i > t} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)} \right]. \end{aligned} \tag{4.1}$$

Here, for any df W we define

$$a_W := \inf\{u : W(u) > 0\} \text{ and } b_W := \sup\{u : W(u) < 1\} \quad (4.2)$$

as the endpoints of the W support. Woodroffe (1985) pointed out that F and G can be completely estimated only if

$$a_G \leq a_F, \quad b_G \leq b_F \text{ and } \int_{a_F}^{\infty} \frac{dF}{G} < \infty.$$

For the unknown probability $\alpha := \mathbb{P}(Y \geq T)$, which cannot be classically estimated by $\hat{\alpha}_n := \frac{n}{N}$ since N is unknown, He and Yang (1998) proposed the estimator

$$\alpha_n(y) := \frac{G_n(y)(1 - F_n(y))}{C_n(y)} =: \alpha_n. \quad (4.3)$$

The authors proved that α_n does not depend on y and its value can be obtained for any y such that $C_n(y) \neq 0$.

Now, let $\{(X_i, Y_i, T_i); i = 1, \dots, n\}$ be the n triplets observed among the N ones such that $Y_i \geq T_i$ and T independent of (X, Y) where $\{(X_i, Y_i); i = 1, \dots, n\}$ are supposed strictly stationary and associated. Then we first consider the joint \mathbf{P} -distribution of (X, Y, T)

$$\begin{aligned} L^*(x, y, t) &:= \mathbf{P}\{X \leq x, Y \leq y, T \leq t\} \\ &= \mathbb{P}\{X \leq x, Y \leq y, T \leq t | Y \geq T\} \\ &= \frac{1}{\alpha} \int_{u \leq x} \int_{a_G \leq w \leq y} G(w \wedge t) F(du, dw), \end{aligned}$$

the joint \mathbf{P} -df $F_{X,Y}^*(\cdot, \cdot)$ of an observed (X, Y) is given by

$$F_{X,Y}^*(x, y) := L^*(x, y, \infty) = \frac{1}{\alpha} \int_{u \leq x} \int_{a_G \leq w \leq y} G(w) F_{X,Y}(du, dw),$$

which gives

$$F_{X,Y}(dx, dy) = \frac{\alpha}{G(y)} F_{X,Y}^*(dx, dy), \text{ for } y \geq a_G. \quad (4.4)$$

By integrating over y , we get the df of X

$$V(x) = \alpha \int_{u \leq x} \int_{y \geq a_G} \frac{1}{G(y)} F^*(du, dy).$$

Following Ould Saïd and Lemdani (2006), an estimator of $V(x)$ is given by

$$\tilde{V}_n(x) := \frac{\alpha}{n} \sum_{i=1}^n \frac{1}{G(Y_i)} \mathbb{I}_{\{X_i \leq x\}}. \quad (4.5)$$

Note that in (4.5) and in the sequel, the sum is taken only over the i 's such that $G(Y_i) \neq 0$. Thus, (4.5) yields the kernel density estimator

$$\tilde{v}_n(x) := \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{1}{G(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right), \quad (4.6)$$

where $K_d : \mathbb{R}^d \rightarrow \mathbb{R}$ is a smooth kernel function and h_n is a positive bandwidth sequence that tends to zero as $n \rightarrow \infty$. Observe that the regression function $m(x)$ can be written as

$$m(x) = \frac{\psi(x)}{v(x)}, \quad (4.7)$$

where $\psi(x) = \int_{\mathbb{R}} y f_{X,Y}(x, y) dy$ and $f_{X,Y}(\cdot, \cdot)$ is the joint density of (X, Y) . Then, assuming that $\tilde{v}_n(x) > 0$ for all $x \in \mathbb{R}^d$, it is well known that the kernel estimator of the regression function $m(x)$ under RLT model is given by

$$\tilde{m}_n(x) := \frac{\tilde{\psi}_n(x)}{\tilde{v}_n(x)},$$

where

$$\tilde{\psi}_n(x) := \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right). \quad (4.8)$$

As in practice α and G are usually unknown, we replace them by their consistent estimators G_n and α_n defined in (5.1) and (5.2), respectively. Thus by plug-in method one can define the feasible kernel estimate for $m(\cdot)$ by

$$\hat{m}_n(x) := \frac{\hat{\psi}_n(x)}{\hat{v}_n(x)}, \quad (4.9)$$

where

$$\begin{aligned} \hat{\psi}_n(x) &:= \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \\ \hat{v}_n(x) &:= \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \end{aligned}$$

4.3 Assumptions and main results

Throughout this paper we assume that $a_G < a_F$ and $b_G \leq b_F$. Let D be a compact set which is included in $\Xi = \{x \in \mathbb{R}^d / v(x) > \delta > 0\}$ for a real $\delta > 0$, and let us define

$$\theta_{i,j} := \sum_{k=1}^d \sum_{l=1}^d \text{cov}(X_{i,k}, X_{j,l}) + 2 \sum_{k=1}^d \text{cov}(X_{i,k}, Y_j) + \text{cov}(Y_i, Y_j), \quad (4.10)$$

where $X_{i,k}$ is the k -th component of X_i . We will make use of the following assumptions gathered here for easy reference.

- (H) The bandwidth h_n satisfies : $h_n \rightarrow 0$, $nh_n^d \rightarrow +\infty$ and $\frac{\log^5 n}{nh_n^d} \rightarrow 0$ as $n \rightarrow +\infty$;
- (A) $\int_{\mathbb{R}} \frac{dF(z)}{G^2(z)} < +\infty$;
- (K1) The kernel K_d is a bounded probability density with compact support;
- (K2) $\int_{\mathbb{R}^d} z_i K_d(z) dz = 0$ for all $i = 1, \dots, d$ and $\int_{\mathbb{R}^d} |z_1^{i_1} \dots z_d^{i_d}| K_d(z) dz < +\infty$ for $i_1 + \dots + i_d = 2$;

- (K3) $\int_{\mathbb{R}^d} |z_1 + \dots + z_d| K_d^2(z) dz < +\infty$ and $\int_{\mathbb{R}^d} K_d^2(z) dz < +\infty$;
- (K4) K_d is Hölder continuous with exponent $\beta > 0$;
- (R) The covariance term defined by $\rho(s) := \sup_{|i-j| \geq s} \theta_{i,j}$ for $s > 0$, satisfies $\rho(s) \leq \gamma_0 e^{-\gamma s}$ for some positive constants γ_0 and γ ;
- (D1) The function $\psi(\cdot)$ is bounded, twice differentiable with $\sup_{x \in D} \left| \frac{\partial^k \psi}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ for $i, j = 1, \dots, d$ and $k = 1, 2$;
- (D2) The function $\psi_1(x) := \int_{\mathbb{R}} \frac{y^2}{G(y)} F(x, dy)$ is bounded and continuously differentiable with $\sup_{x \in D} \left| \frac{\partial \psi_1}{\partial x_i}(x) \right| < +\infty$ for $i = 1, \dots, d$,
- (D3) The probability density $v(\cdot)$ is bounded and twice differentiable with $\sup_{x \in D} \left| \frac{\partial^k v}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ for $i = 1, \dots, d$ and $k = 1, 2$;
- (D4) The conditional joint density $v_{i,j}^*$ of (X_i, X_j) exists and $\sup_{x_1, x_2 \in D} |v_{i,j}^*(x_1, x_2)| < +\infty$;

Remark 4.1. Assumptions (H), (K1), (K4) and (D1) are very common in functional estimation both in independent and dependent cases. Assumptions (A) and (R) imply assumptions of Guessoum et al. (2012) and are needed, among other, to use their results. Furthermore, Assumption (R) quantifies a geometric decay of the covariance terms needed to establish a uniform almost sure convergence type. This condition is similar to the one used in Bulinski (1996), Douge (2007) and Doukhan and Neumann (2007) (without covariate). Assumptions (D3) and (D4) are technical and are used to compute covariances.

The following two theorems give the uniform asymptotic expression of the fluctuation terms for the two estimators $\tilde{\psi}_n(x)$ and $\tilde{v}_n(x)$ defined respectively in (5.5) and (5.6).

Theorem 4.1. Under assumptions (K1)-(K4), (D1)-(D2), (D4), (R) and (H) we have

$$\sup_{x \in D} \left| \tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right| = O \left(\sqrt{\frac{\log n}{nh_n^d}} \right) \text{ P- a.s, as } n \rightarrow +\infty.$$

Theorem 4.2. Under assumptions (K1)-(K4), (D3)-(D4), (R) and (H) we have

$$\sup_{x \in D} \left| \tilde{v}_n(x) - \mathbf{E}(\tilde{v}_n(x)) \right| = O \left(\sqrt{\frac{\log n}{nh_n^d}} \right) \text{ P- a.s, as } n \rightarrow +\infty.$$

The proofs of Theorem 4.1 and Theorem 4.2 are mainly based on a Bernstein-type inequality due to Doukhan and Neumann (2007) recalled in Lemma 4.1 in the appendix.

Now, to state our main result observe that from (4.7) and (4.9) we have

$$\begin{aligned}
\hat{m}_n(x) - m(x) &= \frac{\hat{\psi}_n(x)}{\hat{v}_n(x)} - \frac{\psi(x)}{v(x)} \\
&= \left(\frac{\hat{\psi}_n(x)}{\hat{v}_n(x)} - \frac{\tilde{\psi}_n(x)}{\hat{v}_n(x)} \right) + \left(\frac{\tilde{\psi}_n(x)}{\hat{v}_n(x)} - \frac{\mathbf{E}(\tilde{\psi}_n(x))}{\hat{v}_n(x)} \right) + \left(\frac{\mathbf{E}(\tilde{\psi}_n(x))}{\hat{v}_n(x)} - \frac{\psi(x)}{\hat{v}_n(x)} \right) \\
&\quad + \psi(x) \frac{v(x) - \hat{v}_n(x)}{\hat{v}_n(x)v(x)},
\end{aligned} \tag{4.11}$$

then application of Theorem 4.1 and Theorem 4.2 leads to

Theorem 4.3. *Under assumptions (A), (K1)-(K4), (D1)-(D4), (R) and (H) we have*

$$\sup_{x \in D} |\hat{m}_n(x) - m(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee \left(\frac{\log \log n}{n} \right)^\theta \vee h_n^2 \right\} \mathbf{P}\text{-a.s. as } n \rightarrow \infty.$$

where $0 < \theta < \frac{\gamma}{2\gamma+6+3\kappa/2}$ for any real $\kappa > 0$.

Remark 4.2. *The rate obtained by Ould Saïd and Lemdani (2006) in the independent case for $d = 1$ (their Theorem 4.1) is $O \left(\sqrt{\frac{\log n}{n^{1-2\lambda}h_n^2}} \vee h_n \right)$ where λ is defined in their assumptions (A1) and (A2). If we compare it with our result, we find an added term depending on θ , which is due to the association effect. Note that our rate is slightly better than theirs due to the symmetric condition on the kernel K_d in our assumption (K2). In α -mixing case, Liang et al. (2009) obtained the rate $O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee \hat{w}_n \vee h_n^p \right\}$ where p denotes the order of the kernel K_d and \hat{w}_n quantifies the α -mixing effect. Then if we take $p = 2$ (symmetric kernel), their rate becomes similar to ours.*

4.4 Simulations study

The main purpose of this section is to investigate the finite sample performance of the estimator $\hat{m}_n(x)$ in the case $d = 1$, for some particular regression functions $m(x)$. For that, we generate data as follows

1. The covariate X :
 - Generate $N + 1$ i.i.d. $\mathcal{N}(0, 1)$ rvs $\{W_t; t = -1, 0, \dots, N - 1\}$.
 - Given W_t , generate the associated sequence $\{X_t; t = 1, \dots, N\}$ by $X_t = \exp \left[\frac{1}{2} (W_{t-1} + W_{t-2}) \right]$. This model comes from Chaubey et al. (2011) where it is shown that the autoregressif model of order 2 defined by $S_t = \frac{W_{t-1}}{2} + \frac{W_{t-2}}{2}$ is associated. Furthermore, as the exponential is a nondecreasing function so, from property (P4) of Esary et al. (1967), X_t are associated variables.

2. The interest variable Y :

- Generate N i.i.d. rvs $\{\varepsilon_t; t = 1, \dots, N\}$ with distribution specified below.
- Set $Y_t = m(X_t) + \varepsilon_t$ with $m(\cdot)$ the regression function.

3. The truncated variable T :

- Generate independently the i.i.d. rvs $\{T_t\}$ with distribution $\mathcal{N}(\mu, 1)$ (μ is adapted in order to obtain different values of truncation).

4. The observed data :

- We keep the n observations $\{(X_i, Y_i, T_i); i = 1, \dots, n\}$ of the triplet of rvs (X, Y, T) satisfying the condition $Y_i \geq T_i$.

The supremum is taken over a compact set $D = [a, b]$ for which we consider a subdivision ϖ defined by

$$\varpi = \{a_0 = a, x_1, \dots, x_J = b\}, \quad J \geq 1,$$

and we calculate the values of the estimator $\hat{m}_n(x)$, at each point a_j of ϖ by choosing a Gaussian kernel K and a bandwidth $h_n = O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{3}}\right)$.

This section is divided in three parts. In the first one, we focus on the performance of the estimator when the model $m(x)$ is linear, for different values of α (the rate related to the (no) truncation) and n (the size of the observed sample). In the second part, we compare the performance of the estimator by considering two different distributions of the errors ε_i and thereafter we propose to compare $\hat{m}_n(x)$ to the restricted conditional mean survival time (RCMST) estimate. Finally some nonlinear regression functions are chosen for highlighting the robustness of our estimator.

4.4.1 Performances of $\hat{m}_n(\cdot)$ under linear model

We generate the data with observed sample sizes $n = 50, 100$ and 300 , respectively, from the linear model given by

$$Y_i = 2X_i + 1 + \varepsilon_i, \quad i = 1, \dots, N,$$

where ε_i comes from $\mathcal{N}(0, 0.2)$. Hence, the true underlying regression function is $m(x) = 2x + 1$. In Tables 4.1 and 4.2, we take $\alpha \approx 60\%$ and 80% (obtained for $\mu = 2.5$ and $\mu = 2$ respectively) and we report the bias, variance and MSE of the estimator \hat{m}_n at the bounds $x = 0$ and $x = 2$ and at a middle point $x = 1$, based on $B = 1000$ replications.

From Tables 4.1 and 4.2 it can be seen that :

- The estimator does not have the same behaviour at the bounds of the support and inside this one. In particular in the left we notice, as it is well known, that the estimator has boundary effects which may be attributed to the fact that the support of the kernel exceeds the available range of data.

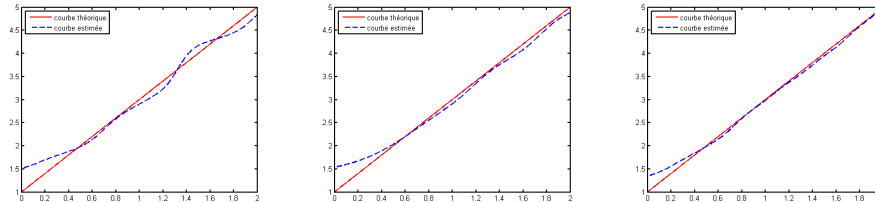
TABLE 4.1 – MSE, bias and variance of \hat{m}_n for $\alpha \approx 60\%$.

	n=50			n=100			n=300		
	MSE	Bias	Var	MSE	Bias	Var	MSE	Bias	Var
$x = 0$	0.3955	0.6028	0.0321	0.2754	0.5053	0.0201	0.1547	0.3787	0.0114
$x = 1$	0.0023	-0.0171	0.0020	0.0008	-0.0117	0.0007	0.0002	-0.0059	0.0002
$x = 2$	0.0293	-0.0998	0.0193	0.0084	-0.0191	0.0080	0.0027	-0.0113	0.0026

TABLE 4.2 – MSE, bias and variance of \hat{m}_n for $\alpha \approx 80\%$.

	n=50			n=100			n=300		
	MSE	Bias	Var	MSE	Bias	Var	MSE	Bias	Var
$x = 0$	0.2073	0.4212	0.0299	0.1182	0.3250	0.0126	0.0901	0.2971	0.0018
$x = 1$	0.0021	-0.0121	0.0019	0.0009	-0.0100	0.0008	0.0002	-0.0046	0.0002
$x = 2$	0.0102	-0.0678	0.0056	0.0044	-0.0324	0.0034	0.0019	-0.0223	0.0014

- Each of the MSE, bias and variance decrease when n increases, so the quality of the estimator is better for high observed sample sizes which is confirmed by Figure 4.1, where we plot $m(\cdot)$ and its estimator $\hat{m}_n(\cdot)$ with $\alpha \approx 80\%$ for $n = 50, 100$ and 300 .

FIGURE 4.1 – $m(\cdot)$ and $\hat{m}_n(\cdot)$ with $\alpha \approx 80\%$ and $n = 50, 100$ and 300 , respectively.

Hereafter, to take into account all points in D , we calculate the MSE along the interval $D = [0, 2]$ by taking the median over $x \in [0, 2]$. The results are given in Table 4.3 for $\alpha \approx 60\%, 70\%$, and 80% (in this case the different values of α are obtained for the same parameter $\mu = 2.5$, in order to make some comparisons using the same law for the simulated variables) for $n = 50, 100$ and 300 . As the MSE gives a pointwise error, we extend the study to the global behaviour of the estimator by using the mean integrated squared error (MISE). The results are given in Table 4.4 with $\alpha \approx 60\%, 70\%$, and 80% (obtained for $\mu = 2.5$) for $n = 50, 100$ and 300 . Tables 4.3 and 4.4 insure that the performance of the estimator is better for high sample size and great value of α . This is also confirmed by Figure 4.2, where we plot $m(\cdot)$ and its estimator $\hat{m}_n(\cdot)$ with $n = 300$ for $\alpha \approx 60\%, 70\%$ and 80% .

TABLE 4.3 – MSE's median of \hat{m}_n .

$\alpha(\%)$	n=50	n=100	n=300
60	0.0033	0.0012	3.0539×10^{-4}
70	0.0028	0.0009	3.0283×10^{-4}
80	0.0027	0.0006	2.7858×10^{-4}

TABLE 4.4 – The MISE of \hat{m}_n .

$\alpha(\%)$	n=50	n=100	n=300
60	0.0539	0.0305	0.0228
70	0.0341	0.0284	0.0128
80	0.0298	0.0198	0.0127

4.4.2 Some comparisons

- We compare the performances of $\hat{m}_n(x)$, via the MISE, based on $B = 1000$ replications, for two different distributions of the error ε_i by generating the data with observed sample sizes $n = 100$ and 300 , respectively, from the linear model given by

$$Y_i = X_i + \varepsilon_i, \quad i = 1, \dots, N,$$

where : (a) ε_i comes from $\mathcal{N}(0, 1)$ and (b) ε_i comes from $\mathcal{St}(2)$.

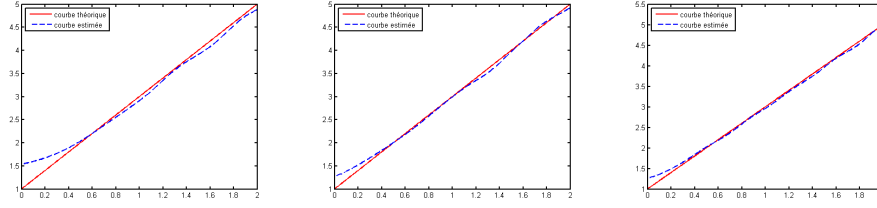
In Table 4.5, it can be seen that the estimation is most robust when the error distribution is normal. One of the reasons may be that in distribution (b) the errors ε_i are heavy-tailed.

- Using the same model with normal error, we compare $\hat{m}_n(x)$ to the RCMST estimate defined below. Recall that in the case of complete and positive data the regression function can be written as a conditional mean survival time that is

$$\mathbb{E}(Y | X = x) = \int_0^{+\infty} S(y|x) dy.$$

where $S(y|x)$ is the conditional survival function of Y given $X = x$. Under the RLT model, we define the RCMST as

$$RCMST = \int_{a_F}^{b_F} S(y|x) dy = \int_{a_F}^{b_F} [1 - F(y|x)] dy,$$

FIGURE 4.2 – $m(\cdot)$ and $\hat{m}_n(\cdot)$ with $n = 300$ and $\alpha \approx 60, 70$ and 80% , respectively.TABLE 4.5 – The MISE of \hat{m}_n .

$\alpha(\%)$	n	Distribution (a)	Distribution (b)
60	100	0.0591	0.9509
	300	0.0281	0.8092
70	100	0.0398	0.9291
	300	0.0194	0.7523
80	100	0.0368	0.9022
	300	0.0131	0.6993

where $F(y|x)$ is the conditional df of Y given $X = x$ and a_F , b_F are defined in (4.2). Then we propose to estimate $RCMST$ by

$$\bar{m}_n(x) = \int_{a_F}^{b_F} [1 - F_n(y|x)] dy,$$

where $F_n(y|x)$ is the estimator of $F(y|x)$ given in Lemdani et al. (2009) by

$$F_n(y|x) = \frac{\sum_{i=1}^n G_n^{-1}(Y_i) K_d \left(\frac{x - X_i}{h_n} \right) K_0 \left(\frac{y - Y_i}{h_n} \right)}{\sum_{i=1}^n G_n^{-1}(Y_i) K_d \left(\frac{x - X_i}{h_n} \right)},$$

with K_0 a smooth df defined on \mathbb{R} .

In order to compare the performance of the two estimators $\hat{m}_n(x)$ and $\bar{m}_n(x)$ through their MISE, we generate the data with observed sample sizes 100 and 300, respectively, from the model given by

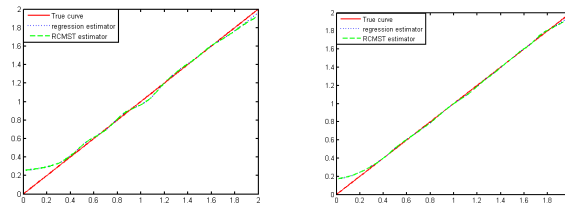
$$Y_i = X_i + \varepsilon_i, \quad i = 1, \dots, N.$$

where $\varepsilon_i \sim \mathcal{N}(0, 1)$.

In Table 4.6, we take $\alpha \approx 60\%$ and 80% (obtained for $\mu = 1$) and we report the MISE of $\hat{m}_n(x)$ and $\bar{m}_n(x)$ based on $B = 1000$ replications. We notice that the two estimators have similar performance which is confirmed by Figure 4.3 where we plot the true regression function $m(x) = x$ together with the estimators $\hat{m}_n(x)$ and $\bar{m}_n(x)$ for $\alpha \approx 80\%$, $n = 100$ and 300 .

TABLE 4.6 – The MISE of \hat{m}_n and $\bar{m}_n(x)$.

$\alpha(\%)$	n	$\hat{m}_n(x)$	$\bar{m}_n(x)$
60	100	0.0585	0.0591
	300	0.0279	0.0281
80	100	0.0368	0.0365
	300	0.0131	0.0132

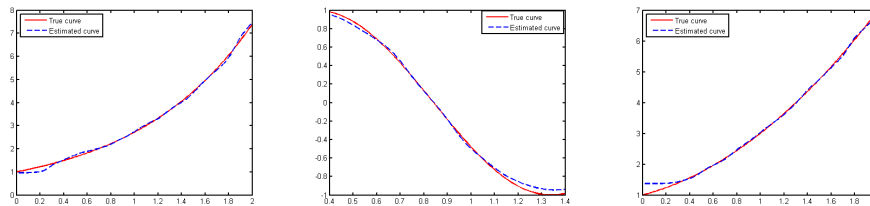
FIGURE 4.3 – $m(\cdot)$, $\hat{m}_n(\cdot)$ and $\bar{m}_n(\cdot)$ with $\alpha \approx 80\%$ and $n = 100, 300$ respectively.

4.4.3 Nonlinear model

We consider the case of nonlinear regression by choosing the following three models :

$$\begin{aligned}
 m(x) &= \exp(x) + \varepsilon_i, & \text{exponential,} \\
 m(x) &= \sin(\pi x + \frac{1}{2}) + \varepsilon_i, & \text{sinus,} \\
 m(x) &= x^2 + x + 1 + \varepsilon_i, & \text{parabolic.}
 \end{aligned}$$

with $D = [0, 2]$ for exponential and parabolic cases and $D = [0.4, 1.4]$ for sinus case. The error ε_i is taken normal, the size n equal to 300 and α equal to 80%.

FIGURE 4.4 – exponential, sinus, and parabolic with $n = 300$ and $\alpha \approx 80\%$

Here again, we can see the good performance of our estimator for nonlinear regression functions.

4.5 Conclusion

This paper has established the uniform strong consistency along with a rate of the kernel regression estimator over a real compact, when the variable of interest is subject to left truncation under association hypothesis. In doing so, a Bernstein-type inequality due to Doukhan and Neumann (2007) has been used. A large simulation study was conducted through which our estimator performance was highlighted in spite of well-known boundary effects of kernel estimation. Alternative methods such as local linear and k-nearest neighbor will be the subject of future research. These methods could improve the bias and the boundary effects. Comparisons of the studied estimator with the RCMST estimate have been made. Regarding this latter estimator, another possible estimation procedure is to consider the conditional product limit estimator defined in Akritas and LaValley (2005), under the assumption that Y is conditionally independent of T given X , for which theoretical results do not exist for associated data. Many other results remain to be established as asymptotic normality and more generally asymptotic behaviour of nonparametric functional estimator.

Appendix 1. Proofs

First, let

$$Z_i(x) := Z(X_i, Y_i; x) = \frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right). \quad (4.12)$$

We recall in the following lemma an exponential inequality stated in Doukhan and Neumann (2007), used in the proofs of Theorems 4.1 and 4.2.

Lemma 4.1. (*Doukhan and Neumann (2007), Theorem 1, p.880*) Suppose that X_1, \dots, X_n are real-valued rvs with zero mean, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that there exist constants $K, M, L_1, L_2 < +\infty, \mu, \nu \geq 0$ and a non-increasing sequence of real coefficients $(\delta(n))_{n \geq 0}$ such that, for all u -tuples (s_1, \dots, s_u) and all v -tuples (t_1, \dots, t_v) with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$, the following inequalities are fulfilled :

- a) $|\text{cov}(X_{s_1} \cdots X_{s_u}, X_{t_1} \cdots X_{t_v})| \leq K^2 M^{u+v-2} ((u+v)!)^\nu uv \delta(t_1 - s_u),$
- b) $\sum_{s=0}^{\infty} (s+1)^k (\delta(s)) \leq L_1 L_2^k (k!)^\mu, \forall k \geq 0,$
- c) $\mathbb{E}(|X_t|^k) \leq (k!)^\nu M^k, \forall k \geq 0.$

Then, for all $t \geq 0$,

$$\mathbb{P} \left(\sum_{i=1}^n X_i \geq t \right) \leq \exp \left(- \frac{t^2/2}{A_n + B_n^{1/(\mu+\nu+2)} t^{(2\mu+2\nu+3)/(\mu+\nu+2)}} \right) \quad (4.13)$$

where A_n can be chosen as any number greater than or equal to $\sigma_n^2 := \text{Var} \left(\sum_{i=1}^n X_i \right)$ and $B_n = 2(K \vee M) L_2 \left(\frac{2^{4+\mu+\nu} n K^2 L_1}{A_n} \vee 1 \right).$

Throughout the proofs, we denote by c (different) constants whose values are allowed to change. The following two lemmas are used to show that the process $Z_i(x)$ satisfies the conditions of Lemma 4.1, that will allow us to use the exponential inequality (4.13). In the first lemma, to lighten the notations, we note $Z_i(x)$ by Z_i .

Lemma 4.2. *Under assumption **(R)**, there exist constants $K, M, L_1, L_2 < +\infty, \mu, \lambda \geq 0$ such that for all $(s_1, \dots, s_u) \in \mathbb{N}^u$ and all $(t_1, \dots, t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \dots \leq s_u \leq t_1 \leq \dots \leq t_v \leq n$, we have*

- a) $\text{cov}(Z_{s_1} \cdots Z_{s_u}, Z_{t_1} \cdots Z_{t_v}) \leq K^2 M^{u+v-2} ((u+v)!)^\lambda uv (\rho(t_1 - s_u))^{\frac{d}{2d+2}},$
- b) $\sum_{s=0}^{\infty} (s+1)^k (\rho(s))^{d/(2d+2)} \leq L_1 L_2^k (k!)^\mu, \forall k \geq 0,$
- c) $\mathbf{E}(|Z_i|^k) \leq (k!)^\lambda M^k, \forall k \geq 0.$

Proof.

Following Bulinski and Shashkin (2007), let us define for a function $\Phi_m : \mathbb{R}^{m(d+1)} \rightarrow \mathbb{R}$, the partial Lipschitz constants of Φ_m , that is

$$\text{Lip}_i(\Phi_m) := \sup_{z_i \neq z'_i} \frac{|\Phi_m(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_{m(d+1)}) - \Phi_m(z_1, \dots, z_{i-1}, z'_i, z_{i+1}, \dots, z_{m(d+1)})|}{|z_i - z'_i|},$$

where $z_j \in \mathbb{R}$. For a function $\Phi : \mathbb{R}^m \rightarrow \mathbb{R}$, let $\text{Lip}(\Phi)$ denote the Lipschitz modulus of continuity of Φ , that is

$$\text{Lip}(\Phi) = \sup_{x \neq y} \frac{|\Phi(x) - \Phi(y)|}{\|x - y\|_1},$$

where $\|x\|_1 = |x_1| + \dots + |x_m|$.

Set $\Phi_u = \prod_{i=s_1}^{s_u} Z_i$ and $\Phi_v = \prod_{j=t_1}^{t_v} Z_j$. Then from one hand, by using Theorem 5.3 in Bulinski and Shashkin (2007), under the Lipschitz condition on K_d and G and from (4.10), we have

$$\text{cov}(Z_{s_1} \cdots Z_{s_u}, Z_{t_1} \cdots Z_{t_v}) \leq \sum_{i=s_1}^{s_u} \sum_{j=w_1}^{w_v} \text{Lip}_i(\Phi_u) \text{Lip}_j(\Phi_v) \theta_{i,j}.$$

As Y is bounded, and from the definition of $\text{Lip}_i(\cdot)$ given above, we get

$$\begin{aligned} \text{Lip}_i(\Phi_u) &\leq \frac{C_1}{h_n} \left(\frac{c}{G(a_F)} \right)^{u-1} \|K_d\|_\infty^{u-1} \\ \text{Lip}_j(\Phi_v) &\leq \frac{C_1}{h_n} \left(\frac{c}{G(a_F)} \right)^{v-1} \|K_d\|_\infty^{v-1}, \end{aligned}$$

where $C_1 = \text{Max} \left\{ \text{Lip}(K_d) \frac{c}{G(a_F)}, \|K_d\|_\infty \left(\frac{1}{G(a_F)} + \text{Lip}(G) \frac{c}{G^2(a_F)} \right) \right\}$. Therefore, the stationarity and assumption **(R)** give

$$\begin{aligned} \text{cov}(Z_{s_1} \cdots Z_{s_u}, Z_{t_1} \cdots Z_{t_v}) &\leq \frac{C_1^2}{h_n^2} \left(\frac{c}{G(a_F)} \right)^{u+v-2} \|K_d\|_\infty^{u+v-2} uv \rho(t_1 - s_u) \\ &\leq C_1^2 \left(\frac{c \|K_d\|_\infty}{G(a_F)} \right)^{u+v-2} \frac{1}{h_n^2} uv \rho(t_1 - s_u). \end{aligned} \quad (4.14)$$

On the other hand, under assumptions **(K1)** and **(D4)** we have

$$\mathbf{E} \left(\frac{\alpha^2 Y_i Y_j}{G(Y_i)G(Y_j)} K_d \left(\frac{x - X_i}{h_n} \right) K_d \left(\frac{x - X_j}{h_n} \right) \right) \leq c \left(\frac{1}{G(a_F)} \right)^2 h_n^{2d} \quad (4.15)$$

and

$$\mathbf{E} \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \leq \frac{c}{G(a_F)} h_n^d. \quad (4.16)$$

Then from (4.15) and (4.16) we deduce

$$\begin{aligned} \text{cov}(Z_{s_1} \cdots Z_{s_u}, Z_{t_1} \cdots Z_{t_v}) &= \mathbf{E}(Z_{s_1} \cdots Z_{s_u} Z_{t_1} \cdots Z_{t_v}) - \mathbf{E}(Z_{s_1} \cdots Z_{s_u}) \mathbf{E}(Z_{t_1} \cdots Z_{t_v}) \\ &\leq c \left(\frac{1}{G(a_F)} \right)^{u+v-2} h_n^{2d}. \end{aligned} \quad (4.17)$$

By combining (4.14) and (4.17), we get

$$\text{cov}(Z_{s_1} \cdots Z_{s_u}, Z_{t_1} \cdots Z_{t_v}) \leq \left(\frac{c \|K_d\|_\infty}{G(a_F)} \right)^{u+v} h_n^d uv (\rho(t_1 - s_u))^{\frac{d}{2d+2}}. \quad (4.18)$$

Choosing $K = \frac{c \|K_d\|_\infty}{G(a_F)} \sqrt{h_n^d}$, $M = \frac{c \|K_d\|_\infty}{G(a_F)}$ and $\lambda = 0$ we get the result in item a). To prove the item b), we follow the same steps as in the proof of Proposition 8 of Doukhan and Neumann (2007). For all $k \geq 0$, under assumption **(R)**, we have

$$\begin{aligned} \sum_{s=0}^{\infty} (s+1)^k (\rho(s))^{d/(2d+2)} &\leq \sum_{s=0}^{\infty} (s+1) \cdots (s+k) e^{-\gamma s d/(2d+2)} \\ &= \frac{d^k}{dt^k} \left(\frac{1}{1-t} \right) \Big|_{t=e^{-\gamma d/(2d+2)}} \\ &= (k)! \left(\frac{1}{1-e^{-\gamma d/(2d+1)}} \right)^{k+1}. \end{aligned}$$

Thus by choosing $\mu = 1$ and $L_1 = L_2 = \frac{1}{1-e^{-\gamma d/(2d+2)}}$ we get the result in item (b). For the proof of item (c), we have for all $k \geq 0$

$$\begin{aligned} \mathbf{E}(|Z_i|^k) &\leq \left(\frac{c \|K_d\|_\infty}{G(a_F)} \right)^k \\ &= (k!)^\lambda M^k, \end{aligned}$$

where λ and M have the same values chosen to get result (a). □

Lemma 4.3. *Under assumptions **(H)**, **(K1)**-**(K3)**, **(R)** and **(D1)**-**(D2)**, we have*

$$\sigma_n^2 := \text{Var} \left(\sum_{i=1}^n Z_i(x) \right) = O(nh_n^d).$$

Proof. We have

$$\begin{aligned}
\sigma_n^2 &= \text{Var} \left(\sum_{i=1}^n Z_i(x) \right) = (nh_n^d)^2 \text{Var} \left(\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right) = n^2 h_n^{2d} \text{Var} \left(\tilde{\psi}_n(x) \right) \\
&= n^2 h_n^{2d} \text{Var} \left(\frac{1}{nh_n^d} \sum_{i=1}^n \frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) \\
&= n \text{Var} \left(\frac{\alpha Y_1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right) + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov} \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right), \frac{\alpha Y_j}{G(Y_j)} K_d \left(\frac{x - X_j}{h_n} \right) \right) \\
&=: V + CV.
\end{aligned}$$

On the one hand, we have

$$\begin{aligned}
V &= n \left[\mathbf{E} \left(\frac{\alpha^2 Y_1^2}{G^2(Y_1)} K_d^2 \left(\frac{x - X_1}{h_n} \right) \right) - \mathbf{E}^2 \left(\frac{\alpha Y_1}{G(Y_1)} K_d \left(\frac{x - X_1}{h_n} \right) \right) \right] \\
&=: n(V_1 - V_2).
\end{aligned}$$

Using (5.3) and a change of variable, we get

$$\begin{aligned}
V_1 &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha^2 y^2}{G^2(y)} K_d^2 \left(\frac{x - u}{h_n} \right) F^*(du, dy) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha y^2}{G(y)} K_d^2 \left(\frac{x - u}{h_n} \right) F(du, dy) \\
&\leq h_n^d \int_{\mathbb{R}^d} K_d^2(z) \psi_1(x - zh_n) dz.
\end{aligned}$$

A Taylor expansion gives

$$\psi_1(x - zh_n) = \psi_1(x) - h_n \left(z_1 \frac{\partial \psi_1}{\partial z_1}(x^*) + \dots + z_d \frac{\partial \psi_1}{\partial z_d}(x^*) \right),$$

where x^* is between $x - h_n z$ and x . Then under assumptions **(H)**, **(K3)** and **(D2)** we have $V_1 = O(h_n^d)$. In the same way, we get

$$\begin{aligned}
V_2 &= \left[\int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha y}{G(y)} K_d \left(\frac{x - u}{h_n} \right) F^*(du, dy) \right]^2 \\
&= \left[\int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) F(du, dy) \right]^2 \\
&= \left[h_n^d \int_{\mathbb{R}^d} K_d(z) \psi(x - zh_n) dz \right]^2.
\end{aligned}$$

Therefore under assumptions **(H)**, **(K1)**-**(K2)** and **(D1)**, a Taylor expansion gives $V_2 = O(h_n^{2d})$.

Thus $V = O(nh_n^d)$. On the other hand, from (4.17) we can write

$$\text{Cov} \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right), \frac{\alpha Y_j}{G(Y_j)} K_d \left(\frac{x - X_j}{h_n} \right) \right) = O(h_n^{2d}). \quad (4.19)$$

To evaluate CV , we use a technique developed in Masry (1986). Let us define

$$\begin{aligned}\mathcal{B}_1 &= \{(i, j); 1 \leq |i - j| \leq \vartheta_n\} \\ \mathcal{B}_2 &= \{(i, j); \vartheta_n + 1 \leq |i - j| \leq n - 1\},\end{aligned}$$

where $\vartheta_n = o(n)$. Then

$$\begin{aligned}CV &= \sum_{i=1}^n \sum_{j \in \mathcal{B}_1} Cov \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right), \frac{\alpha Y_j}{G(Y_j)} K_d \left(\frac{x - X_j}{h_n} \right) \right) \\ &\quad + \sum_{i=1}^n \sum_{j \in \mathcal{B}_2} Cov \left(\frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right), \frac{\alpha Y_j}{G(Y_j)} K_d \left(\frac{x - X_j}{h_n} \right) \right) \\ &=: CV_1 + CV_2.\end{aligned}$$

From (4.19) we get

$$CV_1 = O(\vartheta_n n h_n^{2d}). \quad (4.20)$$

By assumption **(R)** and (4.18) we obtain

$$\begin{aligned}CV_2 &\leq c n h_n^d \sum_{j \in \mathcal{B}_2} e^{-\frac{\gamma_d}{2d+2}|i-j|} \leq c n h_n^d \int_{\vartheta_n}^n e^{-\frac{\gamma_d}{2d+2}u} du \\ &= O \left(n h_n^d e^{-\frac{\gamma_d \vartheta_n}{2d+2}} \right).\end{aligned} \quad (4.21)$$

Now choosing $\vartheta_n = O(h_n^{\nu-d})$ with $0 < \nu < d$, (4.20) and (4.21) become

$$\begin{aligned}CV_1 &= O(n h_n^d h_n^\nu) = o(n h_n^d) \\ CV_2 &= O \left(n h_n^d e^{-\frac{\gamma_d}{2d+2} C h_n^{\nu-d}} \right) = o(n h_n^d).\end{aligned}$$

Finally $CV = o(n h_n^d)$ and $\sigma_n^2 = O(n h_n^d)$. □

Proof of Theorem 4.1. Note that

$$\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) = \frac{1}{n h_n^d} \sum_{i=1}^n Z_i(x). \quad (4.22)$$

We use a classical technique which consists in covering the compact D by a finite number p_n of balls $B_k(x_k, a_n^d)$ centred at $x_k = (x_{k,1}, \dots, x_{k,d})$, for $k \in 1, \dots, p_n$ and where $a_n^d = n^{-\frac{1}{2\beta}} h^{1+\frac{d}{2\beta}}$.

Then for all $x \in D$, there exists $k \in 1, \dots, p_n$ such that $\|x - x_k\| \leq a_n^d$. As D is bounded, there exists a constant $M' > 0$ such that

$$p_n a_n^d \leq M' \Rightarrow p_n \leq \frac{M'}{a_n^d} \Rightarrow p_n = O((a_n^d)^{-1}).$$

Hence we consider the following decomposition :

$$\begin{aligned}
\sup_{x \in D} \left| \tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right| &= \sup_{x \in D} \left| \frac{1}{nh_n^d} \sum_{i=1}^n Z_i(x) - \frac{1}{nh_n^d} \sum_{i=1}^n Z_i(x_k) + \frac{1}{nh_n^d} \sum_{i=1}^n Z_i(x_k) \right| \\
&\leq \max_{1 \leq k \leq p_n} \sup_{x \in B_k} \frac{1}{nh_n^d} \sum_{i=1}^n |Z_i(x) - Z_i(x_k)| + \max_{1 \leq k \leq p_n} \frac{1}{nh_n^d} \left| \sum_{i=1}^n Z_i(x_k) \right| \\
&=: S_1 + S_2.
\end{aligned} \tag{4.23}$$

First, we have under assumptions **(K4)** and **(H)**,

$$\begin{aligned}
\frac{1}{nh_n^d} \sum_{i=1}^n |Z_i(x) - Z_i(x_k)| &\leq \frac{1}{nh_n^d} \sum_{i=1}^n \frac{\alpha |Y_i|}{G(Y_i)} \left| K_d \left(\frac{x - X_i}{h_n} \right) - K_d \left(\frac{x_k - X_i}{h_n} \right) \right| \\
&\quad + \frac{1}{h_n^d} \mathbf{E} \left(\frac{\alpha |Y_i|}{G(Y_i)} \left| K_d \left(\frac{x - X_i}{h_n} \right) - K_d \left(\frac{x_k - X_i}{h_n} \right) \right| \right) \\
&\leq \frac{2\alpha c}{nh_n^d G(a_F)} n \left\| \frac{x - X_1}{h_n} - \frac{x_k - X_1}{h_n} \right\|^\beta \\
&= \frac{c}{h_n^d} \cdot \frac{\|x - x_k\|^\beta}{h_n^\beta} \\
&\leq \frac{c(a_n^d)^\beta}{h_n^{d+\beta}} = O\left(\frac{1}{\sqrt{nh_n^d}}\right).
\end{aligned}$$

So, we get

$$S_1 = O\left(\frac{1}{\sqrt{nh_n^d}}\right). \tag{4.24}$$

We now turn to the term S_2 in (4.23). The use of Lemma 4.2 shows that $Z_i(x)$ defined in (4.12) satisfies the conditions of Lemma 4.1 with $\delta(s) = (\rho(s))^{\frac{d}{2d+2}}$ and we have

$$\mathbf{P} \left(\sum_{i=1}^n Z_i(x_k) \geq \varepsilon \right) \leq \exp \left(- \frac{\varepsilon^2/2}{A_n + B_n^{1/(\mu+\lambda+2)} \varepsilon^{(2\mu+2\lambda+3)/(\mu+\lambda+2)}} \right), \tag{4.25}$$

where A_n can be chosen such that $A_n \leq \sigma_n^2$ with $\sigma_n^2 := \text{Var} \left(\sum_{i=1}^n Z_i(x) \right)$ and $B_n = 2cL_2 \left(\frac{2^{4+\mu+\lambda} n c h_n^d L_1}{A_n} \vee 1 \right)$.

From Lemma 4.3 we have $A_n = O(nh_n^d)$. Furthermore, following the proof of Lemma 4.2, we have $\mu = 1$, $\lambda = 0$, $L_1 = L_2 = \frac{1}{1-e^{-\gamma d/(2d+2)}}$ and $B_n = O(1)$, then in (4.25) we get

$$\mathbf{P} \left(\sum_{i=1}^n Z_i(x_k) \geq \varepsilon \right) \leq \exp \left(- \frac{\varepsilon^2/2}{c n h_n^d + \varepsilon^{5/3}} \right). \tag{4.26}$$

Next, if we choose $\varepsilon = \varepsilon_0 \sqrt{\frac{\log n}{nh^d}}$ for all $\varepsilon_0 > 0$ then from (4.26) we have

$$\begin{aligned}
\mathbf{P} \left(\max_{1 \leq k \leq p_n} \left| \sum_{i=1}^n Z_i(x_k) \right| > nh_n^d \varepsilon_0 \sqrt{\frac{\log n}{nh^d}} \right) &\leq \sum_{k=1}^{p_n} \mathbf{P} \left(\left| \sum_{i=1}^n Z_i(x_k) \right| > nh_n^d \varepsilon_0 \sqrt{\frac{\log n}{nh^d}} \right) \\
&\leq 2p_n \exp \left(- \frac{\frac{\varepsilon_0^2}{2} n^2 h_n^{2d} \frac{\log n}{nh^d}}{cnh_n^d + \varepsilon_0^{\frac{5}{3}} (nh_n^d)^{\frac{5}{3}} \left(\frac{\log n}{nh_n^d} \right)^{\frac{5}{6}}} \right) \\
&\leq 2M'(a_n^d)^{-1} \exp \left(- \frac{\frac{\varepsilon_0^2}{2} \log n}{c + \varepsilon_0^{\frac{5}{3}} \left(\frac{\log^5 n}{nh_n^d} \right)^{\frac{1}{6}}} \right) \\
&\quad - \frac{\varepsilon_0^{2/2}}{c + \varepsilon_0^{\frac{5}{3}} \left(\frac{\log n^5}{nh_n^d} \right)^{\frac{1}{6}}} \\
&= 2M'(a_n^d)^{-1} n^{-\frac{1}{2\beta} h^{-1-\frac{d}{2\beta}} n^{-c\varepsilon_0^2}} \\
&= 2M'(nh)^{-1-\frac{d}{2\beta}} n^{-c\varepsilon_0^2+1+\frac{d+1}{2\beta}} \\
&= 2M' \frac{1}{\sqrt{(nh)^{\frac{2\beta+d}{\beta}}}} n^{-c\varepsilon_0^2+1+\frac{d+1}{2\beta}}. \tag{4.27}
\end{aligned}$$

By assumption **(H)** and for a suitable choice of ε_0 (i.e. $\varepsilon_0^2 > \frac{1}{c} \left(2 + \frac{d+1}{\beta} \right)$) the last term in (4.27) is the general term of a convergent series. Finally, applying Borel-Cantelli's lemma to (4.27) gives

$$S_2 = O \left(\sqrt{\frac{\log n}{nh_n^d}} \right). \tag{4.28}$$

Thus combining (4.24), (4.28) and (4.23) ends the proof. \square

Proof of Theorem 4.2. We follow step by step the proof of Theorem 4.1 with

$$Z_i(x) = \frac{\alpha}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{\alpha}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right),$$

we get the result. \square

Lemma 4.4. Under assumptions **(A)**, **(R)** and **(K1)** we have

$$\sup_{x \in D} \left| \hat{\psi}_n(x) - \tilde{\psi}_n(x) \right| = O \left[\left(\frac{\log \log n}{n} \right)^\theta \right] \quad \mathbf{P}\text{- a.s as, } n \rightarrow +\infty.$$

Proof.

$$\begin{aligned}
\left| \hat{\psi}_n(x) - \tilde{\psi}_n(x) \right| &= \left| \frac{1}{nh_n^d} \sum_{i=1}^n \left[\frac{\alpha_n}{G_n(Y_i)} - \frac{\alpha}{G(Y_i)} \right] Y_i K_d \left(\frac{x - X_i}{h_n} \right) \right| \\
&\leq \frac{1}{nh_n^d} \sum_{i=1}^n \left| \frac{\alpha_n - \alpha}{G_n(Y_i)} - \frac{\alpha}{G_n(Y_i)G(Y_i)} (G_n(Y_i) - G(Y_i)) \right| |Y_i| K_d \left(\frac{x - X_i}{h_n} \right) \\
&\leq \left\{ \frac{|\alpha_n - \alpha|}{G_n(a_F)} + \frac{\alpha}{G_n(a_F)G(a_F)} \sup_{y \geq a_F} |G_n(y) - G(y)| \right\} \frac{1}{nh_n^d} \sum_{i=1}^n |Y_i| K_d \left(\frac{x - X_i}{h_n} \right).
\end{aligned}$$

By using Markov's inequality, assumption **(K1)** and for $\varepsilon > 0$ we get

$$\begin{aligned}
\mathbf{P} \left(\frac{1}{nh_n^d} \sum_{i=1}^n |Y_i| K_d \left(\frac{x - X_i}{h_n} \right) \geq \varepsilon \right) &\leq \frac{\mathbf{E} \left(\frac{1}{nh_n^d} \sum_{i=1}^n |Y_i| K_d \left(\frac{x - X_i}{h_n} \right) \right)}{\varepsilon} \\
&\leq \frac{c}{\varepsilon h_n^d} \mathbf{E} \left(K_d \left(\frac{x - X_1}{h_n} \right) \right) \\
&= \frac{c}{\varepsilon h_n^d} \int_{\mathbb{R}^d} K_d \left(\frac{x - u}{h_n} \right) v^*(u) du \\
&= c\varepsilon^{-1} \int_{\mathbb{R}^d} K_d(z) v^*(x - h_n z) dz \\
&= O(1).
\end{aligned}$$

Now, by following Ould Saïd and Tatachak (2009), Guessoum et al. (2012) and using assumptions **(A)** and **(R)** we get

$$\sup_{y \geq a_F} |G_n(y) - G(y)| = O \left[\left(\frac{\log \log n}{n} \right)^\theta \right],$$

and

$$|\alpha_n - \alpha| = O \left[\left(\frac{\log \log n}{n} \right)^\theta \right].$$

Finally, since $G_n(a_F) \xrightarrow{\mathbf{P}-a.s.} G(a_F)$ and $\frac{1}{nh_n^d} \sum_{i=1}^n |Y_i| K_d \left(\frac{x - X_i}{h_n} \right) = O(1)$, we deduce the result. \square

Lemma 4.5. *Under assumptions **(K1)**-**(K2)**, **(D1)** and **(H)** we have*

$$\sup_{x \in D} |\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x)| = O(h_n^2) \text{ a.s. as } n \rightarrow \infty.$$

Proof.

$$\begin{aligned}
\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) &= \mathbf{E} \left(\frac{1}{nh_n^d} \sum_{i=1}^n \frac{\alpha Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right) - \psi(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{\alpha y}{G(y)} K_d \left(\frac{x - u}{h_n} \right) F^*(du, dy) - \psi(x) \\
&= \frac{1}{h_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}} y K_d \left(\frac{x - u}{h_n} \right) F(du, dy) - \psi(x) \\
&= \int_{\mathbb{R}^d} K_d(z) (\psi(x - zh_n) - \psi(x)) dz.
\end{aligned}$$

Under assumptions **(K1)**-**(K2)**, and **D1**, a Taylor expansion finishes the proof. \square

Lemma 4.6. *Under assumptions **(K1)**-**(K4)**, **(A)**, **(R)**, **(D3)**-**(D4)** and **(H)***

$$\sup_{x \in D} |\hat{v}_n(x) - v(x)| = O \left\{ \sqrt{\frac{\log n}{nh_n^d}} \vee \left(\frac{\log \log n}{n} \right)^\theta \vee h_n^2 \right\} \mathbf{P}\text{-a.s. as } n \rightarrow \infty$$

where $0 < \theta < \frac{\gamma}{2\gamma+6+3/2\kappa}$ for any real $\kappa > 0$.

Proof. We have

$$\sup_{x \in D} |\hat{v}_n(x) - v(x)| \leq \sup_{x \in D} |\hat{v}_n(x) - \tilde{v}_n(x)| + \sup_{x \in D} |\tilde{v}_n(x) - \mathbf{E}(\tilde{v}_n(x))| + \sup_{x \in D} |\mathbf{E}(\tilde{v}_n(x)) - v(x)|.$$

Using the same steps and arguments as in the proof of Lemma 4.4 we get under assumptions **(A)** and **(K1)**

$$\sup_{x \in D} |\hat{v}_n(x) - \tilde{v}_n(x)| = O \left[\left(\frac{\log \log n}{n} \right)^\theta \right]. \quad (4.29)$$

Furthermore, under assumptions **(K1)**-**(K2)**, **(D3)** and **(H)** and using a Taylor expansion, we get

$$\sup_{x \in D} |\mathbf{E}(\tilde{v}_n(x)) - v(x)| = O(h_n^2). \quad (4.30)$$

By combining (4.29), (4.30) and Theorem 4.2 we end the proof. \square

Proof of Theorem 4.3. As already mentioned in section 3, let $\delta > 0$ such that $\inf_{x \in D} |v(x)| > \delta$, hence we have from (4.11)

$$\begin{aligned} \sup_{x \in D} |\hat{m}_n(x) - m(x)| &\leq \frac{1}{\inf_{x \in D} (\hat{v}_n(x))} \left\{ \sup_{x \in D} |\hat{\psi}_n(x) - \tilde{\psi}_n(x)| + \sup_{x \in D} |\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x))| \right. \\ &\quad \left. + \sup_{x \in D} |\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x)| + \frac{\sup_{x \in D} |\psi(x)|}{\inf_{x \in D} v(x)} \sup_{x \in D} |\hat{v}_n(x) - v(x)| \right\} \\ &\leq \frac{1}{\delta - \sup_{x \in D} |\hat{v}_n(x) - v(x)|} \left\{ \sup_{x \in D} |\hat{\psi}_n(x) - \tilde{\psi}_n(x)| + \sup_{x \in D} |\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x))| \right. \\ &\quad \left. + \sup_{x \in D} |\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x)| + \delta^{-1} \sup_{x \in D} |\psi(x)| \sup_{x \in D} |\hat{v}_n(x) - v(x)| \right\}. \end{aligned}$$

Then by using Theorem 4.1, Lemma 4.4, Lemma 4.5 and Lemma 4.6 we get the result. \square

Chapitre 5

Normalité asymptotique de l'estimateur de la fonction de régression pour des données tronquées à gauche et associées

Ce chapitre est consacré à l'étude de la normalité asymptotique de l'estimateur à noyau de la fonction de régression pour des données tronquées à gauche et associées. Nous donnons la version française des résultats tels qu'ils ont été soumis.

5.1 Introduction

Il est connu que les estimateurs non paramétriques de la fonction de régression fournissent un moyen approprié et efficace pour déterminer la relation probablement complexe entre une variable donnée d'intérêt et certaines variables explicatives. En raison de cet intérêt pratique évident, beaucoup d'auteurs ont étudié les propriétés (asymptotiques) de tels estimateurs pendant les dernières décennies. Une des hypothèses de base qui a été généralement considérée est l'indépendance des données observées. Néanmoins dans beaucoup de situations, l'hypothèse d'indépendance n'est pas vérifiée c'est pourquoi l'étude des données dépendantes a suscité beaucoup d'attention pendant ces dernières années. Deux types de données dépendantes sont ainsi largement utilisés dans la littérature sont les données mélangeantes et associées. Dans ce travail, nous sommes intéressés au concept de l'association qui a été introduit et défini par Esary et al. (1967). Rappelons que une suite finie de variables aléatoires (v.a.'s) $Y = (Y_1, Y_2, \dots, Y_N)$ est dite associée si pour toute paire de fonctions f et g définies sur \mathbb{R}^N dans \mathbb{R} non décroissantes coordonnée par coordonnée,

$$Cov(f(Y), g(Y)) \geq 0,$$

lorsque cette covariance existe. Une suite infinie de variables aléatoires est associée si toute sous suite finie est associée. Pour plus de détails sur le concept d'association, nous renvoyons le lecteur à

Bulinski et Shashkin (2007).

Soit Y une variable aléatoire (v.a.) réelle d'intérêt de fonction de répartition (f.d.r.) continue F , et X un vecteur aléatoire de \mathbb{R}^d de covariables de f.d.r. jointe V et de densité jointe v . La fonction de régression au point $x \in \mathbb{R}^d$ est donnée par l'espérance conditionnelle de Y sachant $X = x$, qui est

$$m(x) := \mathbb{E}[Y|X = x].$$

Notons que $m(x)$ peut être écrite sous la forme $m(x) = \frac{\psi(x)}{v(x)}$ où $\psi(x) = \int_{\mathbb{R}} y F(x, dy)$ et $F(., .)$ dénote le f.d.r. jointe de (X, Y) .

Plusieurs méthodes non paramétriques, qui consistent à estimer $m(x)$, existent dans la littérature, ici nous utilisons la méthode à noyau. Sous l'hypothèse que les variables aléatoires observées sont indépendantes, l'estimateur non paramétrique à noyau de $m(x)$ a été étudié par plusieurs auteurs, Nous citons quelques-uns d'entre eux (Devroye et al. (1996), Györfi et al. (1998), Walk(2002), ...). Dans les études de durée de vie, il arrive souvent que la v.a. d'intérêt Y ne puisse pas être complètement observée. La censure et la troncature sont les formes les plus courantes des données incomplètes. La v.a. Y peut être considérée comme la durée de survie des individus dans études biomédicales. Dans cet article, Nous sommes intéressés aux données tronquées à gauche, où l'observation de (X, Y) est interférée par une autre v.a. indépendante T tel que les trois quantités aléatoires X, Y et T sont observables seulement si $Y \geq T$. Telles données apparaissent en astronomie, l'économie, l'épidémiologie, la biométrie, ... (voir, Woodroffe (1985), Feigelson et Babu (1992), Wang et al. (1986), Tsai et al. (1987) et He et Yang (1994)).

Récemment, Ould Saïd et Lemdani (2006) ont construit un nouvel estimateur non paramétrique à noyau de la fonction de régression $m(.)$ pour le modèle aléatoire de troncature à gauche (RLT) et ils ont étudié ses propriétés asymptotiques pour des v.a.'s indépendantes et identiquement distribuées (i.i.d.). Plus tard, Liang et al. (2009) ont obtenu la consistance forte et uniforme de l'estimateur construit et Liang (2011) établit sa normalité asymptotique. Notre objectif dans le présent travail est d'établir la normalité asymptotique de l'estimateur de la fonction de régression introduit par Ould Saïd et Lemdani (2006) pour des données tronquées à gauche et associées. Il complète notre première étude (Guessoum et Hamrani (2016)) sur la convergence presque sûre uniforme du même estimateur. La suite de ce travail est organisée comme suit : la section 2 est consacrée à la présentation du modèle RLT et de ses estimateurs. Les hypothèses requises et le résultat principal sont donnés dans la section 3. Dans la section 4, quelques simulations sont données. Finalement, les preuves sont données à la section 5.

5.2 Présentation du modèle et les estimateurs

Soit $\{(Y_i, T_i); 1 \leq i \leq N\}$, une suite de N copies de couple (Y, T) définie sur $(\Omega, \mathcal{F}, \mathbb{P})$, où la taille N est fixée mais inconnue. Dans tout le reste de ce travail, $\{Y_i; 1 \leq i \leq N\}$ dénote une suite strictement stationnaire associée de Y supposée qui est une v.a. bornée et non négative de f.d.r. continue F et $\{T_i; 1 \leq i \leq N\}$ est une suite strictement stationnaire i.i.d. de la v.a. de troncature T qui a une f.d.r. G lipschitzienne et une densité de probabilité g bornée.

Dans le modèle RLT, la v.a. d'intérêt Y et la v.a. de troncature T sont observables seulement si $Y \geq T$, alors que rien n'est observé si $Y < T$. Sans confusion possible, Nous dénotons toujours $\{(Y_i, T_i); i = 1, \dots, n\}$, ($n \leq N$) la suite réellement observée provenant de la suite d'origine de taille N avec n étant aléatoire mais connu. Notons que puisque la suite de v.a.'s d'intérêt d'origine est associée, celle observée reste associée (par la propriété (P1) dans Esary et al. (1967) et la suite de v.a.'s de troncature observée est aussi i.i.d. (par la proposition 2.1 dans Lemdani et Ould Saïd (2007)).

Comme conséquence de troncature, la taille de la suite réellement observée, n , est une v.a. binomiale $\text{Bin}(N, \alpha)$, avec α est la probabilité de (non) troncature définie par $\alpha := \mathbb{P}(Y \geq T)$. Il est clair que si $\alpha = 0$, aucune donnée peut être observée ainsi, nous supposons dans tous ce travail que $\alpha > 0$.

Comme N est inconnue et n est connu (bien qu'il soit aléatoire), nos résultats ne seront pas établis par rapport à la probabilité \mathbb{P} (relative au N -échantillon) mais par rapport à une nouvelle probabilité \mathbf{P} (relative au n -échantillon) définie par $\mathbf{P}(\cdot) = \mathbb{P}(\cdot | Y \geq T)$. De la même manière, Nous noterons respectivement \mathbb{E} et \mathbf{E} les espérances relatives aux probabilités \mathbb{P} and \mathbf{P} .

Suivant Stute (1993), la f.d.r. jointe d'un couple (Y, T) observé, sous le modèle RLT, devient

$$\begin{aligned} H^*(y, t) &:= \mathbf{P}\{Y \leq y, T \leq t\} = \mathbb{P}\{Y \leq y, T \leq t | Y \geq T\} \\ &= \frac{1}{\alpha} \int_{-\infty}^y G(t \wedge u) dF(u), \end{aligned}$$

où $t \wedge u := \min(t, u)$. Les f.d.r.'s marginales respectives de Y et T et leurs estimateurs empiriques sont définis par

$$\begin{aligned} F^*(y) &= \frac{1}{\alpha} \int_{-\infty}^y G(u) dF(u), \quad G^*(t) = \frac{1}{\alpha} \int_{-\infty}^{\infty} G(t \wedge u) dF(u), \\ F_n^*(y) &:= \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq y\}} \quad \text{et} \quad G_n^*(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{T_i \leq t\}}, \end{aligned}$$

où \mathbb{I}_A dénote la fonction indicatrice de l'ensemble A .

Dans la suite, les f.d.r.'s inconnues F et G sont estimées respectivement par les estimateurs de

Lynden-Bell (1971) F_n et G_n donnés par

$$\begin{aligned} F_n(y) &:= 1 - \prod_{i: Y_i \leq y} \left[\frac{nC_n(Y_i) - 1}{nC_n(Y_i)} \right], \\ G_n(t) &:= \prod_{i: T_i > t} \left[\frac{nC_n(T_i) - 1}{nC_n(T_i)} \right], \end{aligned} \quad (5.1)$$

où $C_n(y) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{T_i \leq y \leq Y_i\}}$ est l'estimateur empirique de $C(y) := \mathbb{P}\{T \leq y \leq Y \mid Y \geq T\}$.

Et pour la probabilité de troncature inconnue α , nous utilisons l'estimateur proposé par He et Yang (1998) donné par

$$\alpha_n := \frac{G_n(y)(1 - F_n(y))}{C_n(y)}, \quad (5.2)$$

pour tout y tel que $C_n(y) \neq 0$, qui est indépendant de y . Woodroffe (1985) pointed out that F et G peuvent être complètement estimées seulement si

$$a_G \leq a_F, \quad b_G \leq b_F \quad \text{et} \quad \int_{a_F}^{\infty} \frac{dF}{G} < \infty.$$

Soit $\{(X_i, Y_i, T_i); i = 1, \dots, n\}$ les n triplets observés parmi les N triplets, tel que $Y_i \geq T_i$ et supposons que T indépendant de (X, Y) où la $\{(X_i, Y_i); i = 1, \dots, n\}$ est supposée strictement stationnaire et associé. Sous le modèle RLT, la f.d.r. jointe d'un triple (X, Y, T) observé est

$$\begin{aligned} L^*(x, y, t) &:= \mathbf{P}\{X \leq x, Y \leq y, T \leq t\} = \mathbb{P}\{X \leq x, Y \leq y, T \leq t \mid Y \geq T\} \\ &= \frac{1}{\alpha} \int_{u \leq x} \int_{a_G \leq w \leq y} G(w \wedge t) F(du, dw). \end{aligned}$$

Prenant $t = +\infty$, un couple observé (X, Y) a donc la f.r. jointe suivante

$$F_{X,Y}^*(x, y) := L^*(x, y, \infty) = \frac{1}{\alpha} \int_{u \leq x} \int_{a_G \leq w \leq y} G(w) F_{X,Y}(du, dw),$$

qui donne

$$F_{X,Y}(dx, dy) = \frac{\alpha}{G(y)} F_{X,Y}^*(dx, dy), \quad \text{pour } y \geq a_G. \quad (5.3)$$

Suivant Ould Saïd et Lemdani (2006), un estimateur à noyau de la fonction de régression $m(x)$ quand les données sont tronquées à gauche est donné par

$$\tilde{m}_n(x) := \frac{\tilde{\psi}_n(x)}{\tilde{v}_n(x)}, \quad (5.4)$$

où

$$\tilde{\psi}_n(x) := \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \quad (5.5)$$

est un estimateur de $\psi(x)$ et $v_n(x)$ est l'estimateur de la densité de X supposé qu'il est strictement positif pour tout $x \in \mathbb{R}^d$ et qui est défini par

$$\tilde{v}_n(x) := \frac{\alpha}{nh_n^d} \sum_{i=1}^n \frac{1}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right). \quad (5.6)$$

$K_d : \mathbb{R}^d \rightarrow \mathbb{R}$ est une fonction noyau et h_n une suite de nombres réels positifs appelée fenêtre tendant vers zéro quand $n \rightarrow \infty$.

Notons que dans (5.5), (5.6) et dans la suite, la somme est sur les i tels que $G_n(Y_i) \neq 0$.

En pratique α et G sont inconnus, d'où il n'est pas possible d'utiliser l'estimateur (5.4). Alors, en les remplaçant par leurs estimateurs G_n et α_n définis dans (5.1) et (5.2), respectivement, nous obtenons l'estimateur faisable de $m(x)$ suivant

$$\hat{m}_n(x) := \frac{\hat{\psi}_n(x)}{\hat{v}_n(x)},$$

où

$$\begin{aligned}\hat{\psi}_n(x) &:= \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{Y_i}{G_n(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right), \\ \hat{v}_n(x) &:= \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{1}{G_n(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right).\end{aligned}\tag{5.7}$$

Dans la suite de ce chapitre, on considérera seulement les $x \in \mathbb{R}^d$ tels que $v(x) > 0$ et la lettre C désignera une constante générique dont sa valeur peut varier d'une expression à une autre.

5.3 Hypothèses et résultats principaux

Définissons

$$\theta_{i,j} := \sum_{k=1}^d \sum_{l=1}^d \text{cov}(X_{i,k}, X_{j,l}) + 2 \sum_{k=1}^d \text{cov}(X_{i,k}, Y_j) + \text{cov}(Y_i, Y_j),$$

où $X_{i,k}$ est la $k^{\text{ème}}$ -composante de X_i . Pour formuler le résultat principal, nous devons imposer les hypothèses suivantes.

(H) La fenêtre h_n satisfait :

- (i) $nh_n^{d+4} \rightarrow 0$ quand $n \rightarrow +\infty$;
- (ii) $nh_n^{d/\tau} (\log \log n)^{1/\tau-1} \rightarrow 0$ quand $n \rightarrow +\infty$, pour certain $0 < \tau < 1$;

(A) $\int \frac{dF(z)}{G^2(z)} < +\infty$;

(K1) Le noyau K_d est une densité de probabilité bornée à support compact et admet des dérivées partielles d'ordre 1 bornées;

(K2) $\int_{\mathbb{R}^d} z_i K_d(z) dz = 0$ pour tout $i = 1, \dots, d$ et $\int_{\mathbb{R}^d} |z_1^{i_1} \cdots z_d^{i_d}| K_d(z) dz < +\infty$ pour $i_1 + \dots + i_d = 2$;

(R) Le terme de covariance définit par $\rho(s) := \sup_{|i-j| \geq s} \theta_{i,j}$ pour $s > 0$, vérifie $\rho(s) \leq \gamma_0 e^{-\gamma s}$ pour certaines constantes positives γ_0 et γ ;

(D1) La fonction $\psi(\cdot)$ est bornée et deux fois différentiable avec $\sup_{x \in \mathbb{R}^d} \left| \frac{\partial^k \psi}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ pour $i, j = 1, \dots, d$ et $k = 1, 2$;

- (D2) La densité $v(\cdot)$ est bornée et deux fois différentiable avec $\sup_{x \in \mathbb{R}^d} \left| \frac{\partial^k v}{\partial x_i \partial x_j^{k-1}}(x) \right| < +\infty$ pour $i = 1, \dots, d$ et $k = 1, 2$;
- (D3) La densité jointe $v_{i,j}^*$ d'un couple (X_i, X_j) observé existe et $\sup_{(r,s) \in \mathbb{R}^{2d}} |v_{i,j}^*(r, s)| \leq C < +\infty$ telle que C est une constante qui ne dépend pas de (i, j) ;
- (B) Soient $0 < p = p_n < n$, $0 < q = q_n < n$ et $k = k_n := \left\lfloor \frac{n}{p+q} \right\rfloor$, ($[x]$ désigne la partie entière de x) des suites de nombres entiers positives tendant vers ∞ , quand $n \rightarrow \infty$, telles que $k(p+q) \leq n$ et $\frac{k(p+q)}{n} \xrightarrow{n \rightarrow +\infty} 1$ satisfaisantes :
- (i) $\frac{pk}{n} \xrightarrow{n \rightarrow +\infty} 1$,
 - (ii) $ph_n^d \xrightarrow{n \rightarrow +\infty} 0$ et $\frac{p^2}{nh_n^d} \xrightarrow{n \rightarrow +\infty} 0$,
 - (iii) $\frac{e^{-\gamma q}}{h_n^{d+2}} \xrightarrow{n \rightarrow +\infty} 0$, où γ est donné dans (R).

Remarque 5.1. (Commentaires sur les hypothèses)

Les différentes hypothèses (H)(i), (K1)-(K2) et (D1)-(D2) sont classiques en estimation fonctionnelle. L'hypothèse (R) quantifie la décroissance exponentielle du terme de covariance. Les hypothèses (A) et (R) impliquent les hypothèses de Guessoum et al. (2012) dont on aura besoin d'utiliser leurs résultats. L'hypothèse (D3) est indispensable pour avoir une bonne majoration des covariances.

La condition (H)(ii) est utile dans nos calculs et elle est vérifiée pour un choix approprié de h_n . Par exemple, si nous choisissons $h_n = C.n^{-\beta}$ pour certain $\beta > 0$ alors elle devient $n^{1-\beta d/\tau} (\log \log n)^{1/\tau-1} \xrightarrow{n \rightarrow +\infty} 0$, ainsi elle est vérifiée, pourvu que $1 - \beta d/\tau < 0$, ceci est équivalent à $\beta > \tau/d$.

L'hypothèse (B) est utilisée lors de l'application de la procédure des grands blocs et petits blocs pour obtenir la normalité asymptotique qui est propre à la dépendance. Les conditions données dans cette hypothèse sont vérifiées facilement, si nous choisissons les suites p_n et q_n comme suit

$$p_n \approx h_n^{-\lambda_1} \text{ et } q_n \approx h_n^{-\lambda_2} \text{ avec } 0 < \lambda_2 < \lambda_1 < 1.$$

Remarque 5.2. Nous donnons ici quelques implications des conditions données dans l'hypothèse (B) dont on aura besoin dans la preuve de nos résultats.

- (a) Comme $\frac{qk}{n} = \frac{(p+q)k}{n} - \frac{pk}{n}$, l'hypothèse B(i) implique $\frac{qk}{n} \xrightarrow{n \rightarrow +\infty} 0$. On a aussi $\frac{q}{p} = \frac{\frac{qk}{n}}{\frac{pk}{n}} \xrightarrow{n \rightarrow +\infty} 0$ de sorte que $q < p$, éventuellement.
- (b) Le premier terme de l'hypothèse B(ii) et (a) implique que $qh_n^d \xrightarrow{n \rightarrow +\infty} 0$, qui implique aussi $h_n \xrightarrow{n \rightarrow +\infty} 0$.
- (c) Le deuxième terme de B(ii) implique $nh_n^d \xrightarrow{n \rightarrow +\infty} \infty$.

Notre résultat principal concernant la normalité asymptotique de $\hat{m}_n(x)$ est donné dans le théorème suivant.

Théorème 5.1. *Sous les hypothèses (H), (A), (K1)-(K2), (R), (D1)-(D3), (B) et pour tout x tel que $v(x) > 0$, nous avons*

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)),$$

où

$$\sigma^2(x) := \frac{\alpha [\psi_0(x)v^2(x) - 2\psi_1(x)\psi(x)v(x) + \psi_2(x)\psi^2(x)]}{v^4(x)} \kappa, \quad (5.8)$$

$$\kappa = \int_{\mathbb{R}^d} K_d^2(z) dz < +\infty \text{ et } \psi_j(x) = \int_{\mathbb{R}} \frac{y^{2-j}}{G(y)} F(x, dy), j = 0, 1, 2.$$

Remarque 5.3. *Un estimateur de type plug-in $\hat{\sigma}_n^2(x)$ pour la variance asymptotique $\sigma^2(x)$ peut être obtenu en utilisant les estimateurs $\alpha_n, \hat{v}_n(\cdot)$ donnés par (5.2), (5.7) et les estimateurs*

$$\hat{\psi}_{j,n}(x) := \frac{\alpha_n}{nh_n^d} \sum_{i=1}^n \frac{Y_i^{2-j}}{G_n^2(Y_i)} K_d\left(\frac{x - X_i}{h_n}\right)$$

de $\psi_j(\cdot), j = 0, 1, 2$.

Comme une application de la normalité asymptotique, nous obtenons l'intervalle de confiance qui est donné dans le corollaire suivant.

Corollaire 5.1. *Sous les hypothèses de Théorème 5.1, nous obtenons pour tout $\xi \in (0, 1)$, l'intervalle de confiance suivant de niveau asymptotique $1 - \xi$ pour $m(x)$*

$$\left[\hat{m}_n(x) - \frac{u_{1-\xi/2} \hat{\sigma}_n^2(x)}{\sqrt{nh_n^d}}, \hat{m}_n(x) + \frac{u_{1-\xi/2} \hat{\sigma}_n^2(x)}{\sqrt{nh_n^d}} \right],$$

où $u_{1-\xi/2}$ dénote le quantile d'ordre $1 - \xi/2$ de la loi normale centrée réduite.

5.4 Étude de simulations

Le but de cette section est d'examiner la normalité asymptotique de notre estimateur pour des échantillons de tailles finies dans le cas $d = 1$. Nous avons à simuler N observations $\{(X_i, Y_i, T_i); i = 1, \dots, N\}$ du triplet de v.a.'s (X, Y, T) avec T indépendante de (X, Y) . Nous considérons la suite associée $\{X_t, Y_t\}$ générée par

$$\begin{cases} X_t = (W_{t-1} + W_{t-2})/2, \\ Y_t = 2X_t + 1 + \varepsilon_t, \end{cases}$$

où

- $\{W_t; t = -1, 0, \dots, N-1\}$ est une suite de $N+1$ v.a.'s i.i.d. suivant une loi exponentielle de paramètre 1.
- $\{\varepsilon_t; t = 1, \dots, N\}$ est une suite de N v.a.'s i.i.d. de loi normale centrée de variance 0.2.

Nous générons également N v.a.'s $\{T_t\}$ i.i.d. suivant une loi exponentielle de paramètre λ (λ est adapté de sorte d'obtenir des valeurs différentes de α). Ensuite, on retient les observations $\{(X_i, Y_i, T_i); i = 1, \dots, n\}$ vérifiant $Y_i \geq T_i$. À partir de ces observations, nous calculons les valeurs de l'estimateur \hat{m}_n en choisissant le noyau gaussien et la fenêtre $h_n = O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{3}}\right)$.

Maintenant, nous examinons d'abord la normalité asymptotique de l'estimateur en comparant la forme de la densité estimée (des écarts normalisés) à celle de la densité de la loi normale centrée réduite. Pour cela, nous estimons la fonction de régression $m(x) = 2x + 1$ comme auparavant et nous calculons l'écart normalisé entre l'estimateur et la fonction de régression théorique pour $x = 0.5$, c'est-à-dire

$$\begin{aligned} \dot{m}_n = \dot{m}_n(0.5) &:= \frac{\sqrt{nh_n}}{\hat{\sigma}_n^2(0.5)}(\hat{m}_n(0.5) - m(0.5)) \\ &= \frac{\sqrt{nh_n}}{\hat{\sigma}_n^2(0.5)}(\hat{m}_n(0) - 2). \end{aligned}$$

Nous générons, en utilisant cette procédure, $B = 1000$ suites indépendantes de taille n . Ceci donne une suite de v.a.'s i.i.d. $\dot{m}_1, \dots, \dot{m}_B$. Nous estimons ensuite, en utilisant la méthode à noyau, sa fonction de densité en choisissant un noyau gaussien et la fenêtre $h_B = C.B^{-1/5}$ (voir Silverman (1986)) où la constante C est choisie convenablement. Puis, Nous représentons dans le même graphe, la densité estimée \dot{m}_n avec la densité de la loi normale centrée réduite pour différentes valeurs de n et pour $\alpha \approx 80\%$ (obtenu pour $\lambda = 1.5$). Nous représentons aussi les histogrammes et les diagrammes quantile-quantile correspondants pour chaque cas.

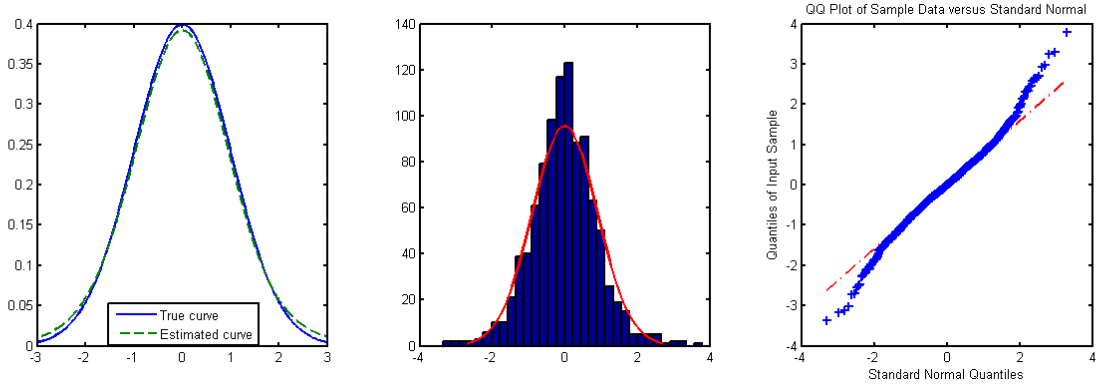
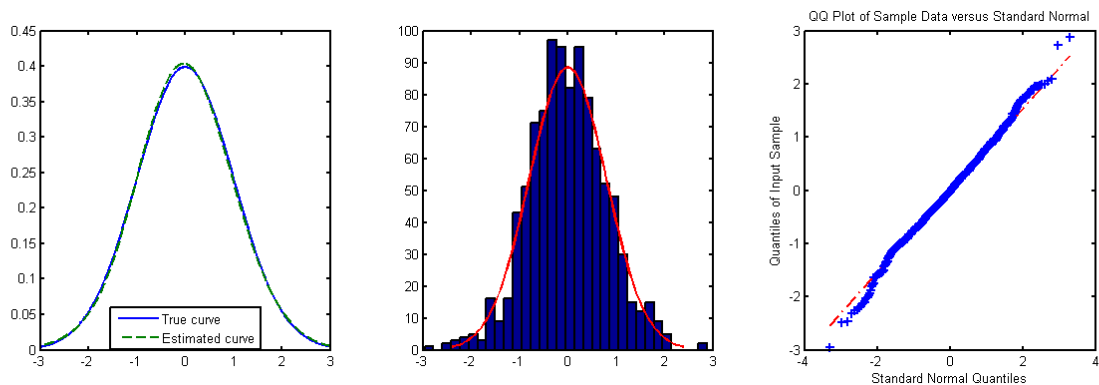


FIGURE 5.1 – Graphes de la normalité pour $n = 100$ et $\alpha \approx 80\%$.

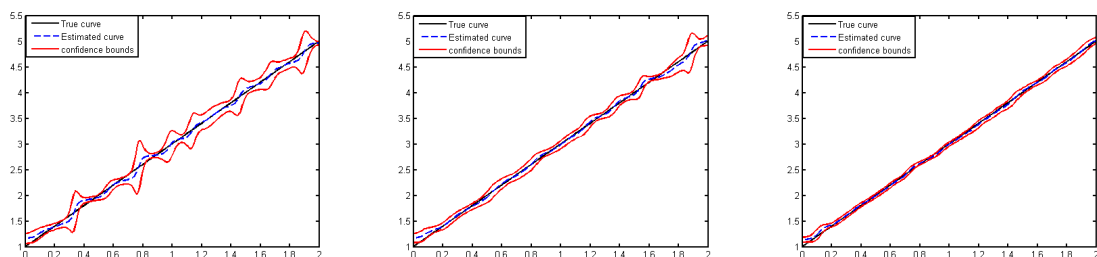
Dans les Figures 5.1 et 5.2, nous remarquons que la qualité de l'ajustement est meilleure quand n est grand. Mais tous les graphes montrent une bonne qualité de l'ajustement et confirme le résultat de la normalité asymptotique théorique.

FIGURE 5.2 – Graphes de la normalité pour $n = 300$ et $\alpha \approx 80\%$.

Ensuite, nous simulons les intervalles de confiance approximatifs à 95% pour $m(x)$ obtenus dans le Corollaire 5.1, donnés par

$$\hat{m}_n(x) \pm 1.96 \times \frac{\hat{\sigma}_n^2(x)}{\sqrt{nh_n}},$$

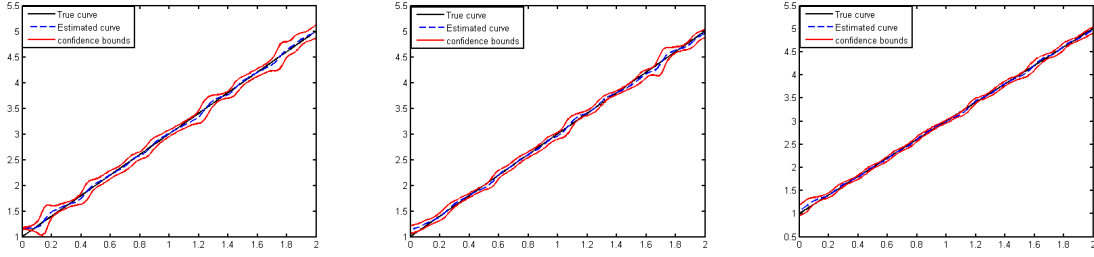
en utilisant les mêmes données et les mêmes arguments que ceux utilisés pour calculer l'estimateur $\hat{m}_n(x)$. Nous générons des suites de taille $n = 50, 100$ et 300 , respectivement, et nous prenons $\alpha \approx 70\%$ et 90% (obtenus pour $\lambda = 2.5$). Nous représentons aux Figures 5.3 et 5.4, la fonction de régression théorique $m(x)$ avec son estimateur $\hat{m}_n(x)$ et les bornes inférieure et supérieure de l'intervalle de confiance approximatif, $x \in [0; 2]$.

FIGURE 5.3 – $\alpha \approx 70\%$, $n = 50$, $n = 100$ et $n = 300$ respectivement.

Dans les Figures 5.3 et 5.4, nous remarquons que

- La longueur des intervalles de confiance estimés devient courtes quand la taille de la suite observée n augmente.
- Pour la même taille de la suite observée, la qualité de l'estimation est meilleure quant le pourcentage de données observées α devient plus grand.

Finalement, nous testons l'efficacité de notre normalité asymptotique en calculant les probabilités de couverture de l'intervalle de confiance de niveau asymptotique 95%. Elles sont obtenues par le pourcentage des vraies valeurs appartenant à l'intervalle de confiance estimé. Nous choisissons

FIGURE 5.4 – $\alpha \approx 90\%$, $n = 50$, $n = 100$ et $n = 300$ respectivement.

$n = 50$, 100 et 300 , $\alpha \approx 70\%$, 90% (obtenus pour $\lambda = 1.5$ et $\lambda = 2.5$ respectivement). Les résultats sont donnés dans le Tableau 5.1 pour 1000 répliquations de suites observées de taille n .

TABLE 5.1 – Les probabilités de couverture de l'intervalle de confiance à 95%.

	n=50	n=100	n=300
$\alpha \approx 70\%$	0.9249	0.9311	0.9551
$\alpha \approx 90\%$	0.9307	0.9389	0.9556

Les valeurs des probabilités de couverture montrent une couverture raisonnable pour l'intervalle de confiance et elle est améliorée quand la valeur de n augmente.

5.5 Preuve du résultat principal

La preuve est principalement basée sur la décomposition suivante

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] = \sqrt{nh_n^d} [(\hat{m}_n(x) - \tilde{m}_n(x)) + (\tilde{m}_n(x) - m(x))]. \quad (5.9)$$

Notons que

$$\begin{aligned}
\tilde{m}_n(x) - m(x) &= \frac{\tilde{\psi}_n(x)}{\tilde{v}_n(x)} - \frac{\psi(x)}{v(x)} \\
&= \frac{1}{\tilde{v}_n(x)} \left[\left(\tilde{\psi}_n(x) - \psi(x) \right) - \frac{\psi(x)}{v(x)} (\tilde{v}_n(x) - v(x)) \right] \\
&= \frac{v(x)}{\tilde{v}_n(x)} \left\{ \left[\left(\tilde{\psi}_n(x) - \mathbf{E}(\tilde{\psi}_n(x)) \right) - \frac{\psi(x)}{v(x)} (\tilde{v}_n(x) - \mathbf{E}(\tilde{v}_n(x))) \right] \times \frac{1}{v(x)} \right\} \\
&\quad + \frac{v(x)}{\tilde{v}_n(x)} \left\{ \left[\left(\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right) - \frac{\psi(x)}{v(x)} (\mathbf{E}(\tilde{v}_n(x)) - v(x)) \right] \times \frac{1}{v(x)} \right\} \\
&=: \frac{v(x)}{\tilde{v}_n(x)} \mathcal{J}_n(x) + \frac{v(x)}{\tilde{v}_n(x)} \mathcal{K}_n(x).
\end{aligned}$$

Soit $\mathcal{I}_n(x) := \hat{m}_n(x) - \tilde{m}_n(x)$, alors (5.9) devient

$$\sqrt{nh_n^d} [\hat{m}_n(x) - m(x)] = \sqrt{nh_n^d} \mathcal{I}_n(x) + \frac{v(x)}{\tilde{v}_n(x)} \left[\sqrt{nh_n^d} \mathcal{J}_n(x) \right] + \frac{v(x)}{\tilde{v}_n(x)} \left[\sqrt{nh_n^d} \mathcal{K}_n(x) \right]. \quad (5.10)$$

Nous montrons d'abord que les deux termes $\mathcal{I}_n(x)$ et $\mathcal{K}_n(x)$ sont négligeables. Ceci est fait dans Lemme 5.1 et Lemme 5.2. Ensuite, nous déterminons la variance asymptotique qui apparaît dans (5.8) (voir Lemme 5.3). Finalement, nous utilisons la procédure de Bernstein des grands blocs et petits blocs pour obtenir la normalité asymptotique du terme principal $\mathcal{J}_n(x)$ qui est une procédure standard pour les données dépendantes.

Lemme 5.1. *Sous les hypothèses (H), (A), (K1) et (R), nous avons*

$$\sqrt{nh_n^d} |\mathcal{I}_n(x)| = o_{\mathbf{P}}(1) \text{ quand } n \rightarrow \infty.$$

Preuve. Observons que

$$\mathcal{I}_n(x) = \frac{\tilde{v}_n(x) [\hat{\psi}_n(x) - \tilde{\psi}_n(x)] - \tilde{\psi}_n(x) [\hat{v}_n(x) - \tilde{v}_n(x)]}{\hat{v}_n(x) \tilde{v}_n(x)}.$$

Ainsi, en suivant la preuve du Lemme A.4 de Guessoum et Hamrani (2016) (Lemma 4.4, Chapitre 4), nous obtenons $|\hat{\psi}_n(x) - \tilde{\psi}_n(x)| = O_{\mathbf{P}} \left[\left(\frac{\log \log n}{n} \right)^{\theta} \right]$ et $|\hat{v}_n(x) - \tilde{v}_n(x)| = O_{\mathbf{P}} \left[\left(\frac{\log \log n}{n} \right)^{\theta} \right]$ avec $0 < \theta < \frac{\gamma}{2\gamma+6+3\kappa/2}$ pour tout réel $\kappa > 0$.

Finalement, en utilisant l'hypothèse (H)(ii), nous obtenons

$$\sqrt{nh_n^d} |\mathcal{I}_n(x)| = \sqrt{nh_n^d} \cdot O_{\mathbf{P}} \left[\left(\frac{\log \log n}{n} \right)^{\theta} \right] = o_{\mathbf{P}}(1).$$

□

Lemme 5.2. *Sous les hypothèses (H), (K1)-(K2) et (D1)-(D2), nous avons*

$$\sqrt{nh_n^d} |\mathcal{K}_n(x)| = o_{\mathbf{P}}(1) \text{ quand } n \rightarrow \infty.$$

Preuve. Rappelons que

$$\mathcal{K}_n(x) = \left[\left(\mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right) - \frac{\psi(x)}{v(x)} \left(\mathbf{E}(\tilde{v}_n(x)) - v(x) \right) \right] \times \frac{1}{v(x)}$$

Ainsi, en suivant la preuve du Lemme A.5 de Guessoum et Hamrani (2016) (Lemma 4.5, Chapitre 4), nous obtenons sous les hypothèses (K1)-(K2) et (D1)

$$\left| \mathbf{E}(\tilde{\psi}_n(x)) - \psi(x) \right| = O(h_n^2)$$

et sous les hypothèses **(K1)**-**(K2)** et **(D2)**

$$|\mathbf{E}(\tilde{v}_n(x)) - v(x)| = O(h_n^2).$$

Ensuite, en utilisant l'hypothèse **(H)(i)**, nous obtenons

$$\sqrt{nh_n^d} |\mathcal{K}_n(x)| = \sqrt{nh_n^{d+4}} = o_{\mathbf{P}}(1).$$

□

Lemme 5.3. *Sous les hypothèses **(H)**, **(K1)**, **(R)** et **(D3)**, nous avons*

$$nh_n^d \text{Var}(\mathcal{J}_n(x)) \rightarrow \sigma^2(x) \quad \text{quand } n \rightarrow \infty.$$

Preuve. Premièrement, définissons les v.a.'s suivantes

$$\begin{aligned} Z_{i,\delta} &:= Z_{i,\delta}(X_i, Y_i; x) = \frac{\alpha Y_i^\delta}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) - \mathbf{E} \left(\frac{\alpha Y_i^\delta}{G(Y_i)} K_d \left(\frac{x - X_i}{h_n} \right) \right), \delta = 0, 1, \\ U_i &:= Z_{i,1} - \frac{\psi(x)}{v(x)} Z_{i,0}. \end{aligned}$$

D'où, le terme $\mathcal{J}_n(x)$ peut être écrit comme suit

$$\mathcal{J}_n(x) = \frac{1}{nh_n^d v(x)} \sum_{i=1}^n \left(Z_{i,1} - \frac{\psi(x)}{v(x)} Z_{i,0} \right) = \frac{1}{nh_n^d v(x)} \sum_{i=1}^n U_i. \quad (5.11)$$

Alors,

$$nh_n^d \text{Var}(\mathcal{J}_n(x)) = \frac{1}{v^2(x)} \left\{ \frac{1}{h_n^d} \text{Var}(U_1) + \frac{1}{nh_n^d} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov}(U_i, U_j) \right\}.$$

Tandis que

$$\text{Var}(U_1) = \text{Var}(Z_{1,1}) + \frac{\psi^2(x)}{v^2(x)} \text{Var}(Z_{1,0}) - 2 \frac{\psi(x)}{v(x)} \text{Cov}(Z_{1,1}, Z_{1,0}),$$

et

$$\text{Cov}(U_i, U_j) = \text{Cov}(Z_{i,1}, Z_{j,1}) + \frac{\psi^2(x)}{v^2(x)} \text{Cov}(Z_{i,0}, Z_{j,0}) - 2 \frac{\psi(x)}{v(x)} \text{Cov}(Z_{i,1}, Z_{j,0}).$$

Donc,

$$\begin{aligned} nh_n^d \text{Var}(\mathcal{J}_n(x)) &= \frac{1}{v^2(x)} \left\{ \frac{1}{h_n^d} \text{Var}(Z_{1,1}) + \frac{1}{nh_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,1}) \right. \\ &\quad + \frac{\psi^2(x)}{v^2(x)} \left[\frac{1}{h_n^d} \text{Var}(Z_{1,0}) + \frac{1}{nh_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,0}, Z_{j,0}) \right] \\ &\quad \left. - 2 \frac{\psi(x)}{v(x)} \left[\frac{1}{h_n^d} \text{Cov}(Z_{1,1}, Z_{1,0}) + \frac{1}{nh_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,0}) \right] \right\}. \quad (5.12) \end{aligned}$$

En utilisant (5.3) et un changement de variables, le théorème de la convergence dominée donne

$$\frac{1}{h_n^d} \text{Var}(Z_{1,1}) \xrightarrow{n \rightarrow +\infty} \alpha \psi_0(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \quad \frac{1}{h_n^d} \text{Var}(Z_{1,0}) \xrightarrow{n \rightarrow +\infty} \alpha \psi_2(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \quad (5.13)$$

$$\frac{1}{h_n^d} \text{Cov}(Z_{1,1}, Z_{1,0}) \xrightarrow{n \rightarrow +\infty} \alpha \psi_1(x) \int_{\mathbb{R}^d} K_d^2(t) dt, \quad (5.14)$$

et sous les hypothèses **(K1)** et **(D3)**, nous avons

$$| \text{Cov}(Z_{i,1}, Z_{j,1}) | \leq C h_n^{2d}, \quad | \text{Cov}(Z_{i,0}, Z_{j,0}) | \leq C h_n^{2d} \quad \text{et} \quad | \text{Cov}(Z_{i,1}, Z_{j,0}) | \leq C h_n^{2d}. \quad (5.15)$$

Pour évaluer chaque une des sommes données dans (5.12), nous introduisons une suite de nombres entiers ϑ_n qui tend vers l'infini quand $n \rightarrow \infty$. Nous commençons par la première. Pour cela, nous dévisons la somme comme suit

$$\frac{1}{n h_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,1}) = \frac{1}{n h_n^d} \sum_{1 \leq |i-j| \leq \vartheta_n} \text{Cov}(Z_{i,1}, Z_{j,1}) + \frac{1}{n h_n^d} \sum_{\vartheta_n < |i-j| \leq n} \text{Cov}(Z_{i,1}, Z_{j,1}),$$

où $\vartheta_n = o(n)$.

À partir du premier terme de (5.15) et pour $|i-j| \leq \vartheta_n$, nous avons

$$\sum_{1 \leq |i-j| \leq \vartheta_n} \text{Cov}(Z_{i,1}, Z_{j,1}) \leq C n h_n^{2d} \vartheta_n.$$

Pour $\vartheta_n < |i-j| \leq n$, en utilisant le Théorème 5.3 de Bulinski et Shashkin (2007) (Théorème 2.1, Chapitre 2) sous l'hypothèse **(R)**, nous obtenons

$$\begin{aligned} \sum_{\vartheta_n < |i-j| \leq n} \text{Cov}(Z_{i,1}, Z_{j,1}) &\leq C h_n^{-2} \sum_{\vartheta_n < |i-j| \leq n} \rho(|i-j|) \\ &\leq C h_n^{-2} \sum_{\vartheta_n < |i-j| \leq n} e^{-\gamma|i-j|} \\ &\leq C n h_n^{-2} \int_{\vartheta_n}^n e^{-\gamma u} du \\ &\leq C n h_n^{-2} e^{-\gamma \vartheta_n}. \end{aligned}$$

Ainsi,

$$\frac{1}{n h_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,1}) \leq C \left[h_n^d \vartheta_n + \frac{e^{-\gamma \vartheta_n}}{h_n^{d+2}} \right]. \quad (5.16)$$

Maintenant, en choisissant $\vartheta_n = O(h_n^{\nu-d})$ avec $0 < \nu < d$ dans (5.16), nous déduisons que

$$\frac{1}{n h_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,1}) \xrightarrow{n \rightarrow +\infty} 0. \quad (5.17)$$

D'une manière similaire, nous montrons aussi que

$$\frac{1}{n h_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,0}, Z_{j,0}) \xrightarrow{n \rightarrow +\infty} 0 \quad \text{et} \quad \frac{1}{n h_n^d} \sum_{1 \leq i, j \leq n, i \neq j} \text{Cov}(Z_{i,1}, Z_{j,0}) \xrightarrow{n \rightarrow +\infty} 0. \quad (5.18)$$

Finalement, (5.13), (5.14), (5.17) et (5.18) complète la preuve du Lemme 5.3. \square

Maintenant, à partir du Lemme 5.1, Lemme 5.2, Lemme 5.3 et (5.10), nous déduisons que pour montrer le Théorème 5.1, il suffit de montrer que

$$\sqrt{nh_n^d} \mathcal{J}_n(x) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)). \quad (5.19)$$

Pour faire cela, comme déjà mentionné ci-dessus, nous utilisons la procédure de Bernstein des grands blocs et petits blocs qui consiste à partager l'ensemble $\{1, 2, \dots, n\}$ à k grands p -blocs et k petits q -blocs, notés respectivement par I_m et $J_m, m = 1, \dots, k$, comme suit

$$\begin{aligned} I_m &= \{i; i = (m-1)(p+q) + 1, \dots, (m-1)(p+q) + p\}, \\ J_m &= \{j; j = (m-1)(p+q) + p + 1, \dots, m(p+q)\}, \end{aligned}$$

les points restants forme l'ensemble $\{l; k(p+q) + 1 \leq l \leq n\}$ qui peut être vide et p, q , et k sont données dans l'hypothèse **(B)**.

Remarque 5.4. *L'ensemble des points restants est l'ensemble des points non inclus dans les grands ou petits blocs ainsi sa taille $n - k(p+q) \leq q$ et $n - k(p+q) \leq p$.*

Soit

$$W_i := \frac{1}{\sqrt{nh_n^d v^2(x)}} U_i$$

et

$$S_n := \sum_{i=1}^n W_i.$$

Puis à partir de (5.11), il est clair que (5.19) est équivalente à

$$S_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)). \quad (5.20)$$

Pour $m = 1, \dots, k$, nous définissons les v.a.'s η_m, η'_m et η''_k comme suit

$$\eta_m := \sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} W_i, \quad \eta'_m := \sum_{j=(m-1)(p+q)+p+1}^{m(p+q)} W_j, \quad \eta''_k := \sum_{l=k(p+q)+1}^n W_l,$$

Ensuite, nous définissons les sommes partielles

$$T_n := \sum_{m=1}^k \eta_m, \quad T'_n := \sum_{m=1}^k \eta'_m, \quad T''_n := \eta''_k.$$

Ainsi, nous avons

$$S_n = T_n + T'_n + T''_n.$$

Alors pour établir la convergence (5.20), il suffit de prouver que

$$T'_n + T''_n \xrightarrow[n \rightarrow +\infty]{\mathbf{P}} 0, \quad (5.21)$$

$$T_n \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)). \quad (5.22)$$

Pour montrer (5.21) et (5.22), nous avons besoin des quatre lemmes suivants.

Lemme 5.4. *Sous les hypothèses (K1), (R), (D3) et (B), nous avons*

- a) $k \operatorname{Var}(\eta'_1) \xrightarrow{n \rightarrow +\infty} 0,$
- b) $\left| \operatorname{Cov}(\eta'_1, \eta'_{l+1}) \right| \leq \frac{qC}{nv^2(x)} \frac{1}{h_n^{d+2}} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |\theta_{1,r+1}|,$
- c) $\sum_{1 \leq i < j \leq k} \left| \operatorname{Cov}(\eta'_i, \eta'_j) \right| \xrightarrow{n \rightarrow +\infty} 0,$
- d) $\operatorname{Var}(T'_n) \xrightarrow{n \rightarrow +\infty} 0.$

Preuve.

a) Nous avons

$$\begin{aligned}
 k \operatorname{Var}(\eta'_1) &= k \operatorname{Var}\left(\sum_{i=p+1}^{p+q} W_i\right) = \frac{k}{nh_n^d v^2(x)} \operatorname{Var}\left(\sum_{i=p+1}^{p+q} U_i\right) \\
 &= \frac{k}{nh_n^d v^2(x)} \left[q \operatorname{Var}(U_1) + 2 \sum_{1 \leq i < j \leq q} |\operatorname{Cov}(U_i, U_j)| \right] \\
 &= \frac{kq}{nh_n^d v^2(x)} \operatorname{Var}(U_1) + \frac{2k}{nh_n^d v^2(x)} \sum_{l=1}^{q-1} (q-l) |\operatorname{Cov}(U_1, U_{l+1})| \quad (\text{par stationnarité}) \\
 &\leq \frac{kq}{nh_n^d v^2(x)} \operatorname{Var}(U_1) + \frac{2kq}{nh_n^d v^2(x)} \sum_{l=1}^{q-1} |\operatorname{Cov}(U_1, U_{l+1})| \\
 &\leq \frac{kq}{n} \frac{1}{v^2(x)} \left\{ \frac{1}{h_n^d} \operatorname{Var}(Z_{1,1}) + \frac{\psi^2(x)}{v^2(x)} \frac{1}{h_n^d} \operatorname{Var}(Z_{1,0}) - 2 \frac{\psi(x)}{v(x)} \frac{1}{h_n^d} \operatorname{Cov}(Z_{1,1}, Z_{1,0}) \right\} + \\
 &\quad \frac{2}{v^2(x)} \frac{kq}{nh_n^d} \sum_{l=1}^{q-1} \left| \operatorname{Cov}(Z_{1,1}, Z_{l+1,1}) + \frac{\psi^2(x)}{v^2(x)} \operatorname{Cov}(Z_{1,0}, Z_{l+1,0}) - 2 \frac{\psi(x)}{v(x)} \operatorname{Cov}(Z_{1,1}, Z_{l+1,0}) \right|.
 \end{aligned}$$

Notons qu'en utilisant (5.13) et (5.14), nous obtenons

$$\frac{1}{v^2(x)} \left\{ \frac{1}{h_n^d} \operatorname{Var}(Z_{1,1}) + \frac{\psi^2(x)}{v^2(x)} \frac{1}{h_n^d} \operatorname{Var}(Z_{1,0}) - 2 \frac{\psi(x)}{v(x)} \frac{1}{h_n^d} \operatorname{Cov}(Z_{1,1}, Z_{1,0}) \right\} \xrightarrow{n \rightarrow +\infty} \sigma^2(x). \quad (5.23)$$

Par conséquent, (5.15) et (5.23) entraînent que

$$k \operatorname{Var}(\eta'_1) \leq C \frac{kq}{n} + C \frac{kq}{n} q h_n^d.$$

et grâce aux hypothèses **B(i)** et **(ii)**, $k \operatorname{Var}(\eta'_1) \rightarrow 0$ quand $n \rightarrow \infty$.

b) Par stationnarité, nous avons

$$\begin{aligned}
 \left| \operatorname{Cov}(\eta'_1, \eta'_{l+1}) \right| &= \left| \sum_{i=p+1}^{p+q} \sum_{j=l(p+q)+p+1}^{(l+1)(p+q)} \operatorname{Cov}(W_i, W_j) \right| \\
 &= \left| \sum_{r=1}^q (q-r+1) \operatorname{Cov}(W_1, W_{l(p+q)+r}) + \sum_{r=1}^{q-1} (q-r) \operatorname{Cov}(W_{r+1}, W_{l(p+q)+1}) \right|
 \end{aligned}$$

Ensuite,

$$\begin{aligned}
\left| Cov\left(\eta'_1, \eta'_{l+1}\right) \right| &= \left| \sum_{r=1}^q (q-r+1) Cov(W_1, W_{l(p+q)+r}) + \sum_{r=1}^{q-1} (q-r) Cov(W_1, W_{l(p+q)-r+1}) \right| \\
&\leq q \sum_{r=1}^q |Cov(W_1, W_{l(p+q)+r})| + q \sum_{r=1}^{q-1} |Cov(W_1, W_{l(p+q)-r+1})| \\
&\leq q \sum_{r=l(p+q)+1}^{l(p+q)+q} |Cov(W_1, W_r)| + q \sum_{r=l(p+q)-(q-2)}^{l(p+q)} |Cov(W_1, W_r)| \\
&\leq q \sum_{r=l(p+q)-(q-2)}^{l(p+q)+q} |Cov(W_1, W_r)| = q \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |Cov(W_1, W_{r+1})| \\
&\leq \frac{q}{nh_n^d v^2(x)} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |Cov(U_1, U_{r+1})| \\
&\leq \frac{q}{nh_n^d v^2(x)} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} \left| Cov(Z_{1,1}, Z_{r+1,1}) + \frac{\psi^2(x)}{v^2(x)} Cov(Z_{1,0}, Z_{r+1,0}) \right. \\
&\quad \left. - 2 \frac{\psi(x)}{v(x)} Cov(Z_{1,1}, Z_{r+1,0}) \right|.
\end{aligned}$$

En utilisant le Théorème 5.3 de Bulinski et Shashkin (2007) (Théorème 2.1, Chapitre 2), nous pouvons montrer que

$$Cov(Z_{1,1}, Z_{r+1,1}) \leq \frac{C_1^2}{h_n^2} \theta_{1,r+1}, \text{ avec } C_1 = \text{Max} \left\{ Lip(K_d) \frac{M_1}{G(a_F)}, \|K_d\|_\infty \left(\frac{1}{G(a_F)} + Lip(G) \frac{M_1}{G^2(a_F)} \right) \right\},$$

$$Cov(Z_{1,0}, Z_{r+1,0}) \leq \frac{C_2^2}{h_n^2} \theta_{1,r+1}, \text{ avec } C_2 = \text{Max} \left\{ Lip(K_d) \frac{1}{G(a_F)}, \|K_d\|_\infty \frac{Lip(G)}{G^2(a_F)} \right\},$$

$$Cov(Z_{1,1}, Z_{r+1,0}) \leq \frac{C_1 C_2}{h_n^2} \theta_{1,r+1}.$$

Ainsi $b)$ est obtenu.

$c)$ En utilisant la stationnarité, nous avons

$$\begin{aligned}
\sum_{1 \leq i < j \leq k} \left| Cov\left(\eta'_i, \eta'_j\right) \right| &= \sum_{l=1}^{k-1} (k-l) \left| Cov\left(\eta'_1, \eta'_{l+1}\right) \right| \\
&\leq k \sum_{l=1}^{k-1} \left| Cov\left(\eta'_1, \eta'_{l+1}\right) \right| \\
&\leq C \frac{qk}{nv^2(x)} \frac{1}{h_n^{d+2}} \sum_{l=1}^{k-1} \sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)} |\theta_{1,r+1}| \quad (\text{en utilisant b)}) \\
&\leq \frac{C}{v^2(x)} \frac{qk}{n} \frac{1}{h_n^{d+2}} \sum_{r=p}^{\infty} |\theta_{1,r+1}|.
\end{aligned}$$

Les hypothèses **(R)**, **(B)(i)** et **(iii)** impliquent que

$$\begin{aligned} \sum_{1 \leq i < j \leq k} \left| \text{Cov}(\eta'_i, \eta'_j) \right| &\leq \frac{C}{v^2(x)} \frac{qk}{n} \frac{\gamma_0}{h_n^{d+2}} \sum_{r=p}^{\infty} e^{-\gamma r} \\ &\leq \frac{C}{v^2(x)} \frac{qk}{n} \frac{\gamma_0}{1 - e^{-\gamma}} \frac{e^{-\gamma p}}{h_n^{d+2}} \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

d) Puisque

$$\text{Var}(T'_n) = \text{Var}\left(\sum_{m=1}^k \eta'_m\right) = k \text{Var}(\eta'_1) + 2 \sum_{1 \leq i < j \leq k} \left| \text{Cov}(\eta'_i, \eta'_j) \right|.$$

Alors, a) et c) donnent le résultat.

□

Lemme 5.5. *Sous les hypothèses **(K1)**, **(D3)** et **(B)**, nous avons*

$$\text{Var}(T''_n) \xrightarrow{n \rightarrow +\infty} 0.$$

Preuve.

$$\begin{aligned} \text{Var}(T''_n) &= \text{Var}(\eta''_k) \\ &= \text{Var}\left(\sum_{l=k(p+q)+1}^n W_l\right) \\ &= (n - k(p+q)) \text{Var}(W_1) + 2 \sum_{k(p+q)+1 \leq i < j \leq n} |\text{Cov}(W_i, W_j)| \\ &= (n - k(p+q)) \text{Var}(W_1) + 2 \sum_{1 \leq i < j \leq n-k(p+q)} |\text{Cov}(W_i, W_j)| \quad (\text{par stationnarité}) \\ &= \frac{n - k(p+q)}{nh_n^d v^2(x)} \text{Var}(U_1) + \frac{2}{nh_n^d v^2(x)} \sum_{1 \leq i < j \leq n-k(p+q)} |\text{Cov}(U_i, U_j)| \\ &\leq \frac{p}{nh_n^d v^2(x)} \text{Var}(U_1) + \frac{2}{nh_n^d v^2(x)} \sum_{1 \leq i < j \leq p} |\text{Cov}(U_i, U_j)| \quad (\text{voir Remarque 5.4}) \\ &\leq \frac{p}{n v^2(x)} \left\{ \frac{1}{h_n^d} \text{Var}(Z_{1,1}) + \frac{\psi^2(x)}{v^2(x)} \frac{1}{h_n^d} \text{Var}(Z_{1,0}) - 2 \frac{\psi(x)}{v(x)} \frac{1}{h_n^d} \text{Cov}(Z_{1,1}, Z_{1,0}) \right\} + \\ &\quad \frac{2}{v^2(x)} \frac{p}{nh_n^d} \sum_{r=1}^{p-1} \left| \text{Cov}(Z_{1,1}, Z_{r+1,1}) + \frac{\psi^2(x)}{v^2(x)} \text{Cov}(Z_{1,0}, Z_{r+1,0}) - 2 \frac{\psi(x)}{v(x)} \text{Cov}(Z_{1,1}, Z_{r+1,0}) \right|. \end{aligned}$$

Grâce à (5.15) et (5.23), nous obtenons

$$\text{Var}(T''_n) \leq C \frac{p}{n} + C \frac{p^2}{nh_n^d} h_n^{2d}.$$

Ainsi, les hypothèses **(B)(i)** et **(ii)** donnent $\text{Var}(T''_n) \xrightarrow{n \rightarrow +\infty} 0$.

□

Lemme 5.6. *Sous les hypothèses (K1), (R), (D3) et (B), nous avons*

- a) $k \text{Var}(\eta_1) \xrightarrow{n \rightarrow +\infty} \sigma^2(x),$
- b) $|\text{Cov}(\eta_1, \eta_{l+1})| \leq \frac{pC}{nv^2(x)} \frac{1}{h_n^{d+2}} \sum_{r=l(p+q)-p}^{l(p+q)+p} |\theta_{1,r+1}|,$
- c) $\sum_{1 \leq i < j \leq k} |\text{Cov}(\eta_i, \eta_j)| \xrightarrow{n \rightarrow +\infty} 0,$
- d) $\text{Var}(T_n) \xrightarrow{n \rightarrow +\infty} \sigma^2(x).$

Preuve.

a) Nous avons

$$\begin{aligned} k \text{Var}(\eta_1) &= k \text{Var}\left(\sum_{i=1}^p W_i\right) \\ &= \frac{k}{nh_n^d v^2(x)} \text{Var}\left(\sum_{i=1}^p U_i\right) \\ &= \frac{k}{nh_n^d v^2(x)} \left[p \text{Var}(U_1) + 2 \sum_{1 \leq i < j \leq p} |\text{Cov}(U_i, U_j)| \right]. \end{aligned}$$

Par des arguments similaires à ceux utilisés dans la preuve du Lemme 5.4 a) et en utilisant les hypothèses (B)(i) et (ii), nous obtenons le résultat.

b) Nous avons

$$\begin{aligned} |\text{Cov}(\eta_1, \eta_{l+1})| &= \left| \sum_{i=1}^p \sum_{j=l(p+q)+1}^{l(p+q)+p} \text{Cov}(W_i, W_j) \right| \\ &= \left| \sum_{r=1}^p (p-r+1) \text{Cov}(W_1, W_{l(p+q)+r}) + \sum_{r=1}^{p-1} (p-r) \text{Cov}(W_{r+1}, W_{l(p+q)+1}) \right| \\ &= \left| \sum_{r=1}^p (p-r+1) \text{Cov}(W_1, W_{l(p+q)+r}) + \sum_{r=1}^{p-1} (p-r) \text{Cov}(W_1, W_{l(p+q)-r+1}) \right| \\ &\leq p \sum_{r=1}^p |\text{Cov}(W_1, W_{l(p+q)+r})| + p \sum_{r=1}^{p-1} |\text{Cov}(W_1, W_{l(p+q)-r+1})| \\ &\leq p \sum_{r=l(p+q)-(p-2)}^{l(p+q)+p} |\text{Cov}(W_1, W_r)| = p \sum_{r=l(p+q)-(p-1)}^{l(p+q)+(p-1)} |\text{Cov}(W_1, W_{r+1})| \\ &\leq \frac{p}{nh_n^d v^2(x)} \sum_{r=l(p+q)-(p-1)}^{l(p+q)+(p-1)} |\text{Cov}(U_1, U_{r+1})|. \end{aligned}$$

Ainsi,

$$\begin{aligned} |Cov(\eta_1, \eta_{l+1})| &\leq \frac{p}{nh_n^d v^2(x)} \sum_{r=l(p+q)-p}^{l(p+q)+p} \left| Cov(Z_{1,1}, Z_{r+1,1}) + \frac{\psi^2(x)}{v^2(x)} Cov(Z_{1,0}, Z_{r+1,0}) \right. \\ &\quad \left. - 2 \frac{\psi(x)}{v(x)} Cov(Z_{1,1}, Z_{r+1,0}) \right|. \end{aligned}$$

En suivant maintenant étape par étape la preuve du Lemme 5.4 b), nous obtenons le résultat.

c) Nous avons

$$\begin{aligned} \sum_{1 \leq i < j \leq k} |Cov(\eta_i, \eta_j)| &= \sum_{l=1}^{k-1} (k-l) |Cov(\eta_1, \eta_{l+1})| \quad (\text{par stationnarité}) \\ &\leq k \sum_{l=1}^{k-1} |Cov(\eta_1, \eta_{l+1})| \\ &\leq C \frac{pk}{nv^2(x)} \frac{1}{h_n^{d+2}} \sum_{l=1}^{k-1} \sum_{r=l(p+q)-p}^{l(p+q)+p} |\theta_{1,r+1}| \quad (\text{en utilisant b)}) \\ &\leq \frac{C}{v^2(x)} \frac{pk}{n} \frac{1}{h_n^{d+2}} \sum_{r=q}^{\infty} |\theta_{1,r+1}| \\ &\leq \frac{C}{v^2(x)} \frac{pk}{n} \frac{\gamma_0}{h_n^{d+2}} \sum_{r=q}^{\infty} e^{-\gamma r} \\ &\leq \frac{C}{v^2(x)} \frac{pk}{n} \frac{\gamma_0}{1 - e^{-\gamma}} \frac{e^{-\gamma q}}{h_n^{d+2}} \xrightarrow{n \rightarrow +\infty} 0, \end{aligned}$$

par les hypothèses **(B)(i)** et **(iii)**.

d)

$$Var(T_n) = Var\left(\sum_{m=1}^k \eta_m\right) = k Var(\eta_1) + 2 \sum_{1 \leq i < j \leq k} |Cov(\eta_i, \eta_j)|.$$

D'après a) et c), $Var(T_n) \rightarrow \sigma^2(x)$ quand $n \rightarrow \infty$.

□

Preuve du Théorème 5.1. Comme nous l'avons indiqué ci dessus, pour établir le résultat donné dans le théorème 5.1, il suffit d'établir (5.21) et (5.22). Selon le Lemme 5.4 d) et le Lemme 5.5, (5.21) est obtenue par l'inégalité de Tchebychev. Ainsi, il reste à montrer que (5.22) et pour cela il suffit de montrer que

- les v.a.'s η_m , $m = 1, \dots, k$ sont asymptotiquement indépendantes, i.e. quand $n \rightarrow \infty$,

$$\left| \mathbf{E} \left(e^{it \sum_{m=1}^k \eta_m} \right) - \prod_{m=1}^k \mathbf{E} \left(e^{it \eta_m} \right) \right| \rightarrow 0,$$

- pour tout $\epsilon > 0$, $k \mathbf{E} \left(\eta_1^2 \mathbf{1}_{\{|\eta_1| \geq \epsilon \sigma(x)\}} \right) \xrightarrow{n \rightarrow +\infty} 0$ (le critère de Lindeberg-Feller pour la normalité asymptotique de T_n sous l'indépendance).

Pour le premier point, nous avons

$$\begin{aligned}
\left| \mathbf{E} \left(e^{it \sum_{m=1}^k \eta_m} \right) - \prod_{m=1}^k \mathbf{E} (e^{it \eta_m}) \right| &= \left| \mathbf{E} \left(e^{it \sum_{m=1}^k \eta_m} \right) - \prod_{m=1}^{k-1} \mathbf{E} (e^{it \eta_m}) \mathbf{E} (e^{it \eta_k}) \right| \\
&\leq \left| \mathbf{E} \left(e^{it \sum_{m=1}^k \eta_m} \right) - \mathbf{E} (e^{it \eta_k}) \mathbf{E} \left(e^{it \sum_{m=1}^{k-1} \eta_m} \right) \right| + \left| \mathbf{E} (e^{it \eta_k}) \mathbf{E} \left(e^{it \sum_{m=1}^{k-1} \eta_m} \right) \right. \\
&\quad \left. - \mathbf{E} (e^{it \eta_k}) \prod_{m=1}^{k-1} \mathbf{E} (e^{it \eta_m}) \right| \\
&\leq \left| \text{Cov} \left(e^{it \sum_{m=1}^{k-1} \eta_m}, e^{it \eta_k} \right) \right| + \left| \mathbf{E} \left(e^{it \sum_{m=1}^{k-1} \eta_m} \right) - \prod_{m=1}^{k-1} \mathbf{E} (e^{it \eta_m}) \right|.
\end{aligned}$$

Par induction, nous avons

$$\begin{aligned}
\left| \mathbf{E} \left(e^{it \sum_{m=1}^k \eta_m} \right) - \prod_{m=1}^k \mathbf{E} (e^{it \eta_m}) \right| &\leq \left| \text{Cov} \left(e^{it \sum_{m=1}^{k-1} \eta_m}, e^{it \eta_k} \right) \right| + \left| \text{Cov} \left(e^{it \sum_{m=1}^{k-2} \eta_m}, e^{it \eta_{k-1}} \right) \right| + \dots + \\
&\quad \left| \text{Cov} (e^{it \eta_1}, e^{it \eta_2}) \right|. \tag{5.24}
\end{aligned}$$

Ensuite, nous appliquons le Lemme 1 de Bulinski (1996) pour chaque terme du second membre de (5.24). Pour cela, nous avons à déterminer les dérivées partielles d'ordre 1 de la fonction $\varphi_m : \mathbb{R}^{p(d+1)} \rightarrow \mathbb{R}$, $m = 1, \dots, k$ définie par

$$\begin{aligned}
&\varphi_m(x_{(m-1)(p+q)+1}, \dots, x_{(m-1)(p+q)+p}, y_{(m-1)(p+q)+1}, \dots, y_{(m-1)(p+q)+p}) = e^{it \eta_m} = \\
&\exp \left\{ it \left(\sum_{l=(m-1)(p+q)+1}^{(m-1)(p+q)+p} \frac{1}{\sqrt{nh_n^d v^2(x)}} \left[Z_{l,1}(x_l, y_l; x) - \frac{\psi(x)}{v(x)} Z_{l,0}(x_l, y_l; x) \right] \right) \right\},
\end{aligned}$$

avec $x_l \in \mathbb{R}^d$ et $y_l \in \mathbb{R}$ pour $l = (m-1)(p+q)+1, \dots, (m-1)(p+q)+p$.

La dérivée partielle d'ordre 1 de la fonction φ_m par rapport à $x_{l,j}$, pour $l = (m-1)(p+q)+1, \dots, (m-1)(p+q)+p$ et $j = 1, \dots, d$ est donnée par

$$\begin{aligned}
\frac{\partial \varphi_m}{\partial x_{l,j}} &= \frac{ite^{it \eta_m}}{\sqrt{nh_n^d v^2(x)}} \frac{\partial}{\partial x_{l,j}} \left[Z_{l,1}(x_l, y_l; x) - \frac{\psi(x)}{v(x)} Z_{l,0}(x_l, y_l; x) \right] \\
&= \frac{ite^{it \eta_m}}{\sqrt{nh_n^d v^2(x)}} \frac{\alpha}{G(y_l)} \left(y_l - \frac{\psi(x)}{v(x)} \right) \frac{\partial}{\partial x_{l,j}} K_d \left(\frac{x - x_l}{h_n} \right).
\end{aligned}$$

La dérivée partielle d'ordre 1 par rapport à y_l , pour $l = (m-1)(p+q)+1, \dots, (m-1)(p+q)+p$ est comme suit

$$\begin{aligned}
\frac{\partial \varphi_m}{\partial y_l} &= \frac{\partial}{\partial y_l} e^{it \eta_m} \\
&= ite^{it \eta_m} \frac{\partial}{\partial y_l} \left(\sum_{i=(m-1)(p+q)+1}^{(m-1)(p+q)+p} \frac{1}{\sqrt{nh_n^d v^2(x)}} \left[Z_{i,1}(x_i, y_i; x) - \frac{\psi(x)}{v(x)} Z_{i,0}(x_i, y_i; x) \right] \right) \\
&= \frac{ite^{it \eta_m}}{\sqrt{nh_n^d v^2(x)}} \frac{\partial}{\partial y_l} \left[Z_{l,1}(x_l, y_l; x) - \frac{\psi(x)}{v(x)} Z_{l,0}(x_l, y_l; x) \right] \\
&= \frac{ite^{it \eta_m} \alpha}{\sqrt{nh_n^d v^2(x)}} \left[\frac{G(y_l) - y_l g(y_l)}{G^2(y_l)} + \frac{\psi(x)}{v(x)} \frac{g(y_l)}{G^2(y_l)} \right] K_d \left(\frac{x - x_l}{h_n} \right).
\end{aligned}$$

Comme g est bornée, alors en utilisant l'hypothèse **(K1)**, nous avons

$$\left| \frac{\partial \varphi}{\partial y_l}(x_l, y_l) \right| \leq C \frac{t}{\sqrt{nh_n^d}} \quad \text{et} \quad \left| \frac{\partial \varphi}{\partial x_l^j}(x_l, y_l) \right| \leq C \frac{t}{\sqrt{nh_n^{d+2}}}.$$

Maintenant, en appliquant le Lemme 1 de Bulinski (1996), nous obtenons

$$|Cov(e^{it\eta_1}, e^{it\eta_2})| \leq \frac{Ct^2}{nh_n^{d+2}} \sum_{i \in I_1} \sum_{j \in I_2} |\theta_{i,j}| \quad (5.25)$$

et en généralisant (5.25), nous obtenons

$$\left| Cov\left(e^{it \sum_{m=1}^{k-1} \eta_m}, e^{it\eta_k}\right) \right| \leq \frac{Ct^2}{nh_n^{d+2}} \sum_{i \in (I_1 \cup I_2 \cup \dots \cup I_{k-1})} \sum_{j \in I_k} |\theta_{i,j}|.$$

Alors,

$$\left| \mathbf{E}\left(e^{it \sum_{m=1}^k \eta_m}\right) - \prod_{m=1}^k \mathbf{E}\left(e^{it\eta_m}\right) \right| \leq \frac{Ct^2}{nh_n^{d+2}} \left[\sum_{i \in I_1} \sum_{j \in I_2} |\theta_{i,j}| + \sum_{i \in (I_1 \cup I_2)} \sum_{j \in I_3} |\theta_{i,j}| + \dots + \sum_{i \in (I_1 \cup I_2 \cup \dots \cup I_{k-1})} \sum_{j \in I_k} |\theta_{i,j}| \right]. \quad (5.26)$$

Par stationnarité, l'inégalité (5.26) devient

$$\left| \mathbf{E}\left(e^{it \sum_{m=1}^k \eta_m}\right) - \prod_{m=1}^k \mathbf{E}\left(e^{it\eta_m}\right) \right| \leq \frac{Ct^2}{nh_n^{d+2}} \left[(k-1) \sum_{i \in I_1} \sum_{j \in I_2} |\theta_{i,j}| + (k-2) \sum_{i \in I_1} \sum_{j \in I_3} |\theta_{i,j}| + \dots + \sum_{i \in I_1} \sum_{j \in I_k} |\theta_{i,j}| \right]. \quad (5.27)$$

Encore une fois, par stationnarité, pour chaque $2 \leq l \leq k$,

$$\begin{aligned} \sum_{i \in I_1} \sum_{j \in I_l} |\theta_{i,j}| &= |\theta_{1,(l-1)(p+q)-(p-2)}| + 2|\theta_{1,(l-1)(p+q)-(p-3)}| + \dots + (p-1)|\theta_{1,(l-1)(p+q)}| \\ &\quad + p|\theta_{1,(l-1)(p+q)+1}| + (p-1)|\theta_{1,(l-1)(p+q)+2}| + \dots + |\theta_{1,(l-1)(p+q)+p}| \\ &\leq p \sum_{j=(l-1)(p+q)-(p-2)}^{(l-1)(p+q)+p} |\theta_{1,j}|. \end{aligned}$$

D'où, l'inégalité (5.27) devient

$$\begin{aligned} \left| \mathbf{E}\left(e^{it \sum_{m=1}^k \eta_m}\right) - \prod_{m=1}^k \mathbf{E}\left(e^{it\eta_m}\right) \right| &\leq \frac{Ct^2}{nh_n^{d+2}} pk \sum_{j=q+1}^{(k-1)(p+q)+p-1} |\theta_{1,j+1}| \\ &\leq Ct^2 \frac{pk}{n} \frac{1}{h_n^{d+2}} \sum_{j=q}^{\infty} |\theta_{1,j+1}| \\ &\leq Ct^2 \frac{pk}{n} \frac{\gamma_0}{h_n^{d+2}} \sum_{j=q}^{\infty} e^{-\gamma j} \\ &\leq Ct^2 \frac{pk}{n} \frac{\gamma_0}{1 - e^{-\gamma}} \frac{e^{-\gamma q}}{h_n^{d+2}} \xrightarrow{n \rightarrow +\infty} 0, \end{aligned}$$

d'après les hypothèses **(B)(i)** et **(iii)**.

Maintenant, il reste à montrer le deuxième point. Pour cela, nous utilisons le fait que $|\eta_1| \leq C \frac{p\|K_d\|_\infty}{G(a_F)\sqrt{nh_n^d v^2(x)}}$ et l'inégalité de Tchebychev pour avoir

$$\begin{aligned} k\mathbf{E} \left(\eta_1^2 \mathbf{1}_{\{|\eta_1| \geq \varepsilon \sigma(x)\}} \right) &\leq C \frac{kp^2 \|K_d\|_\infty^2}{G^2(a_F)nh_n^d v^2(x)} \mathbf{P}(|\eta_1| > \varepsilon \sigma(x)) \\ &\leq C \frac{p^2 \|K_d\|_\infty^2}{G^2(a_F)nh_n^d v^2(x)} \frac{k\text{Var}(\eta_1)}{\varepsilon^2 \sigma^2(x)} \xrightarrow{n \rightarrow +\infty} 0 \quad (\text{par l'hypothèse } \mathbf{(B)(ii)}), \end{aligned}$$

ce qui complète la preuve du théorème. □

Chapitre 6

Conclusion et Perspectives

6.1 Conclusion

Au cours de cette recherche, nous avons réalisé deux travaux portant sur les propriétés asymptotiques de l'estimateur non paramétrique de type noyau de la fonction de régression pour des données tronquées à gauche et présentant une forme de dépendance de type associé.

Le premier concerne la caractérisation de la vitesse de convergence uniforme presque sûre sur un compact de l'estimateur de la fonction de régression introduit par Ould Saïd et Lemdani (2006). Pour cela, nous avons appliqué une inégalité de type Bernstein due à Doukhan et Neumann (2007) pour obtenir une vitesse de convergence presque sûre et nous avons utilisé les techniques de recouvrement d'un ensemble compact par un nombre fini de boules pour obtenir l'uniformité.

Le second travail porte sur la normalité asymptotique du même estimateur. Pour l'établir, nous utilisons la procédure de Bernstein des grands blocs et petits blocs qui est la technique généralement utilisée pour l'obtention des théorèmes centrales limites pour des estimateurs non paramétriques de fonctionnelles pour des données dépendantes. Cette technique nous permet de vérifier la condition de Lindeberg-Feller et d'obtenir la normalité asymptotique.

Nous avons illustré les résultats obtenus dans ces travaux par des simulations.

Ces travaux sont une extension au cas des données associées de ceux de Ould Saïd et Lemdani (2006) dans le cas des données indépendantes et ceux de Liang, Li et Qi (2009) et Liang (2011) pour des données α -mélangeantes sur le comportement asymptotique de l'estimateur à noyau de la fonction de régression sous le modèle aléatoire tronqué à gauche.

6.2 Perspectives

Pour conclure les travaux de cette thèse, nous présentons quelques perspectives de recherche possibles en vue d'améliorer et d'étendre nos résultats.

- Dans nos travaux, nous n'avons traité que des covariables à valeurs dans \mathbb{R}^d , or nous pensons qu'il est possible d'adapter nos résultats au cas de données tronquées à gauche et associées avec une covariable X fonctionnelle. Le cas des données tronquées à gauche et indépendantes a été traité récemment par Derrar, Laksaci et Ould Saïd (2015) pour la régression robuste dont la régression classique est un cas particulier.

- La structure de dépendance considérée dans cette thèse est l'association qui est un cas particulier de la dépendance faible définie par Doukhan et Louhichi (1999). Ainsi, les résultats établis au chapitre 4 et 5 peuvent être étendus au cas de données tronquées à gauche et faiblement dépendantes. Le cas où les données sont complètement observées est traité par Ango Nze, Bühlmann et Doukhan (2002).

- Il est connu que le choix du paramètre de lissage (la fenêtre) en estimation non paramétrique influence beaucoup l'estimation ainsi, un travail sur la recherche de la fenêtre optimale permettra d'améliorer la qualité de l'estimation.

- Suite au résultat concernant la normalité asymptotique établi au chapitre 5, un résultat de type Berry Esseen pour l'estimateur à noyau de la fonction de régression permettrait de quantifier la vitesse avec laquelle s'effectue la convergence vers la loi normale.

Bibliographie

- [1] Akritas, M. G., and LaValley, M. P. (2005), “ A generalized product-limit estimator for truncated data," *J. Nonparametr. Statist.*, 17, 643 – 663.
- [2] Andrew, D.W.K., (1984), “ Non-strong mixing autoregressive processes," *J. Appl. Probab.*, 21, 930 – 934.
- [3] Ango Nze, P., Bühlmann, P., and Doukhan, P. (2002), “ Weak dependence beyond mixing and asymptotics for nonparametric regression," *Ann. Statist.*, 30(2), 397 – 430.
- [4] Antoniodis, A., Grégoire, G., and McKeague, I. W. (1994), “Wavelet methods for curve estimation," *J. Am. Statist. Ass.*, 89, 1340 – 1353.
- [5] Azevedo, C. (2010), “A note on convergence rates in the strong law of large numbers for associated sequences," *Proceedings of the 13th WSEAS, International Conference on Applied Mathematics (MATH'08)*, 98 – 100.
- [6] Bagai, I., and Prakasa Rao, B.L.S. (1995), “Kernel-type density and failure rate estimation for associated sequences," *Ann. Inst.Statist. Math.*, 47, 253 – 266.
- [7] Birkel, T. (1988), “On the convergence rate in the central limit theorem for associated processes," *Ann. Probab.*, 16(4), 1685 – 1698.
- [8] Birkel, T. (1989), “A note on strong law of large numbers for positively dependent random variables," *Statist. Probab. Lett.*, 7, 17 – 20.
- [9] Bosq, D. (1998), “Nonparametric Statistics for Stochastic Processes : Estimation and Prediction (Second Edition)," *Lecture Notes in Statistics, 110, Springer Verlag, New York*.

-
- [10] Bosq, D., Lecoutre, J. P. (1987), "Théorie de l'estimation fonctionnelle," *ECONOMICA, Paris*.
- [11] Bradley, R.C. (2007), " Introduction to strong mixing conditions," *Vol 1-3, Kendrick Press, Utah*.
- [12] Bulinski, A. (1996), " On the convergence rates in the CLT for positively and negatively dependent random fields," *I. A. Ibragimov and A. Yu. Zaitsev (Eds.), Gordon and Breach*, 3 – 14.
- [13] Bulinski A., and Shabanovich, E. (1998), "Asymptotical behaviour of some functionals of positively and negatively dependent random fields(in Russian)," *Fundam. Appl. Math.*, 4, 479 – 492.
- [14] Bulinski, A., and Shashkin, A. (2007), " Limit theorems for associated random fields and related systems," *Vol 10, Advanced series on statistical science & applied probability*.
- [15] Bulinski, A., Suquet, C. (2001), "Normal approximation for quasi-associated random fields," *Statist. Probab. Lett.*, 54, 215 – 226.
- [16] Burton, R. M., Dabrowski, A. R., and Dehling, H. (1986), "An invariance principle for weakly associated random vectors," *Stochastic Process. Appl.*, 23(2), 301 – 306.
- [17] Carbonez, A., Györfi, L., and Van der Meulin, E.C. (1995), " Partitioning-estimate of a regression function under random censoring," *Statist. & Decisions*, 13, 21 – 37.
- [18] Carrasco, M., Chernov M., Florens J. and Ghysels E. (2007), "Efficient Estimation of General dynamic Models with a Continuum of Moment Conditions," *Journal of Econometrics*, 140, 529 – 573.
- [19] Cai, Z., and Roussas, G.G. (1998), "Kaplan-Meier estimator under association," *Journal. Multivariate Anal.*, 67, 318 – 348.
- [20] Chao, M.T. (1987), "Influence curves for randomly truncated data," *Biometrika*, 74, 426 – 429.
- [21] Chaubey, Y.P., Dewan, I., and Li, J. (2011), "Smooth estimation of survival and density functions for a stationary associated process using Poisson weights," *Statistics and Probability Letters*, 81, 267 – 276.

-
- [22] Chen, K., Chao, M.T., and Lo, S.W. (1995), "On strong uniform consistency of the Lynden-Bell estimator for truncated data," *Ann. Statist.*, 23, 440 – 449.
- [23] Collomb, G. (1981), "Estimation non-paramétrique de la régression ," *Revue Bibliographique. Internat. Statist. Rev.*, 49, 75 – 93.
- [24] Collomb, G. (1983), "Méthodes non paramétriques en régression, analyse de séries temporelles, prédiction et discrimination," *Doctorat d'état, Toulouse 3*.
- [25] Collomb, G. (1984), "Propriétés de convergence presque complète du prédicteur à noyau," *Z. W. Gebiete*, 66, 441 – 460.
- [26] Collomb, G. (1985), "Nonparametric regression : an up to date bibliography Statistics. 16, 309 – 324.
- [27] Cox, J. T., and Grimmett, G. (1984), "Central limit theorem for associated random variables and the percolation model," *Ann. Probab.*, 12, 514 – 528.
- [28] Dabrowski, A. R. (1985), "A functional law of the iterated logarithm for associated sequences," *Statist. Probab. Lett.*, 3, 209 – 212.
- [29] Derrar, S., Laksaci, A., and Ould Saïd, E. (2015), "On the nonparametric estimation of the functional regression for a random left-truncation model," *Journal of statistical Theory and Practice*, 9, 823 – 849.
- [30] Devroye, L. (1978), "The uniform convergence of the Nadaraya-Watson regression function estimate," *Canad. J. Statist.*, 6, No.2, 179 – 191.
- [31] Devroye, L., Györfi, L., Krzyzak, A., and Lugosi, G. (1994), "On the strong universal consistency of nearest neighbor regression function estimates," *Ann. of Statist.*, 22, 1371 – 1385.
- [32] Devroye, L., Györfi, L., and Lugosi, G. (1996), " A probabilist theory of pattern Recognition," *Springer Verlag*.

-
- [33] Donoho, D., Johnstone, I. M., Kerkycharian, G., and Picard, D. (1995), "Wavelet shrinkage : Asymptopia? " *Journal of the Royal Statistical Society*, Series B, 57, 301 – 369.
- [34] Douge, L. (2007), "Vitesses de convergence dans la loi forte des grands nombres et dans l'estimation de la densité pour des variables aléatoires associées," *C. R. Acad. Sci. Paris*, Ser. I 344, 515 – 518.
- [35] Doukhan, P. (1994), "Mixing : Properties and Examples," *vol. 85. Springer-Verlag, New York*.
- [36] Doukhan P. and Louhichi S. (1999), "A new weak dependence condition and applications to moment inequalities", *Stochastic Processes and their Applications*, 84, 313 – 342.
- [37] Doukhan, P. and Neumann, M. (2007), "Probability and moment inequalities for sums of weakly dependent random variables, with applications," *Stochastic Processes and their Applications*, 117, 878 – 903.
- [38] Esary, J., Proschan, F., and Walkup, D. (1967), "Association of random variables with applications," *Ann. Math. Stat.*, 38, 1466 – 1476.
- [39] Eubank, R. L. (1999), "Nonparametric Regression and Spline Smoothing," *Marcel Dekker, New York* .
- [40] Ferrani, Y., Ould Saïd, E., and Tatachak, A. (2016), " On kernel density and mode estimates for associated and censored data," *Com. Statist. Theory & Meth.*, 45, 1853 – 1862.
- [41] Fortuin, C., Kasteleyn, P., and Ginibre, J. (1971), "Correlation inequalities on some partially ordered sets," *Comm. Math. Phys.*, 22, 89 – 103.
- [42] Guessoum, Z., and Hamrani, F. (2016), "Convergence rate of the kernel regression estimator for associated and truncated data," *Journal of Nonparametric Statistics*, DOI : 10.1080/10485252.2017.1303
- [43] Guessoum, Z., and Ould Saïd, E. (2008), "On the nonparametric estimation of the regression function under censorship model," *Statist. & Decisions*, 26, 159 – 177.

- [44] Guessoum, Z., and Ould Saïd, E. (2010), “ Kernel regression uniform rate estimation for censored data under α -mixing condition," *Electronic J. Stat.*, 4, 117 – 132.
- [45] Guessoum, Z., and Ould Saïd, E. (2012), “ Central limit theorem for the kernel estimator of the regression function for censored time series," *J. of Nonparam. Stat.*, 24, 379 – 397.
- [46] Guessoum, Z., Ould Saïd, E., Sadki, O., and Tatachak, A. (2012), “A note on the Lynden-Bell estimator under association," *Statist. Probab. lett.*, 82, 1994 – 2000.
- [47] Gill, R. D. (1994), “Lectures on survival analysis," *In Lectures on Probability Theory*, Bernard, P., editor, Springer-Verlag, 115 – 241.
- [48] Gross, S.T., and Lai, T.L. (1996), “Nonparametric estimation regression analysis with left truncated and censored data," *J. Amer. Statist. Assoc.*, 91, 1166 – 1180.
- [49] Györfi, L., Härdle, W., Sarda, P. and Vieu, P. (1989), “Nonparametric curve estimation for time series," *Lecture Notes in Statistics 60*, Springer- Verlag.
- [50] Györfi, L., Kohler, M., Walk, H. (1998). Weak and strong universal consistency semi-recursive partitioning and kernel regression estimates," *Statistics and Decisions*, 16, 1 – 18.
- [51] He, S., and Yang, G. (1994), “ Estimating a lifetime distribution under different sampling plan," *Statist. Decision theory and related topics*, 5, 73 – 85.
- [52] He, S., and Yang, G. (1998), “Estimation of the truncation probability in the random truncation model," *The Annals of Statistics*, 26, 1011 – 1027.
- [53] He, S., and Yang, G. (2003), “Estimation of regression parameters with left truncated data," *J.Statist. Plann. Inference*, 117, 99 – 122.
- [54] Ioannides, D. A., and Roussas, G. G. (1999), “Exponential inequality for associated random variables," *Statist. Probab. Lett.*, 42(4), 423 – 431.
- [55] Joag-Dev, K., and Proschan, F. (1983), “Negative association of random variables, with applications," *Ann. Statist.*, 11(1), 286 – 295.

-
- [56] Kalbfleisch J.D., and Lawless J.F (1989), "Inference based on retrospective ascertainment : an analysis of the data on transfusion related AIDS," *J. Amer. Stat. Assoc.*, 84, 360 – 372.
- [57] Kaplan, E.M. and Meier, P. (1958), "Nonparametric estimation from incomplete observations," *J. Amer. Statist. Assoc.*, 53, 457 – 481.
- [58] Keiding, N. and Gill, R.D. (1990), "Random truncation models and Markov processes," *Ann. Statist.*, 18, 582 – 602.
- [59] Klein J.P. and Moeschberger M.L. (2003). "Survival Analysis : Techniques for Censored and Truncated Data," *Springer, second edition*.
- [60] Kohler, M., Máthé, K., and Pintér, M. (2002), "Prediction from randomly right censored data," *J. Multivariate Anal.*, 80, 73 – 100.
- [61] Lee, M.T., Rachev, S.T., and Samorodnitsky, G. (1990), "Association of stable random variables," *Ann. Probab.*, 18, 387 – 397.
- [62] Lehmann, E. L. (1966), "Some concepts of dependence," *Ann. Math. Statist.*, 37, 1137 – 1153.
- [63] Liebscher, E. (2001), " Estimation of the density and the regression function under mixing conditions," *Statist. Decisions*, 19(1), 9 – 26.
- [64] Lemdani, M., and Ould Saïd, E. (2007), "Asymptotic Behavior of the Hazard Rate Kernel Estimator Under Truncated and Censored Data," *Communications in Statistics - Theory and Methods*, 36, Issue 1, 155 – 173.
- [65] Lemdani, M., Ould Saïd, E., and Poulin, N. (2009). "Asymptotic properties of a conditional quantile estimator with randomly truncated data," *Journal of Multivariate Analysis*, 100, 546 – 559.
- [66] Liang, H.Y., Li, D. L., and Qi, Y.C. (2009), "Strong convergence in nonparametric regression with truncated dependent data," *J. Multivariate Anal.*, 100, 162 – 174.

-
- [67] Liang, H.Y. (2011), "Asymptotic normality for regression function estimate under truncation and α -mixing conditions," *C. Statistics-Theory and Methods*, 40, 1999 – 2021.
 - [68] Lynden-Bell, D. (1971), "A method of allowing for known observational selection in small samples applied to 3CR quasars," *Monthly Notices Royal Astronomy Society*, 155, 95 – 118.
 - [69] Marshall, A.W., Olkin, I. (1967), "A multivariate exponential distribution," *J. Amer. Statist. Assoc.*, 62, 30 – 44.
 - [70] Masry, E. (1986), "Recursive probability density estimation for weakly dependent processes," *IEEE Trans.Inform.Theor.*, 32, 254 – 267.
 - [71] Masry, E. (2002), "Multivariate probability density estimation for associated processes : strong consistency and rates," *Statist. Probab. Lett.*, 58, 205 – 219.
 - [72] Menni, N., and Tatachak, A. (2016), "A note on estimating the conditional expectation under censoring and association : strong uniform consistency," *Stat. Papers*, DOI 10.1007/s00362-016-0801-8.
 - [73] Nadaraya, E.A. (1964), "On estimating regression," *Theor. Probab. Appl.*, 9, 141 – 142.
 - [74] Nagaraj, N.K. and Reddy, C.S. (1993), "Asymptotic null distributions of tests for change in level in correlated data," *Sankhya : The Indian Jour. of Stat.*, A, 55, 37 – 48.
 - [75] Newman, C. M. and Wright, A. L. (1981), "An invariance principle for certain dependent sequences," *Ann. Probab.*, 9(4), 671 – 675.
 - [76] Newman, C. M. (1984), "Asymptotic independence and limit theorems for positively and negatively dependent random variables," *Inequalities in statistics and probability (Lincoln, Neb., 1982)*, *IMS Lecture Notes Monogr. Ser.*, 5, *Inst. Math. Statist. (Hayward, CA)*, 127 – 140.
 - [77] Oliveira, P.E. (2005), "An exponential inequality for associated variables," *Stat. Probab. Lett.*, 73, 189 – 197.
 - [78] Oliveira, P.E. (2012), "Asymptotics for associated random variables," *Springer-Verlag Berlin Heidelberg*.

-
- [79] Ould Saïd, E. (2006), "A strong uniform convergence rate of kernel conditional quantile estimator under random censorship," *Statist. Prob. Lett.*, 76, 579 – 586.
 - [80] Ould Saïd, E. and Lemdani, M. (2006), "Asymptotic properties of a nonparametric regression function estimator with randomly truncated data," *Ann.Inst.Statist.Math.*, 58, 357 – 378.
 - [81] Ould Saïd, and Sadki, O., (2008), "Prediction via the conditional quantile for right censorship model," *Far East J. Theo. Statist.*, 25(2), 145 – 179.
 - [82] Ould Saïd, E., and Tatachak, A. (2007), "Asymptotic properties of the kernel estimator of the conditional mode for the left truncated model," *C. R. Math. Acad. Sci. Paris*, 344(10), 651 – 656.
 - [83] Ould Saïd, E., and Tatachak, A. (2009), "Strong consistency rate for the kernel mode under strong mixing hypothesis and left truncation," *Com.Statist.Theory & Meth.*, 38, 1154 – 1169.
 - [84] Prakasa Rao, B.L.S. (1993), "Bernstein-type inequality for associated sequences," *In : Ghosh, J., Mitra, S., Parthasarathy, K., Prakasa Rao, B.L.S. (eds.) Statistics and Probability : A Raghu Raj Bahadur Festschrift, Wiley Eastern, New Delhi*, 499 – 509.
 - [85] Peterson, A. V. (1977), "Expressing the Kaplan-Meier estimator as a function of empirical sub survival functions," *J. Amer. Statist. Assoc.*, 72, 854 – 858.
 - [86] Pham, T.D., and Tran, L.T., (1985), "Some mixing properties of time series models," *Stoch. Process. Appl.*, 19, 297 – 303.
 - [87] Pitt, L.D. (1982), "Positively correlated normal variables are associated," *Ann. Probab.*, 10, 496 – 499.
 - [88] Pollard, D. (1984), "Convergence of Stochastic Processes," *Springer*.
 - [89] Rao. C. R. (1965), "Linear Statistical Inference and Its Applications," *Wiley*.
 - [90] Rio, E. (2000), "Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants (In French)," *Mathématiques et Applications*, 31, Springer- Verlag, New York.

-
- [91] Rosenblatt, M. (1956), "A central limit theorem and a strong mixing condition," *Proc. Nat. Ac. Sc. U.S.A.*, 42, 43 – 47.
 - [92] Roussas, G. G. (1991), "Kernel estimates under association : strong uniform consistency," *Statist. Probab. Lett.*, 12, 393 – 403.
 - [93] Roussas, G.G. (2000), "Asymptotic normality of the kernel estimate of a probability density function under association," *Statist.Probab.Lett.*, 50, 1 – 12.
 - [94] Khardani, S., Lemdani, M., and Ould Saïd, E. (2010) "Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship," *J. Korean Statist. Soc.*, 39(4), 455 – 469.
 - [95] Sarda, P., Vieu, P. (2000), "Kernel regression, " *In : M. Schimek (ed.) Smoothing and regression; Approaches, Computation, and Application. Wiley Series in Probability and Statistics, Wiley, New York.*
 - [96] Shao, J. (2003), "Mathmetical Statistics, " *2nd ed. New York, Springer.*
 - [97] Shashkin, A. P (2002), " Quasi-associatedness of a gaussian system of random vectors," *Russ. Math. Surv.* 57, 1243 – 1244.
 - [98] Schumaker, L. (1981), "Spline Functions : Basic Theory," *Wiley, New York.*
 - [99] Schimek, M. (2000), "Smoothing and Regression : Approaches, computation, and application," *Ed. M.G. Schimek, Wiley Series in Probability and Statistics.*
 - [100] Silverman, B.W. (1986), "Density estimation for statistics and data analysis," *Monographs on Statist. and Appl. Probab., Chapman and Hill.*
 - [101] Stone, C. J. (1977), "Consistent nonparametric regression," *Ann. Statist.*, 5, 595 – 645.
 - [102] Stone, C. (1980), "Optimal rates of convergence for nonparametric estimators," *Ann. Statist.* 8, No.6, 1348 – 1360.

-
- [103] Stone, C. (1982). "Optimal global rates of convergence for nonparametric regression," *Ann. Statist.* 10, 1040 – 1053.
- [104] Stute, W. (1982), "Law of logarithm for kernel density estimators," *Ann. Probab.*, 10, 414 – 422.
- [105] Stute, W. (1993), "Almost sure representation of the product-limit estimator for truncated data," *Ann. Statist.*, 21, 146 – 156.
- [106] Stute, W., and Wang, J.L. (1993), "The strong law under random censorship," *The Annals of statistics*, 21(3), 1591 – 1607.
- [107] Tsai, W.Y., Jewell, N.P., and Wang, M.C. (1987), "A note on the product-limit estimator under right censoring and left truncation," *Biometrika*, 74, 883 – 886.
- [108] Vapnik, V. N., and Chervonenkis, A. Y. (1971), "On the uniform convergence of relative frequencies of events to their probabilities," *Theory of Probability and its Applications*, 16, 264 – 280.
- [109] Vieu, P. (1991), "Quadratic errors for nonparametric estimates under dependence," *J. Multivariate Anal.*, 39(2), 324 – 347.
- [110] Walk, H. (1997), "Strong universal consistency of kernel and partitioning regression estimates," *Mathematisches Institut A. Preprint 97-1, Universität Stittgart*.
- [111] Walk, H. (2002), "Almost sure convergence of Nadaraya- Watson regression estimates," *In M. Dror, P. l'Ecuyer F. Szidarovsky (Eds.) Assays on uncertanily-S. Yakowitz memorial volume*, 201 – 223.
- [112] Wang, M.C., Jewell, N.P., and Tsai, W.Y.(1986), "Asymptotic properties of the product-limit estimate under random truncation," *Ann.Statist.*, 14, 1597 – 1605.
- [113] Watson, G.S. (1964), "Smooth regression analysis," *Sankhya Series A*, 26, 359 – 372.

- [114] Woodroffe, M. (1985), "Estimating a distribution function with truncated data," *Ann. Statist.*, 13, 163 – 177.
- [115] Yu, H. (1986), "The law of the iterated logarithm for associated random variables," *Acta Math. Sinica*, 29, 507 – 511 (in Chinese).
- [116] Yu, H. (1996), "A strong invariance principle for associated random variables," *Ann. Probab.*, 24, 2079 – 2097.
- [117] Zhou, Y. (1996), "A note on the TJW product-limit estimator for truncated and censored data," *Statis. Probab. Lett.*, 26, 381 – 387.